研究課題名：生体高分子タンパク質の X 線結晶構造解析

研究代表者：宍倉　文夫（日本大学医学部自然科学系）
研究従事者：宍倉　文夫（日本大学医学部自然科学系）
桑田　隆生（日本大学理工学部 PDF）
共同研究者：高木　尚（東北大学大学院生命科学研究科）
奥　忠武（日本大学生物資源学部農芸化学科）
杉田博昭（筑波大学生物科学系）
鈴木知彦（高知大学理学部生化学教室）
後藤寿夫（徳島大学総合科学部生命科学教室）
石川紘一（日本大学医学部薬理学教室）

【研究目的】

日本大学量子科学研究所は平成 12 年度ー平成 16 年度の文部科学省の学術フロンティアに選定された。同研究所電子線利用研究施設に生体高分子 X 線結晶構造解析実験室を立ち上げ，学内並び学外の共同利用研究者と結晶作製から構造解析までの技術を共有化する。そのために次の 3 つのプロセスを可及的すみやかに達成することを第 1 目標とする。

1　標準タンパク質としてリゾチームを指標に，導入した機器類を検証する。同時に，X 線結晶解析技術を習得する。
2　ゾウガメのヘモグロビン（Hb）を新規タンパク質に，結晶化から構造の精密化並び構造モデルを決定する。
3　論文を公表する。

1〜3 を達成した上で，さらに，X 線結晶構造解析研究者に対する研究支援体制を軌道にのせ，共同研究並び利用実験施設として共同活動を展開する。

また，同研究所電子線利用研究施設で学術フロンティア推進事業として開発する高輝度単波長光源のパラメトリック X 線光源の高度利用研究に連係して，パラメトリック X 線を利用した X 線結晶構造解析の技術を開発し，その技術移転をはかる。

【研究概要】

平成 12 年度から平成 16 年度の研究概要は次の通りである。

1　研究準備
　平成 12 年 4 月〜平成 13 年 3 月　建屋の建設。平成 13 年 3 月竣工式
　平成 13 年 6 月 X 線結晶構造解析実験室の設備完了。
　研究支援者（桑田隆生博士）雇用。
2　リゾチームの結晶化並び構造解析（導入した機器類の検証と結晶構造解析技術の習得）
　平成 13 年 6 月〜平成 14 年 5 月（終了）
3　新規タンパク質ゾウガメの Hb D の結晶構造解析
　平成 14 年 5 月〜平成 15 年 8 月（終了）
4　学外共同研究アカムシの Hb D の結晶構造解析
　東北大学大学院 高木尚教授との共同研究

平成 15 年 11 月ー平成 16 年 3 月（終了）

5　学外共同研究の継続

高知大学理学部 鈴木知彦教授との共同研究（シロウリガイの Hb の構造研究）

静岡大学理学部 竹内浩昭助教授との共同研究（アホロートルの Hb の研究）

筑波大学生物科学系 杉田博昭教授との共同研究（カブトガニのヘモグロビンの構造研究）

６　施設利用研究の開始

麻布大学環境保健学部 佐俣哲郎教授の X 線構造解析実験

筑波大学生物科学系 沼田治教授の X 線構造解析実験

【まとめ】

文部科学省学術フロンティア推進事業の研究年度（平成 12 年度ー平成 16 年度）における研究成果は次の通りである。

1　日本大学量子科学研究所電子線利用研究施設に生体高分子 X 線結晶構造解析実験室を構築した。

2　X 線結晶構造解析実験を実施できるように同上実験室を整備した。その結果，結晶作製から構造の精密化までの作業を実施できる。

3　X 線結晶構造解析実験を終了し，立体構造のモデルを決定したタンパク質には，ゾウガメの Hb D，アカムシの Hb V と Hb VII（投稿準備中）がある。また，シロウリガイの 2 種類の Hb，アホロートルの 2 種類の Hb，カブトガニのヘモシアニンなどの結晶化を進めている。

4　共同研究及び施設利用研究が開始されて，成果をあげている。

研究目的として掲げた計画は, パラメトリック X 線を利用した X 線結晶構造解析の技術を開発すること（現在開発中）以外すべて達成した。現在，共同研究と施設利用研究を展開している。

【研究業績等】

・発表論文

(1) F. Shishikura and K. Takami (2001) The amino acid sequences of the α- and β-globin chains of hemoglobin from the Aldabra giant tortoises, *Geochelone gigantea*. Zoological Science **18**, 515-526.

(2) F. Shishikura (2002) The primary structure of hemoglobin D from the Aldabra giant tortoise, *Geochelone gigantea*. *Zoological Science* **19**, 197-206.

(3) 宍倉文夫 (2002) 生体高分子立体構造解析実験室の立ち上げ. *日医雑誌* **61**, 179.

(4) 宍倉文夫 (2002) 爬虫類のヘモグロビン：ゾウガメ（*Geochelone gigantea* と *Geochelone nigra*）グロビンの分子進化. *日医雑誌* **61**, 263 – 276.

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

(6) F. Shishikura (2004) Leech hemoglobin: primary structures of four kinds of globins from *Haemadipsa zeylanica* var. *japonica*. *Micron* **35**, 123-125.

(7) F. Shishikura (2004). The complete amino acid sequences of four globins from the land leech *Haemadipsa zeylanica* var. *japonica*. *Comparative Biochemistry and Physiology* **137B**, 443-454.

・図書

(1) 宍倉文夫，桑沢清明（2002）血液・ヘモリンパ液／無脊椎動物　生物学データ大百科事典（石原勝敏　編集代表）朝倉書店，pp. 1019 – 1093，東京

・学会発表

(1) 桑田隆生，宍倉文夫 (2002) アルダブラゾウガメ（*Geochelone gigantea*）のヘモグロビン D の結晶化および X 線回折．日本結晶学会平成 14 年度年会講演要旨集，21 頁.

(2) 桑田隆生，宍倉文夫 (2003) アルダブラゾウガメ（*Geochelone gigantea*）のヘモグロビン D の X 線結晶解析．日本結晶学会平成 15 年度年会講演要旨集，122 頁.

(3) F. Shishikura (2003) Leech hemoglobin: Primary structures of four kinds of globins from *Haemadipsa zeylanica* var. *japonica*. XIIIth International Conference on Invertebrate Dioxygen Binding Proteins. Johannes Gutenberg Universität Mainz, Mainz, Germany, September.

(4) T. Kuwada, T. Hasegaw S. Sato, I. Sato, K. Ishikawa, T. Takagi and F. Shishikura (2004) *Tokunagayusurika akamusi* (Diptera) hemoglobin: Crystallographic analyses of Hb V and Hb VII at 1.65 Å and 1.50 Å resolution, respectively. Zoological Science **21**, 1327.

(5) F. Shishikura and H-a, Takeuchi (2004) *Ambystoma mexicanum* (axolotl, Amphibia) hemoglobin: The complete amino acid sequences of two types of α globins and one β globin. Zoological Science **21**, 1327.

・発表論文(奥 忠武・発表分)

(1) Suruga, K., Murakami, K., Taniyama, Y., Hama, T., Chida, H., Satoh, T., Yamada, S., Hakamata, W., Kawachi, R., Nishio, T. and Oku, T., A novel microperoxidase activity: methyl-viologen-linked nitrite reducing activity of microperoxidase, ***Biochemical and Biophysical Research Communications***, **315**, 815-822 (2004)

(2) Ishida, M., Dohmae, N., Shiro, Y., Oku, T., Iizuka, T. and Isogai, Y., Design and synthesis of de novo cytochrome *c*, ***Biochemistry***, **43**(30), 9823-9833 (2004)

(3) Suruga, K., Nagasawa, N., Yamada, S., Satoh, T., Kawachi, R., Nishio, T., Kume, T. and Oku, T., Radiation-induced enhancement of nitrite reducing activity of cytochrome *c*, *Journal of* ***Agricultural and Food Chemistry***, **51**(23), 6835-6843 (2003)

(4) Nishio, T., Hoshino, S., Kondo, A., Ogawa, M., Matsuishi, Y., Kitagawa, M., Kawachi, R. and Oku, T., Glycosidase-catalyzed deoxy oligosaccharide synthesis. Practical synthesis monodeoxy analogs of ethyl β-thioisomaltoside using *Aspergillus niger* α-glucosidase, ***Bioscience Biotechnology Biochemistry***, **67**(5), 1024-1029 (2003)

ZOOLOGICAL SCIENCE **18**: 515–526 (2001)

The Amino Acid Sequences of the α- and β-Globin Chains of Hemoglobin from the Aldabra Giant Tortoises, *Geochelone gigantea*

Fumio Shishikura[1]* and Kazutoshi Takami[2]

[1]*Department of Biology, Nihon University School of Medicine, Oyaguchi-kamimachi, Itabashi-ku, Tokyo 173-8610*
[2]*Osaka Municipal Tennoji Zoo, 1-108, Chausuyama, Tennoji, Osaka 543-0063, Japan*

ABSTRACT—Tetrameric hemoglobins (Hbs) A and D were isolated from red blood cells of the Aldabra giant tortoises, *Geochelone gigantea*, by a hydrophobic interaction chromatography. After reduction and *S*-pyridylethylation, two sets of two types of α-chains (α-1 and α-2) and one β-chain were purified from the major Hb A and minor Hb D in molar ratios of about 1:1:2, respectively, by a reversed-phase column chromatography. The complete amino acid sequences of the three globin-chains from Hb A were determined: 141 amino acid residues for the two α-chains and 146 amino acid residues for the β-chain. Using computer analysis (amino acid maximum homology), the two α-chains shared a 96.5% sequence identity and had low sequence identities (37.8% for α-1 and 35.8% for α-2) with the β-chain of the same species, *G. gigantea*.

We constructed a phylogenetic tree of 28 primary globin structures from *Reptilia* (7 species of squamates, 4 species of turtles, 3 species of crocodiles and 1 species of sphenodontids), including the three globins of *G. gigantea* Hb A. The following results were obtained: (1) The two terrestrial species of *Geochelone* (*G. gigantea* and *G. carbonaria*) were closely related: 139 amino acid residues (95.2%) of the two β-globin chains were conserved; (2) Based on the divergence patterns of globin-chains, the sea turtle *Caretta caretta* was shown to be unusual relatedness form the groups of terrestrial and freshwater species in turtles. The molecular relationships appearing on the phylogenetic tree also support the traditional classification of reptiles and partly confirm previous molecular studies of reptilian hemoglobin evolution.

INTRODUCTION

Hemoglobin, the major respiratory protein, has been extensively investigated in animals, plants, protozoans, fungi and bacteria (Keilin, 1956; Kleinschmidt and Sgouros, 1987). At the molecular level the protein has provided much information in both functional and evolutionary aspects (Bunn and Forget, 1986; Goodman *et al*., 1988; Vinogradov *et al*., 1993). In reptiles, to our knowledge, 35 globin-chains from 19 species have been sequenced (Gorr *et al*., 1998; Kleinschmidt and Sgouros, 1987; Fushitani *et al*., 1996). Among these studies, there are four investigations on turtle hemoglobin: one for a land tortoise, *Geochelone carbonaria* (Bordin *et al*., 1997), one for a sea turtle, *Caretta caretta* (Petruzzelli *et al*., 1996), and two for freshwater turtles, *Chrysemys picta bellii* (Rücknagel *et al*., 1984) and *Phrynops hilarii* (Rücknagel *et al*., 1984).

This study aimed to establish the complete primary structures of both types of globins (α and β) from the land tortoises *Geochelone gigantea*, and analyze phylogenetic relationships among reptiles including two species of *Geochelone*, *G. gigantea* and *G. carbonaria*. First, the two hemoglobin components (Hb A and Hb D) were separated from *G. gigantea* hemoglobin under native conditions and purified constitutive polypeptide globin-chains from each of the two hemoglobin components for use in analyzing their primary structures. This study provides complete primary structures of the three globins from Hb A of *G. gigantea*. Here, we also describe a phylogenetic tree constructed for 14 complete amino acid sequences of both α-type globins and β-type globins of reptilian hemoglobins, including the two α-globins and one β-globin of *G. gigantea* Hb A. The phylogenetic tree supports previous studies on the classification, phylogeny and molecular evolution of reptiles (Benton, 1990, Fushitani *et al*., 1996: Gorr *et al*., 1998).

* Corresponding author: Tel. +81-3-3972-8111;
FAX. +81-3-3972-0027.
E-mail: fshishi@med.nihon-u.ac.jp.

MATERIALS AND METHODS

Materials

Blood from a male Aldabra giant tortoise, *G. gigantea*, weighing approximately 36 kg, was collected in heparin-Tris-HCl buffer, pH 8.0, at Osaka Municipal Tennoji Zoo where the animal died just before

bleeding. The animal came from the Aldabra Atoll located only about 350 km from the northern coast of Madagascar and about 600 km away from East Africa's coast. Acetonitrile, ammonium sulfate, ammonium bicarbonate, *tri-n*-butyl phosphine, *4*-vinyl pyridine and *V8* protease (from *Staphylococcus aureus* strain *V8*) were purchased from Nakalai Tesque, Inc. (Kyoto, Japan). Trifluoroacetic acid (TFA) was obtained from Sigma Chemical Co. (St. Louis, MO. USA). Lysyl endopeptidase (*Achromobactor* protease I) was purchased from Wako Pure Chemicals Co. (Tokyo, Japan). Separation columns, Alkyl Superose column HR5/5 and Resource column (3 ml prepackaged with source 15 RPC gel matrix), were purchased from Pharmacia Biotech (Upsala, Sweden), and placed in a fast protein liquid chromatography (FPLC) system (Pharmacia Biotech). All other chemicals and solvents used were of the most purified grade commercially available.

Preparation of hemoglobin solution

Red blood cells were washed three times in 10 vol. physiological saline, lysed with 1 mM Tris-HCl, pH 8.0, and centrifuged at 3000 × g for 15 minutes to remove cell debris. All procedures were done at 4°C. The cell lysate, hemoglobin solution, was aliquoted and stored at –80°C until use.

Separation of hemoglobin components

The hemoglobin solution which had been saturated at 40% by adding 60% saturated ammonium sulfate was subjected to an Alkyl Superose column equilibrated with 60% saturated ammonium sulfate (183 g/ 500 ml) in 50 mM ammonium bicarbonate, pH 8.0. Elution was carried out with a gradient of 60–0% saturated ammonium sulfate in the 50 mM ammonium bicarbonate buffer. The flow rate was maintained at 0.5 ml/min and fractions of protein peaks were collected. The fractions were monitored at 415 and 280 nm by spectrophotometers (Model 115, Gilson and UV-1, Pharmacia Biotech).

Protein modification

Reduction and *S*-pyridylethylation of globins were performed by the method described previously (Friedman *et al*., 1970). After each reaction, the modified protein was dialyzed against 0.1M ammonium bicarbonate and lyophilized. Finally, the remaining reagents were completely removed from the sample by reversed-phase column chromatography on Resource from water containing 0.1% TFA to 80% acetonitrile containing 0.08% TFA. Flow rates were maintained at 0.5 ml/min. The fractions were monitored at 214 and 280 nm by a spectrophotometer (Model 116, Gilson).

Separation and purification of globin-chains

To separate α-type and β-type globin-chains, the *S*-pyridylethylated hemoglobin was subjected to a Resource column and eluted with a 0.1% TFA buffered gradient to 60% acetonitrile in 0.08% TFA. All fractions were monitored at 214 and 280 nm by a spectrophotometer (Model 116, Gilson). 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).

Enzymatic digestion

Lysyl endopeptidase digestion was performed essentially as described (Jekel *et al*., 1983). Briefly, samples (15–20 nmoles) of the *S*-pyridylethylated protein were first dissolved in 8 M urea and incubated at 37°C for 30 min, followed by the addition of 0.5 M ammonium bicarbonate at a final concentration of 4 M urea in 0.1 M ammonium bicarbonate. Lysyl endopeptidase digestion of the sample was performed at an enzyme/substrate ratio of 1:30 (mol/mol) for 4 hr at 37°C in 0.1 M ammonium bicarbonate solution, pH 8.2 containing 4 M urea. To obtain overlapping peptides, the sample (20 nmoles) was digested with the *V8* protease at a ratio of 1:100 (w/w, enzyme/substrate) for 48 hr at 37°C in a 0.1M Tris-HCl solution, pH 8.5 containing 1 M urea.

Peptide separation

All peptides derived from their parent molecules were separated using a reversed-phase column, Resource, in a 0.1% TFA buffered gradient to 60% acetonitrile in 0.08% TFA. Flow rates were maintained at 0.5 ml/min. All fractions were monitored at 214 nm and 280 nm by a spectrophotometer (Model 116, Gilson). Re-chromatography of selected peptides, when necessary, was performed as previously described (Shishikura *et al*., 1987).

Sequence determination

Sequence analysis was performed using a Shimadzu gas phase protein sequencer, PPSQ-10, equipped with a PTH-10 amino acid analyzer (Shimadzu Co., Kyoto, Japan). Phenylthiohydantoin (PTH)-derivatives from the sequencer were separated and quantified. PTH-cysteine was detected as pyridylethylated-PTH-cysteine, the elution point of which was determined as described in the manufacturer's manual.

Computer analysis

A multiple alignment program, Clustal W (Thompson *et al*., 1994), was used in the alignment of 28 primary structures of globins from *Reptilia*. Pair-wise distances among the 28 globin sequences were analyzed using a computer program PROTDIST stored in the PHYLIP package (v. 3.51c; Felsenstein, 1993) under the Kimura-formula option. Based on the pair-wise distances, Neighbor-Joining/UPGMA in NEIGHBOR (Felsenstein, 1993) was used to construct the phylogenetic tree of hemoglobins from *Reptilia*.

RESULTS AND DISCUSSION

Two components of hemoglobins

The red blood cells of the Aldabra giant tortoises, *G. gigantea*, contain two main hemoglobin components, major and minor, which were successfully separated under native conditions (50 mM ammonium bicarbonate, pH 8.0) by use of an Alkyl Superose column HR 5/5 (Fig. 1). Two peaks were detected at 280 nm and 415 nm. They exist at a ratio of about 5:1 based on chromatogram area calculation. This value may vary from 5:1 to 5:3 depending on sample preparation. Braunitzer and coworkers have succeeded in separating two main components, Hb A and Hb D, from the adult Western Painted Turtle (*Chrysemys picta bellii*) by use of polyacrylamide gel electrophoresis under alkaline (pH 8.3) and dissociating conditions (Rücknagel and Braunitzer, 1988). Also other investigators successfully separated two components from sphenodontid hemoglobin on DEAE at pH 8.5, but in this case they added 0.1% mercaptoethanol and 0.1% dithiothreitol to the elution buffer (Abbasi *et al*, 1988) and lysis buffer of red blood cells (Weber *et al*., 1989), respectively. Brittain (1988) determined the existence of three carbomonoxy-form hemoglobin components, T1, T2 and T3, from *Sphenodon punctatus* by using DEAE-Sephadex eluted by a gradient of Tris-HCl buffer (pH 8.5) *versus* Bistris-HCl buffer (pH 5.5). Bonilla *et al* (1994) used preparative isoelectric focusing and agarose gels with ampholines for separation of intact hemoglobin components. They purified two hemoglobins from the South American snake Riotropical Racer, *Mastigodryas bifossatus*, however, whose protein bands were closely separated with isoelectric points of 8.02 and 8.07, respectively. Hence, the Alkyl Superose column, a kind of hydrophobic interaction col-

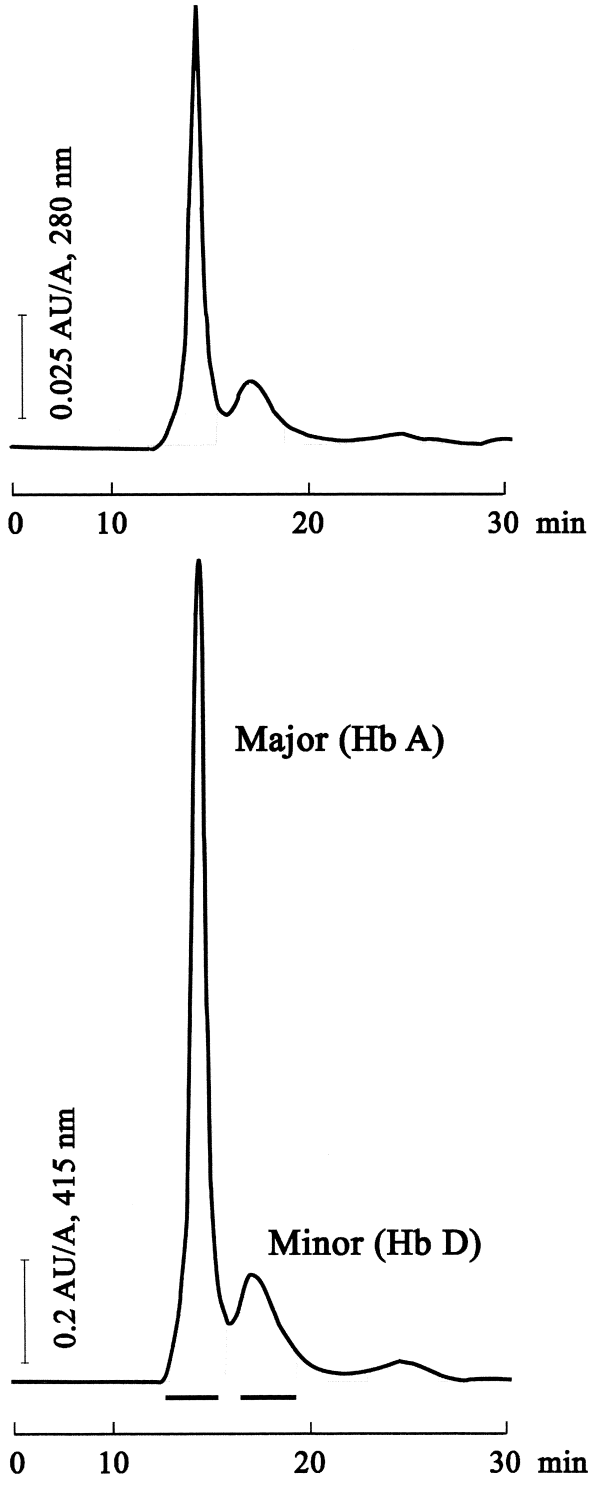


Fig. 1. Alkyl Superose HR5/5 column chromatography of the Aldabra giant tortoises, *G. gigantea*, intact hemoglobin. The hemoglobin solution was saturated at 40% by adding 60% saturated ammonium sulfate and applied to an Alkyl Superose HR5/5 column equilibrated with 60% saturated ammonium sulfate buffer (buffer A). After washing with buffer A, the adsorbed proteins were eluted with linearly decreasing ammonium sulfate concentration from 60% to 0% in 50 mM ammonium bicarbonate, pH 8.0. Elution was monitored at 280 nm (top) and 415 nm (bottom). Flow rate was maintained at 0.5 ml/min. Bars indicate the pooled fractions of major and minor peak, Hb A and Hb D, respectively.

umn, should provide excellent resolution in separation and purification of intact hemoglobin components of *G. gigantea*. Successful separation of intact hemoglobin components enables the manufacture of crystals from individual hemoglobin components for future study of the relationships between physiological functions and crystal structures of hemoglobin components.

Chain separation

After separation of individual components of *Geochelone* hemoglobin, the constituents of the major and minor hemo-

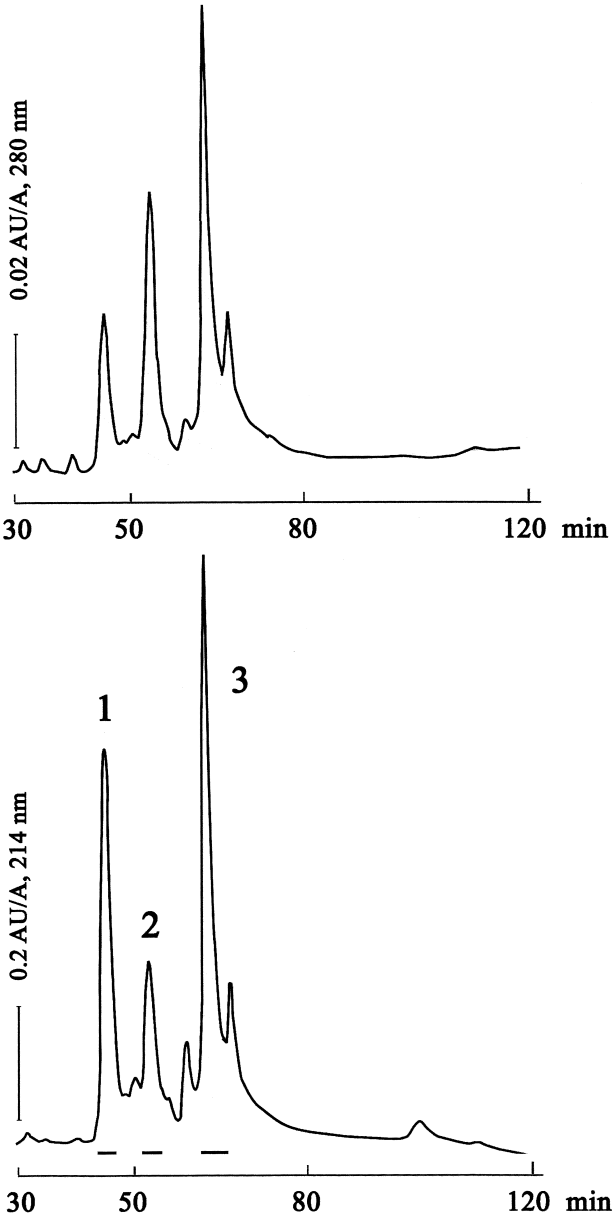


Fig. 2. Separation of globin-chains from reduced and *S*-pyridylethylated Hb A on Resource column. A linear gradient was used between 0.1% TFA in water and 0.08% TFA in 60% acetonitrile at a flow rate of 0.5 ml/min. Elution was monitored at 280 (top) and 214 nm (bottom). Globins from peaks 1, 2, 3 are α-1, α-2 and β globin-chains, respectively. Bars indicate fractions used for sequencing.

globins could have been further separated under intact conditions by a reversed-phase column using prepackaged Resource resins but their resolution proved to be inadequate. However, after reduction and *S*-pyridylethylation of the hemoglobin, the major hemoglobin yielded three more widely separated main peaks and several additional minor peaks on a chromatogram obtained by the same reversed-phase column described above. Many investigators have used reversed-phase semi-microbore type columns (Abbasi *et al*., 1988; Matsuura *et al*., 1989; Fushitani *et al*., 1996; Petruzzelli *et al*., 1996) or CM-columns (Liu, 1975; Leclercq *et al*., 1982; Rücknagel and Braunitzer, 1988; Rücknagel *et al*., 1988; Islam *et al*., 1990: Naqvi *et al*., 1994) for separation of globin-chains. Our method described here has a similar efficiency and gives excellent resolution in separation at high flow rates with low backpressure (*versus* flow rates). Figure 2 shows a typical separation profile of three main peaks of globin-chains from the major hemoglobin. The advantages to modify the protein by reduction and *S*-pyridylethylation were also true for separation of globin-chains from the minor hemoglobin (data not shown). As the results, we prepared six individual globin-chains in total from the two hemoglobin components of *G. gigantea*.

Table 1 shows the results of the first 20 amino-terminal amino acid sequences of the six globin-chains. The nomenclature of α- and β-globin chains depends on amino acid sequence similarities to those of the known sequences of reptilian hemoglobins (Kleinschmidt and Sgouros, 1987). Consequently, the major hemoglobin is a kind of Hb A and the minor is Hb D. Two α-types of globins in Hb A or Hb D have identical sequences each other so far sequenced (Table 1), while their patterns on reversed-phase column chromatography are shown distinctly different (Fig. 2, data not shown for those of Hb D). Figure 2 also shows that two kinds of α-types of globin-chains (α-1 and α-2) and one β-type globin-chain are separated at molar ratio of about 1:1:2. This indicates that the Hb A is tetrameric hemoglobin comprised α-1/α-2/β_2. The presence of two subtypes of α-globins is completely confirmed by their primary structures shown in Fig. 3.

In birds (Dolan *et al*., 1997) and crocodiles (Kleinschmidt and Sgouros, 1987) only one type of β-globin chain has been demonstrated, while lizards and snakes express two types of β-globin chains (Fushitani *et al*., 1996; Gorr *et al*., 1998). It is still unknown yet whether the β-globin chains of Hb A and Hb D from *G. gigantea* are identical or not, but peptide maps digested with lysyl endopeptidase (data not shown) and the first 20 amino-terminal amino acid residues (Table 1) suggest that they might be the same. It was also indicated that the blood of adult Western Painted Turtles, *Chrysemys picta bellii* (Rücknagel and Braunitzer, 1988) were sharing the same β-globin chains when compared with the two complete amino acid sequences of β-types of globin-chains from *Chrysemys* Hb A and Hb D. On the contrary, in frogs there have been reported to present two subtypes of β-globin chains (Knöchel *et al*., 1983; Patient *et al*., 1983; Oberthür *et al* 1983 and 1986). On the numbers of subtypes of β-type globin-chains among amphibians, reptiles, birds and mammals, thus, reinvestigations are needed, in particular, in view of evolution of Tetrapoda.

Table 1. Amino-terminal amino acid residues of six globin chains from the Aldabla giant tortoises, *Geochelone gigantea*

		1									10									20
Hb A	α-1	V	L	T	A	G	D	K	A	N	V	K	T	V	W	S	K	V	G	S	H
	α-2	V	L	T	A	G	D	K	A	N	V	K	T	V	W	S	K	V	G	S	H
	β	V	H	W	T	S	E	E	K	Q	Y	I	T	A	L	Q	W	A	K	V	N
		1									10									20
Hb D	α-1	M	L	T	E	D	D	K	Q	L	I	Q	H	V	W	E	K	V	L	E	H
	α-2	M	L	T	E	D	D	K	Q	L	I	Q	H	V	W	E	K	V	L	E	H
	β	V	H	W	T	S	E	E	K	Q	Y	I	T	A	L	Q	W	A	K	V	N

Sequencing and alignments

In general, Amniota (reptiles, birds, and mammals) has two or more hemoglobin components (Ikehara *et al*., 1997; Gorr *et al*, 1998) which are expressed under different physiological conditions. The presence of α-type (α^D) globin-chain in Hb D is, in particular, of interest in the study of the molecular phylogeny of Amniota because α^D-globin chain was first studied in birds such as chickens (Hagopian and Ingram, 1971; Brown and Ingram, 1974; Kleinshmidt and Sgouros, 1987). The nomenclature of Hb A and Hb D was adopted in Ingram's laboratory (Hagopian and Ingram, 1971; Brown and Ingram, 1974) to describe the various domestic fowl hemoglobins: The embryonic and adult definitive erythrocytes contain the major adult (A) hemoglobin and the minor definitive (D) hemoglobin.

Hb D was also reported in the tuatara *Sphenodon* (Abbasi *et al*., 1988). As for the presence of Hb D in turtles, it was first found in the adult Western Painted Turtle, *Chrysemys picta bellii* (Rücknagel *et al*., 1984) and the Hilaire's Sideneck Turtle, *Phrynops hilarii* (Rücknagel *et al*., 1984). This study describes the presence of Hb D in the Aldabra giant tortoises, *G. gigantea*, and also demonstrates the presence of two subtypes of α-type globin-chains. To ascertain the presence of α^D globin-chains in the Hb D of *G. gigantea*, a study on the primary structures of the α^D globin-chains is in progress (the primary structure of α^D –1 globin chain has been submitted to the JIPID with an accession number PC7116). To date, all reptiles sequenced (Abbasi *et al*., 1988; Matsuura *et al*., 1989; Islam *et al*., 1990; Rücknagel *et al*., 1988; Abbasi and Braunitzer, 1991; Fushitani *et al*., 1996) except crocodiles (Leclercq *et al*., 1981) have been clarified to possess two hemoglobin components, Hb A and Hb D.

Reptilian phylogenetic tree

Geochelone is a unique group among turtles since it includes two big-size tortoises; *G. gigantea*, the Aldabra giant tortoises, and *G. elephantopus*, the Galapagos giant tortoises. These species may weigh up to 250 kg and measure 150 cm over the curve of their carapaces (Jackson, 1984). The Galapagos tortoises are, in general, known as the world's largest living tortoises. Recently, Bordin *et al* (1997) have studied

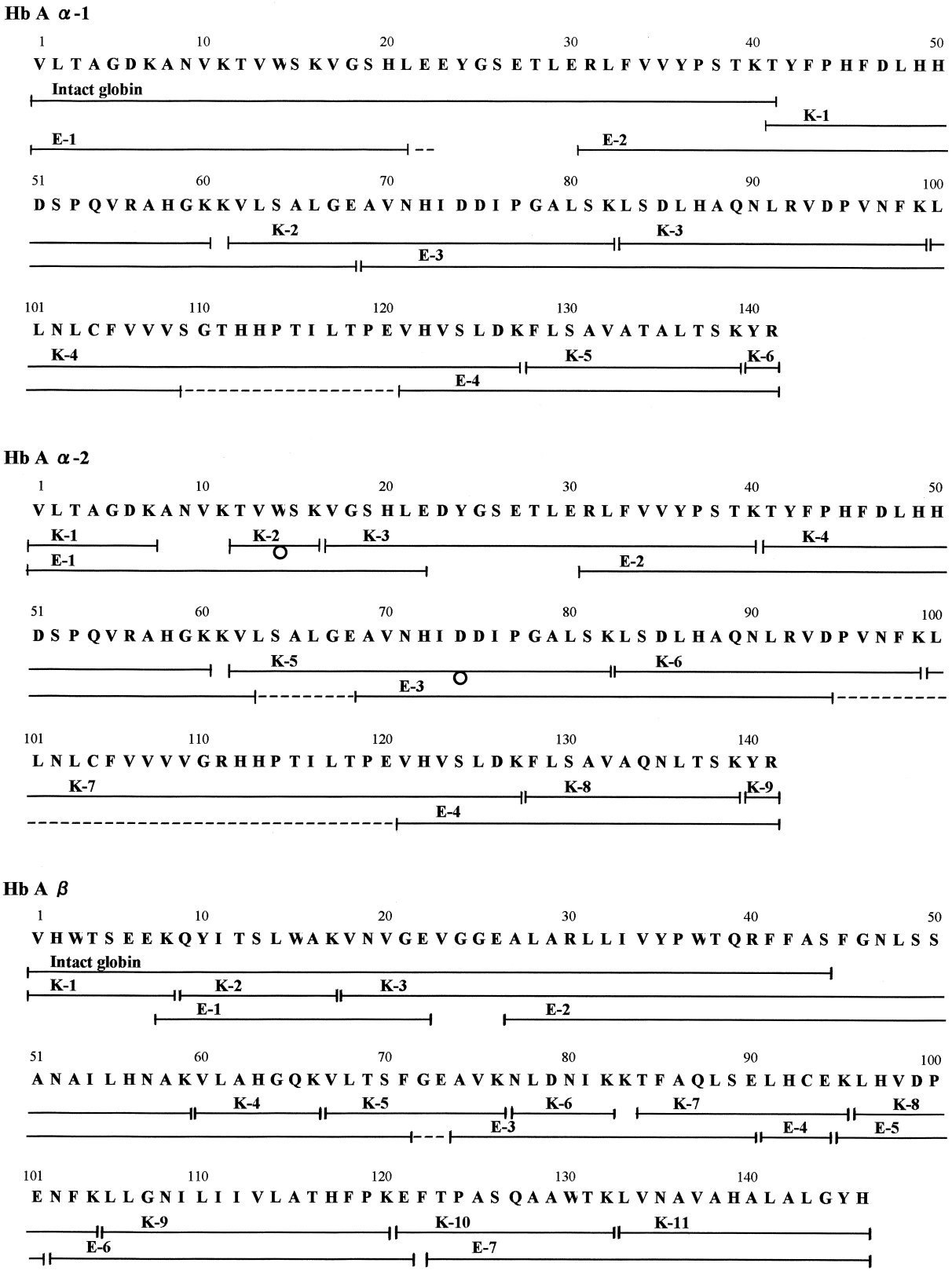


Fig. 3. Strategies and complete sequences of α-1 (top), α-2 (middle) and β (bottom) globin-chains of Hb A from the Aldabra giant tortoises, *G. gigantea*. The complete amino acid sequences of the α-1 (top), α-2 (middle) and β (bottom) globin-chains of *G. gigantea* Hb A are established from overlapping peptides and fragments. The residues marked with continuous lines are those identified by Edman degradation method. Vertical lines represent the beginning and the end of sequencing. Dashed lines indicate the residues not determined but which might be included in the fragment. Small open circles indicate residues incompletely identified using the fragment. Peptide nomenclatures are as follows: lysyl endopeptidase, K; V8 protease, E.

one of the *Geochelone* species, *G. carbonaria*, whose carapace size measures at most about 40 cm, and reported the primary structure of the β-globin deduced from its cDNA analysis. The distribution range of *G. carbonaria* stretches throughout mainland South America: Panama, Colombia, Venezuela, Brazil, Paraguay, and Argentina. Thus, the present habitats of the three *Geochelone* species are remote and isolated from one another. Many questions arises, such as "When did they diversify from their ancestor?", "Is there any correlation between diversity of morphological characteristics and evolution of protein structure?", "How do they differ in their primary structures?", and "How have they adapted protein functions to environmental and physical circumstances? " On these points, comparing the primary structures gives fairly important clues for understanding and elucidating the evolution and improvements of molecular structures of proteins as well as genes.

Figure 3 summarizes the strategies used to establish the complete amino acid sequences of the three globin-chains from Hb A. Appendix 1 provides the data supporting the sequences of Fig. 3. The two α-globin chains are composed of 141 amino acid residues and the β-chain is composed of 146 residues. All overlaps were quantitatively confirmed by duplicated analyses of amino acid residues, with the exceptions of residues 22 through 30 and 109 through 120 for α-1 globin-chain, 23 through 30, 61, 63 through 68, and 95 through 120 for α-2 globin chain, and 72, 73 and 122 for β-globin chain.

The two α-globin chains and the β-globin chain of *G. gigantea* are aligned with those previously reported for reptilian hemoglobins (Appendix 2). When the globin sequences of the two α-chains and the β-chain of *G. gigantea* are compared with those of known sequences, there are 19 invariant amino acids among the 28 globins from reptilian hemoglobins. As for the invariant amino acid residues among the 14

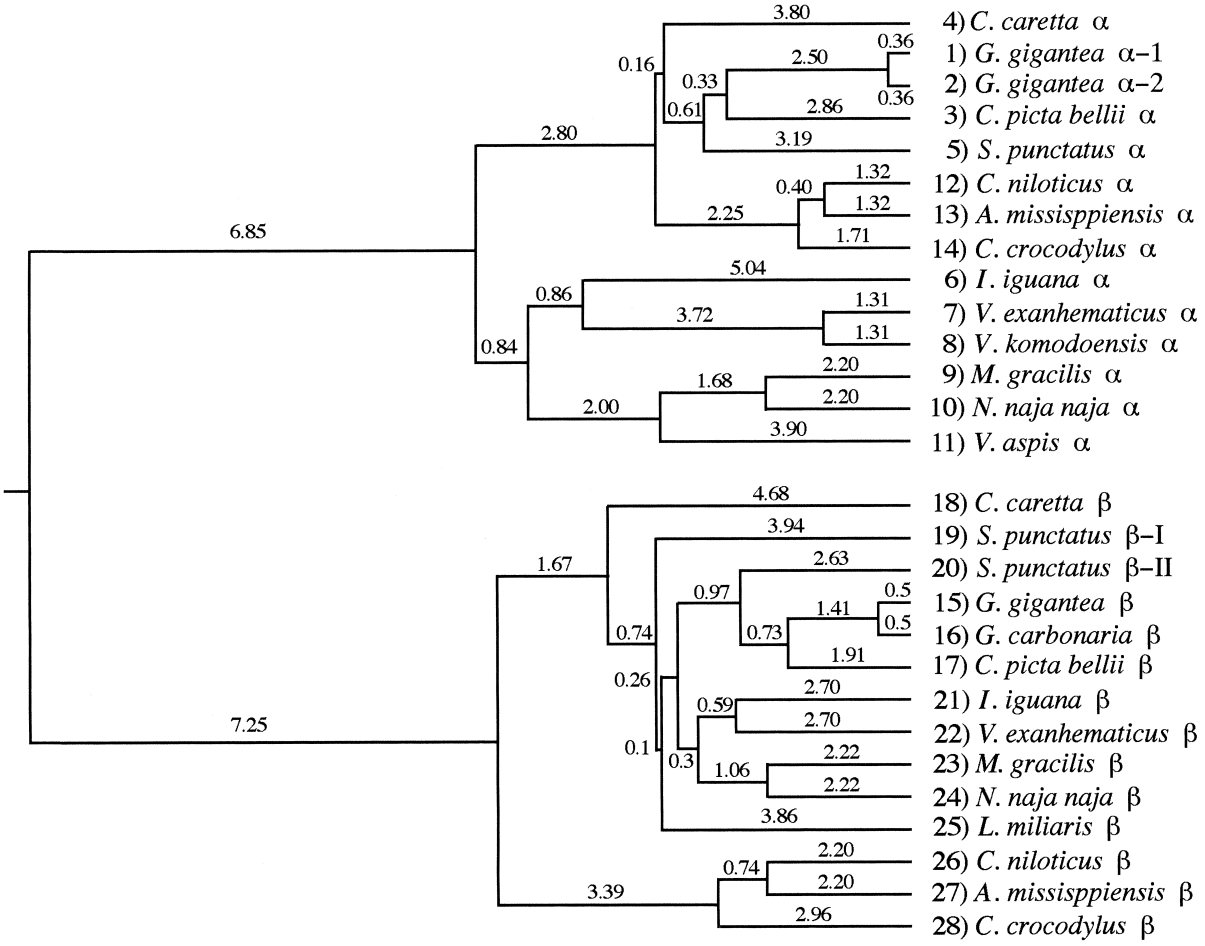


Fig. 4. The phylogenetic tree of 28 reptilian globins. Distance matrices estimated by PRTODIST (option: Kimura formula/data not shown) was used for construction of a rooted tree by NEIGHBOR under the UPGMA method in the package of PHYLIP (version 3.51c: Felsenstain, 1993). Branch lengths are proportional to protein distances and shown on the individual branches of the tree. The abscissa is a time scale in Myr (million years) ago based on the separations of the α- and β- globin chains described by Goodman *et al.*, 1975. The references of globin-chains used in the present analysis are as follows: 1) this study, 2) this study, 3) Rücknagel *et al.*, 1988, 4) Petruzzelli *et al.*, 1996, 5) Abbasi *et al.*, 1988, 6) Rücknagel *et al.*, 1988, 7) Abbasi *et al.*, 1988, 8) Fushitani *et al.*, 1996, 9) Islam *et al.*, 1990, 10) Naqvi *et al.*, 1994, 11) Duguet *et al.*, 1974, 12) Leclercq *et al.*, 1981, 13) Leclercq *et al.*, 1981, 14) Leclercq *et al.*, 1982, 15) this study, 16) Bordin *et al.*, 1997, 17) Rücknagel *et al.*, 198, 18) Petruzzelli *et al.*, 1996, 19) Abbasi *et al.*, 1988, 20) Brittain, 1988, 21) Rücknagel *et al.*, 1988, 22) Abbasi *et al.*, 1991, 23) Islam *et al.*, 1990, 24) Naqvi *et al.*, 1994, 25) Matsuura *et al.*, 1989, 26) Leclercq *et al.*, 1981, 27) Leclercq *et al.*, 1981, 28) Leclercq *et al.*, 1982.

α-globins and 14 β-globins from reptiles, there are 44 invariant (31.2%) and 41 invariant (28.1%), respectively. When compared sequence similarities of globin chains within a species (*G. gigantea*), the sequence identities of 37.8% (α-1 *versus* β) and 35.8% (α-2 *versus* β) were obtained. These resemble the value (42.5% identity) obtained from the comparison with human α- and β-globin chains (Bunn and Forget, 1986). On the contrary, comparing the sequence of the β-globin chain of *G. gigantea* with that of *G. carbonaria*, there are 139 identical amino acid residues. This similarity (95.2%) coincides with the sequence similarity (96.5%) of the two α-globin chains of *G. gigantea*. This finding suggests that the two species are definitely very closed related to each other, and their protein structures, though they are limited, have somehow been conserved even when their morphological characteristics have greatly changed.

Reptilian phylogeny and diversity based on α- and β-globin chains

Phylogenetic analyses of 28 globin-chains including 4 species of Testudinata, 7 species of Squamata (snakes and lizards), 3 species of Crocodylia and 1 species of Rhynchocephalia were conducted by PROTDIST (Felsenstein, 1993). The rooted tree (Fig. 4) deduced by NEIGHBOR under the UPGMA method (PHYLIPS; Felsenstein, 1993) is highly correlated at the level of orders with the reptilian traditional phylogeny established mainly depending on morphological characteristics (Carroll, 1969; Benton, 1990). This partly supports the previous molecular studies on the evolution of reptilian hemoglobins (Goodman *et al*., 1975; Fushitani *et al*., 1996; Gorr *et al*., 1998; Vinogradov *et al*., 1993). The molecular relationships appearing on our phylogenetic tree are summarized as follows: (1) the two species of *Geochelone* have separated very recently (estimated to be about 17 million years ago): divergence dates are estimated 2.6–4.4 times later than those of the two species of Varanas and the two species of crocodiles (*Crocodylus niloticus* and *Alligator mississippiensis*); (2) the species Sphenodon is closely related to the group of tortoises; (3) the primary structures of β-globin chains from the sea turtle *Caretta* and the sea snake *Liophis miliaris* hemoglobins were, in particular, shown to be unusual relatedness from the group of terrestrial species in turtles and squamates, respectively; (4) in the branches of α-globin chains the squamates (snakes and lizards) diverged from the groups of turtles and crocodiles, but in the branches of β-globin chains the crocodiles first separated from the other groups of reptiles (turtles and squamates).

One of the most interesting objectives is, therefore, determining when and how the Galapagos giant tortoises, *G. elephantopus*, diversified from their sister species, the Aldabra giant tortoises *G. gigantea*. At the present time, the habitats of the two giant tortoises are remote oceanic islands and separated by two continents, Africa and South America, and the Atlantic Ocean. According to our β-globin data, the divergence time of *G. gigantea* and *G. carbonaria* was estimated as 17 myr (million years) ago. This coincides with the recent study of Caccone *et al* (1999) who have estimated that the colonization of Madagascar by tortoises occurred in 22–14 myr ago based on tortoises mtDNA sequences. It is not unreasonable to make a scenario that the two living giant species had diversified less than 17 myr ago from their common ancestor and reached oceanic islands, one is volacinc (the Galapagos) and the other is atoll (Aldabra), by rafting, the most reliable way of migration for terrestrial animals. Future studies on globin structures of *G. elephantopus* may well explain the divergence times and molecular relationships of hemoglobins among the three *Geochelone* species.

ACKNOWLEDGMENTS

We thank the staff of veterinary surgeons of Animal Hospital of Osaka Municipal Tennoji Zoo for their cooperation in collecting blood of the Aldabra giant tortoises, *G. gigantea*. We also thank Dr. M. Esumi, Department of Pathology, Nihon University School of Medicine, for her generosity in helping to prepare the equipment for analyzing amino acid sequences.

REFERENCES

Abbasi A, Wells RMG, Brittain T, Braunitzer G (1988) Primary structure of the hemoglobins from Sphenodon (*Sphenodon punctatus*, Tuatara, Rhyncocephalia). Biol Chem Hoppe-Seyler 369: 755–764

Abbasi A, Braunitzer G (1991) Primary structure of hemoglobin from Monitor Lizard (*Varanus exanthematicus albigularis*-Squamata). Biol Chem Hoppe-Seyler 372: 473–479

Benton MJ (1990) Phylogeny of the major tetrapod groups: Morphological data and divergence dates. J Mol Evol 30: 409–424

Bonilla GO, Focesi A Jr, Bonaventura C, Bonaventura J, Cashon RE (1994) Functional properties of the hemoglobin from the South American snake *Mastigodryas bifossatus*. Comp Biochem Physiol 109A: 1085–1095

Bordin S, Meza AN, Saad STO, Ogo SH, Costa FF (1997) cDNA-derived amino-acid sequence of a land turtle (*Geochelone carbonaria*) β-chain hemoglobin. Biochem Mol Biol Int 42: 255–260

Brittain T (1988) An investigation of the functioning of the two major haemoglobins of the Sphenodon using fast reaction kinetic methods. Biochem J 251: 771–776

Brown, JL, Ingram, VM (1974) Structural studies on chick embryonic hemoglobins. J Biol Chem 249: 3960–3972

Bunn HF, Forget BG (1986) In Hemoglobin: Molecular, genetic and clinical aspects. Chap 2, Hemoglobin Structure, pp 13–35, Chap 3, Hemoglobin function, pp 37–60, Saunders WB Co, Philadelphia

Caccone A, Amato G, Gratry OC, Behler J, Powell JR (1999) A molecular phylogeny of four endangered Madagascar tortoises based on mtDNA sequences. Mol Phylogenet Evol 12: 1–9

Carroll RL (1969) Origin of reptiles. In; Biology of Reptilia, Gans C *et al*, Ed, Vol 1, p.1–44, Academic Press, London & New York

Dolan M, Dodgson JB, Engel JD (1983) Analysis of the adult chicken β-globin gene. Nucleotide sequence of the locus, microheterogeneity at the 5'-end of β-globin mRNA, and aberrant nuclear RNA species. J Biol Chem 258: 3983–3990

Duguet M, Chauvet J-P, Acher R (1974) Phylogeny of hemoglobins: The complete amino acid sequence of an α-chain of viper (*Vipera aspis*) hemoglobin. FEBS Let 47: 333–337

Felsenstein J (1993) PHYLIP: Phylogeny Inference Package, Version 3.5c. Department of Genetics SK-50, University of Wash-

ington, Seattle, WA 98195, USA
Friedman M, Krull LH, Cavins JF (1970) The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl) cysteine. J Biol Chem 245: 3568–3871
Fushitani K, Higashiyama K, Moriyama EN, Imai K, Hosokawa K (1996) The amino acid sequences of two α chains of hemoglobins from Komodo dragon *Varanus komodoensis* and phylogenetic relationships of amniotes. Mol Biol Evol 13: 1039–1043
Goodman M, Moore GW, Matsuda G (1975) Darwinian evolution in the genealogy haemoglobin. Nature 253: 603–608
Goodman M, Pedwaydon J, Czelusniak J, Suzuki T, Gotoh T, Moens L, Shishikura F, Walz DA, Vinogradov SN (1988) An evolutionary tree for invertebrate globin sequences. J Mol Evol 27: 236–249
Gorr TA, Mable BK, Kleinschmidt T (1998) Phylogenetic analysis of reptilian hemoglobins: Trees, rates, and divergences. J Mol Evol 47: 471–485
Hagopian HK, Ingram VM (1971) Developmental changes of erythropoiesis in cultured chick blastoderms. J Cell Biol 51: 440–451
Ikehara T, Eguchi Y, Kayo S, Takei H (1997) Isolation and sequencing of two α-globin genes α^A and α^D in pigeon and evidence for embryo-specific expression of the α^D-globin gene. Biochem Biophys Res Commun 234: 450–453
Islam A, Persson B, Zaidi ZH, Jörnvall H (1990) Sea snake (*Microcephalophis gracilis*) hemoglobin: Primary structure and relationships to other forms. J Prot Chem 9: 533–541
Jackson MH (1984) Galapagos, a natural history, University of Calgary Press, Alberta, Canada
Jekel PA, Weijer WJ, Beintema JJ (1983) Use of endoproteinase Lys-C from *Lysobacter enzymogenes* in protein sequence analysis. Anal Biochem 134: 347–354
Keilin D (1956) Distribution and diversity of haemoglobin. Acta Biochim Polon 3: 439–457
Kleinschmidt T, Sgouros JG (1987) Hemoglobin sequences. Biol Chem Hoppe-Seyler, 368: 579-615
Knöchel W, Meyerhof W, Hummel S, Grundmann U (1983) Molecular cloning and sequencing of mRNAs coding for minor adult globin polypeptides of *Xenopus laevis*. Nucl Acids Res 11: 1543–1553
Leclercq F, Schnek AG, Braunitzer G, Stangl A, Schrank B (1981) Direct reciprocal allosteric interaction of oxygen and hydorgen carbonate sequence of the haemoglobins of the caiman (*Caiman crocodylus*), the Nile crocodile (*Crocodylus niloticus*) and the Mississippi crocodile (*Alligator mississippiensis*). Hoppe-Seyler's Z Physiol Chem 362; 1151–1158
Leclercq F, Stangl A, Schrank B, Vandecasserie C, Paul C, Schnek AG, Leonis J, Braunitzer G (1982) *Caiman crocodylus* hemoglobin. Complete primary structures of α and β chains; phylogenic and regulatory aspects. Int J Peptide Protein Res 20: 337–350
Liu C-S (1975) Preparation and chemical characterization of the three chains of the major hemoglobin of the sea snake, *Pelamis platurus*. J Biochem 78: 19–29
Matsuura MSA, Fushitani K, Riggs AF (1989) The amino acid sequences of the α and β chains of hemoglobin from the snake, *Liophis miliaris*. J Biol Chem 264: 5515–5521
Naqvi S, Zaidi ZH, von Bahr-Lindström H, Carlquist M, Jörnvall H (1983) Characterization of hemoglobin from the lizard *Uromastix hardwickii*. FEBS Lett 162: 290–295
Naqvi S, Abbasi A, Zaidi ZH (1994) Primary structure of hemoglobin from cobra *Naja naja naja*. J Prot Chem 13: 669–679
Oberthür W, Braunitzer G, Baumann R, Wright PG (1983) Die Primärstruktur der α- und β-Ketten der Hauptokomponenten der Hämoglobine des Straußes (*Struthio camelus*) und des Nandus (*Rhea americana*) (*Struthioformes*). Aspekte zur Atmungsphysiologie und Systematik. Hope-Seyler's Z Physiol Chem 364: 119–134
Oberthür W, Godovac-Zimmermann J, Braunitzer G (1986) The expression of α^D-chains in the hemoglobin of adult ostrich (*Struthio camelus*) and American rhea (*Rhea americana*). The different evolution of adult bird α^A-, α^D- and β-chains. Biol Chem Hoppe-Seyler 367: 507–514
Patient RK, Harris R, Walmsley ME, Williams JG (1983) The complete nucleotide sequence of the major adult β globin gene of *Xenopus laevis*. J Biol Chem 258: 8521–8523
Petruzzelli R, Aureli G, Lania A, Galtieri A, Desideri A, Giardina B (1996) Diving behaviour and haemoglobin function: the primary structure of the α- and β-chains of the sea turtle (*Caretta caretta*) and its functional implications. Biochem J 316: 959–965
Rücknagel KP, Reischl E, Braunitzer G (1984) Hämoglobine der Reptilian. Expression von α^D-Genen bei Schildkröten, *Chrysemys picta bellii* und *Phrynops hilarii* (Testudines). Hoppe-Seyler's Z Physiol Chem 365: 1163–1171
Rücknagel KP, Braunitzer G, Wiesner H (1988) Hemoglobins of reptiles. The primary structures of the α^I- and β^I-chains of common iguana (*Iguana iguana*). Biol Chem Hoppe-Seyler 369: 1143–1150
Rücknagel KP, Braunitzer G (1988) Hemoglobins of reptiles. 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
Shishikura F, Snow JW, Gotoh T, Vinogradov SN, Walz DA (1987) Amino acid sequence of the monomer subunit of the extracellular hemoglobin of *Lumbricus terrestris*. J Biol Chem 262: 3123–3131
Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequences weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acid Res 22: 4673–4680
Vinogradov SN, Walz DA, Pohajdak B, Moens L, Kapp OH, Suzuki T, Trotman, CNA (1993) Adventitious variability? The amino acid sequences of nonvertebrate globins. Comp Biochem Physiol 106B: 1–26
Weber RE, Kleinschmidt T, Abbasi A, Wells RMG, Braunitzer G (1989) Allosteric transition in hemoglobin ($\alpha^A{}_2\beta^I{}_2$) from the Phynchocephalian reptile relict *Sphenodon punctatus*. Hemoglobin 13: 625–636

(Received December 3, 2000 / Accepted January 22, 2001)

Appendix 1. Sequence analyses of peptides obtained by cleavage with lysyl endopeptidase and *V8* protease.

Step	Hb A α-1 Amino Acid (p moles)			Hb A α-2 Amino Acid (p moles)			Hb A β Amino Acid (p moles)			
	Intact globin		**E-1**	**K-1**		**E-1**	**Intact globin**		**K-1**	
1	V (1650)		V (3281)	V (2099)		V (243)	V (576)		V (2689)	
2	L (1610)		L (2685)	L (2733)		L (272)	H (206)		H (991)	
3	T (666)		T (1189)	T (1174)		T (145)	W (264)		W (1952)	
4	A (1520)		A (2077)	A (2233)		A (252)	T (171)		T (814)	
5	G (1008)		G (1504)	G (1715)		G (160)	S (60)		S (241)	
6	D (1006)		D (2096)			D (116)	E (289)		E (1323)	
7	K (1690)		K (2744)			K (142)	E (349)		E (1460)	**E-1**
8	A (1185)		A (1690)			A (191)	K (231)	**K-2**	K (812)	K (1622)
9	N (1092)		N (1598)			N (169)	Q (213)	Q (1643)		Q (1375)
10	V (1048)		V (1414)			V (149)	Y (171)	Y (1231)		Y (1079)
11	K (1438)		K (1720)	**K-2**		K (189)	I (196)	I (1782)		I (1261)
12	T (464)		T (645)	T (1715)		T (100)	T (85)	T (770)		T (1537)
13	V (1006)		V (1092)	V (2333)		V (137)	S (31)	S (242)		S (218)
14	W (556)		W (693)	N.D.		W (60)	L (160)	L (1377)		L (1050)
15	S (147)		S (179)	S (303)		S (24)	W (50)	W (849)		W (516)
16	K (777)		K (961)	K (1431)	**K-3**	K (140)	A (179)	A (1208)		A (945)
17	V (665)		V (652)		V (2182)	V (114)	K (49)	K (1019)	**K-3**	K (780)
18	G (395)		G (437)		G (1589)	G (77)	V (142)		V (1625)	V (660)
19	S (86)		S (98)		S (343)	S (18)	N (123)		N (1397)	N (625)
20	H (183)		H (214)		H (786)	H (60)	V (148)		V (1738)	V (549)
21	L (484)		L (397)		L (1634)	L (62)	G (93)		G (1015)	G (332)
22	E (384)				E (1685)		E (103)		E (1387)	E (258)
23	E (172)				D (1179)		V (126)		V (1146)	
24	Y (266)				Y (1041)		G (83)		G (872)	
25	G (211)				G (842)		G (104)		G (882)	
26	S (48)				S (200)		E (82)		E (991)	**E-2**
27	E (217)				E (1029)		A (131)		A (1032)	A (1119)
28	T (111)				T (412)		L (101)		L (1020)	L (922)
29	L (304)				L (755)		A (130)		A (979)	A (1476)
30	E (183)		**E-2**		E (763)	**E-2**	R (60)		R (944)	R (565)
31	R (152)		R (768)		R (723)	N.D.	L (97)		L (852)	L (780)
32	L (261)		L (2331)		L (567)	L (232)	L (114)		L (900)	L (882)
33	F (157)		F (2060)		F (489)	F (208)	I (61)		I (747)	I (569)
34	V (239)		V (2060)		V (475)	V (232)	V (76)		V (706)	V (600)
35	V (264)		V (2004)		V (496)	V (206)	Y (38)		Y (558)	Y (453)
36	Y (97)		Y (1510)		Y (334)	Y (153)	P (45)		P (505)	P (473)
37	P (138)		P (1670)		P (349)	P (174)	W (11)		W (326)	W (171)
38	S (27)		S (286)		S (56)	S (37)	T (13)		T (249)	T (256)
39	T (57)		T (653)		T (126)	T (94)	Q (45)		Q (373)	Q (383)
40	K (89)	**K-1**	K (1543)	**K-4**	K (231)	K (194)	R (29)		R(437)	R (375)
41		T (763)	T (515)	T (1520)		T (98)	F (51)		F (405)	F (395)
42		Y (1690)	Y (740)	Y (2342)		Y (87)	F (70)		F (466)	F (398)
43		F (2475)	F (965)	F (2797)		F (103)	A (76)		A (343)	A (444)
44		P (1501)	P (679)	P (1796)		P (99)			S (53)	S (48)
45		H (460)	H (375)	H (850)		H (58)			F (270)	F (319)

46		F (1739)	F (724)	F (1967)		F (74)			G (173)	G (164)
47		D (893)	D (619)	D (1810)		D (102)			N (214)	N (184)
48		L (1451)	L (577)	L (1788)		L (72)			L (177)	L (228)
49		H (451)	H (326)	H (851)		H (56)			S (37)	S (34)
50		H (658)	H (380)	H (1066)		H (61)			S (37)	S (29)
51		D (893)	D (411)	D (1283)		D (71)			A (162)	A (169)
52		S (171)	S (64)	S (211)		S (14)			N (110)	N (111)
53		P (674)	P (295)	P (1372)		P (72)			A (144)	A (163)
54		Q (636)	Q (226)	Q (824)		Q (54)			I (82)	I (82)
55		V (668)	V (268)	V (869)		V (59)			L (90)	L (110)
56		R (527)	R (240)	R (936)		N.D.			H (39)	H (43)
57		A (684)	A (258)	A (798)		A (62)			N (83)	N (85)
58		H (238)	H (122)	H (341)		H (36)			A (88)	A (129)
59		G (405)	G (131)	G (421)		G (46)		**K-4**	K (14)	K (90)
60		K (443)	K (167)	K (592)		K (57)		V (2182)		V (68)
61	**K-2**		K (214)		**K-5**	K (60)		L (2201)		L (81)
62	V (3256)		V (121)		V (2749)	V (45)		A (2052		A (113)
63	L (3474)		L (105)		L (2873)			H (969)		H (31)
64	S (489)		S (19)		S (459)			G (1171)		G (42)
65	A (2602)		A (95)		A (2381)			Q (1220)		Q (65)
66	L (2452)		L (62)		L (1972)		**K-5**	K (965)		K (45)
67	G (1693)		G (38)		G (1269)		V (2425)			V (36)
68	E (2120)	**E-3**	E (22)		E (1948)	**E-3**	L (2591)			L (42)
69	A (2070)	A (1907)			A (1654)	A (1931)	T (1186)			T (33)
70	V (1820)	V (1171)			V (1453)	V (214)	S (312)			S (4)
71	N (1489)	N (1100)			N (1381)	N (193)	F (1730)			F (19)
72	H (618)	H (454)			H (606)	H (89)	G (1129)			
73	I (1521)	I (970)			I (1227)	I (105)	E (1487)		**E-3**	
74	D (911)	D (911)			N.D.	D (166)	A(1403)		A (4964)	
75	D (994)	D (968)			D (1046)	D (194)	V (1108)		V (4570)	
76	I (731)	I (791)			I (816)	I (86)	K (901)	**K-6**	K (5937)	
77	P (578)	P (838)			P (667)	P (119)		N (2636)	N (3659)	
78	G (476)	G (628)			G (509)	G (101)		L (2788)	L (3700)	
79	A (635)	A (956)			A (686)	A (118)		D (2238)	D (2842)	
80	L (577)	L (1293)			L (620)	L (118)		N (2286)	N (2972)	
81	S (80)	S (144)			S (90)	S (26)		I (1890)	I (2807)	
82	K (301)	K (1044)	**K-3**	**K-6**	K (480)	K (115)		K (1684)	K (4045)	
83		L (956)	L (2787)	L (3202)		L (96)	**K-7**		K (4113)	
84		S (125)	S (405)	S (351)		S (17)	T (1096)		T (1122)	
85		D (598)	D (1876)	D (1593)		D (66)	F (2400)		F (2071)	
86		L (954)	L (1901)	L (1619)		L (73)	A (2179)		A (1994)	
87		H (262)	H (922)	H (912)		H (39)	Q (1670)		Q (1462)	
88		A (566)	A (1709)	A (1089)		A (70)	L (1911)		L (1253)	
89		Q (298)	Q (1358)	Q (1203)		Q (59)	S (250)		S (173)	
90		N (398)	N (1427)	N (1172)		N (53)	E (1924)		E (297)	**E-4**
91		L (671)	L (1407)	L (1087)		L (56)	L (1254)			L (1461)
92		R (357)	R (1561)	R (1200)		R 68)	H (578)			H (1110)
93		V (634)	V (1052)	V (1003)			C/pe-Cys			C/pe-Cys
94		D (322)	D (999)	D (1344)			E (1037)		**E-5**	E (291)
95		P (368)	P (640)	P (1372)			K (979)	**K-8**	K (672)	
96		V (532)	V (648)	V (553)				L (2331)	L (1064)	

97		N (271)	N (585)	N (469)				H (850)	H (567)	
98		F (300)	F (494)	F (415)				V (2029)	V (1121)	
99	**K-4**	K (482)	K (433)	K (473)	**K-7**			D (1509)	D (710)	
100	L (2512)	L (472)			L (879)			P (1450)	P (441)	
101	L (2298)	L (597)			L (914)			E (1794)	E (597)	**E-6**
102	N (1617)	N (216)			N (758)			N (1350)		N (153)
103	L (1774)	L (498)			L (740)			F (1234)		F (462)
104	C/pe-Cys	C/pe-Cys			C/pe-Cys			K (1129)	**K-9**	K (365)
105	F (1569)	F (212)			F (606)				L (1695)	L (284)
106	V (1478)	V (359)			V (585)				L (1550)	L (259)
107	V (2016)	V (424)			V (692)				G (977)	G (104)
108	V (2133)	V (451)			V (697)				N (999)	N (74)
109	S (241)				V (694)				I (647)	I (80)
110	G (943)				G (360)				L (392)	L (147)
111	T (483)				R (502)				I (387)	I (54)
112	H (417)				H (190)				I (563)	I (71)
113	H (633)				H (257)				V (258)	V (150)
114	P (834)				P (327)				L (252)	L (162)
115	T (457)				T (166)				A (237)	A (121)
116	I (695)				I (244)				T (94)	T (55)
117	L (784)				L (426)				H (106)	H (37)
118	T (337)				T (141)				F (137)	F (43)
119	P (457)				P (176)				P (86)	P (39)
120	E (541)		**E-4**		E (235)	**E-4**	**K -10**		K (63)	K (44)
121	V (513)		V (401)		V (195)	V (242)	E (2396)		**E-7**	E (25)
122	H (205)		H (144)		H (78)	H (70)	F (1684)		N.D.	
123	V (539)		V (371)		V (284)	V (149)	T (962)		T (1215)	
124	S (46)		S (62)		S (17)	S (24)	P (1487)		P (1081)	
125	L (289)		L (266)		L (95)	L (74)	A (1793)		A (1427)	
126	D (194)		D (338)		D (71)	D (105)	S (316)		S (87)	
127	K (206)	**K-5**	K (318)	**K-8**	K (63)	K (125)	Q (1099)		Q (440)	
128		F (842)	F (244)	F (1051)		F (58)	A (1265)		A (681)	
129		L (856)	L (247)	L (1116)		L (83)	A (1201)		A (575)	
130		S (140)	S (51)	S (176)		S (18)	W (797)		W (141)	
131		A (670)	A (326)	A (772)		A (104)	T (407)		T (143)	
132		V (588)	V (166)	V (723)		V (83)	K (490)	**K-11**	K (318)	
133		A (665)	A (208)	A (1088)		A (77)		L (2011)	L (309)	
134		T (255)	T (103)	Q (1203)		Q (61)		V (1763)	V (183)	
135		A (511)	A (153)	N (1172)		N (51)		N (1325)	N (37)	
136		L (1003)	L (216)	L (1087)		L (41)		A (1803)	A (260)	
137		T (191)	T (68)	T (216)		T (62)		V (1596)	V (159)	
138		S (61)	S (28)	S (69)		S (14)		A (1740)	A (281)	
139	**K-6**	K (230)	K (116)	K (377)	**K-9**	K (72)		H (485)	H (60)	
140	Y (2846)		Y (66)		Y (1715)	Y (33)		A (1530)	A (220)	
141	R (698)		R (61)		R (310)	R (39)		L (1522)	L (153)	
142								A (1459)	A (154)	
143	Microheterogeneity:							L (1347)	L (119)	
144	α-1; 104 (L)							G (901)	G (56)	
145	β;115(L), 116(A), 119(F)							Y (873)	Y (50)	
146								H (96)	H (20)	

Appendix 2. Alignment of amino acid sequences of 28 reptilian globins.
Clustal W (Thompson *et al.* 1994), a multiple alignment program, was used. The invariant amino acid residues are indicated by asterisks. The nomenclatures of globin-chains with a numerical order are the same as Fig. 4.

```
                                1         10         20         30         40         50         60
 1) G. gigantea α-1             -VLTAGDKANVKTVWSKVGSHLEE---YGSETLERLFVVYPSTKTYFPHFDLHH------DSPQVR
 2) G. gigantea α-2             -VLTAGDKANVKTVWSKVGSHLED---YGSETLERLFVVYPSTKTYFPHFDLHH------DSPQVR
 3) C. picta bellii α           -VLNAGDKANVKAVWNKVAAHVEE---YGAETLERMFTVYPQTKTYFPHFDLHH------GSAQIR
 4) C. caretta α                -VLSSGDKANVKSVWSKVQGHLED---YGAETLDRMFTVFPQTKTYFSHFDVHH------GSTQIR
 5) S. punctatus α              -MLSASDKANVKAIWSKVCVHAEE---YGAETLERMFTVYPSTKTYFPHFDLTH------GSAQVK
 6) I. iguana α                 -VLTEDDKNHIRAIWGHVDNNPEA---FGVEALTRLFLAYPATKTYFAHFDLNP------GSAQIK
 7) V. exanthematicus α         -VLTEDDKNHVKGLWAHVHDHIDE---IAADALTRMFLAHPASKTYFAHFDLSP------DNAQIK
 8) V. komodoensis α            -VLTEDDKTHVKTLWGHVHNHAEE---IAADALTRMFLAHPTSKTYFAHFDFSP------NSANIK
 9) M. gracilis α               -VLTEEDKARVRVAWVPVSKNAEL---YGAETLTRLFAAHPTTKTYFPHFDLSP------GSNDLK
10) N. naja naja α              -VLTDEDKARVRASWVPVGKNAEL---YGSETLTRMFAAHPTTKTYFPHFDLSP------GSNNLR
11) V. aspis α                  -VLSEDDKNRVRTS---VGKNPELPGEYGSETLTRMFAAHPTTKTYFPHFDLSS------GSPNLK
12) C. niloticus α              -VLSSDDKCNVKAVWSKVAGHLEE---YGAEALERMFCAYPQTKIYFPHFDLSH------GSAQIR
13) A. missippiensis α          -VLSMEDKSNVKAIWGKASGHLEE---YGAEALERMFCAYPQTKIYFPHFDMSH------NSAQIR
14) C. crocodylus α             -VLSEEDKSHVKAIWGKVAGHLEE---YGAESLERMFCAYPQTKIYFPHFDMSH------NSAQIR

15) G. gigantea β               VHWTSEEKQYITSLWAKVNVGEVG-----GEALARLLIVYPWTQRFFASFGNLSSANAILHNAKVL
16) G. carbonaria β             VHWSCEEKQFITSLWAKVNVEEVG-----GEALARLLIVYPWTQRFFSSFGNLSSPNAILHNAKVL
17) C. picta bellii β           VHWTADEKQLITSLWGKVNVEECG-----SEALARLLIVYPWTQRFFSTFGNLSNAEAILHNPHVH
18) C. caretta β                THWTAEERHYITSMWDKINVAEIG-----GESLARMLIVYPWTQKFFSDFGNLTSSSAIMHNVKIQ
19) S. punctatus β-I            VHWTAEEKHLLGSLWAKVDVADIG-----GEALGRLLVVYPWTQRFFADFGNLSSATAICGNPRVK
20) S. punctatus β-II           VHWTAEEKQLVTSLWTKVNVDECG-----GEALGRLLIVYPWTQRFFSSFGNLSSSTAICGNPRVK
21) I. iguana β                 VHWTAEEKQLITQVWGKIDVAQIG-----GETLACLLVVYPWTQRFFPDFGNLSNAAAICGNAKVK
22) V. exanthematicus β         VHWTAEEKQLICSLWGKIDVGLIG-----GETLAGLLVIYPWTQRQFSHFGNLSSPTAIAGNPRVK
23) M. gracilis β               VHWSAEEKQLITGLWGKVDVAEVG-----GATLGKLLVVFPWTQRFFAHFGNLSSANAIICNPVVK
24) N. naja naja β              VHWSAEEKQLITSLWAKVDVPEVG-----AATLGKMMVMYPWTQRFFAHFGNLSGPSALCGNPQVR
25) L. miliaris β               VHWTAEEKSAITAIWGKVDVAAIG-----GEALCRLLIVYPWTQRFFTSFGNLSNAAAIQSNAQVK
26) C. niloticus β              ASFDPHEKQLIGDLWHKVDVAHCG-----GEALSRMLIVYPWKRRYFENFGDISNAQAIMHNEKVQ
27) A. missippiensis β          ASFDAHERKFIVDLWAKVDVAQCG-----ADALSRMLIVYPWKRRYFEHFGKMCNAHDILHNSKVQ
28) C. crocodylus β             SPFSAHEESLIVDLWAKVDVASCG-----GDALSRMLIIYPWKRRYFEHFGKLSTDQDVLHNEKIR
                                                            *         *     *   *

          70         80         90         100        110        120        130        140        150
 1)   AHGKKVLSALGEAVNHIDDIPGALSKLSDLHAQNLRVDPVNFKLLNLCFVVVSGTHHPTILTPEVHVSLDKFLSAVATALTSKYR
 2)   AHGKKVLSALGEAVNHIDDIPGALSKLSDLHAQNLRVDPVNFKLLNLCFVVVVGRHHPTILTPEVHVSLDKFLSAVAQNLTSKYR
 3)   THGKKVLTALGEAVNHIDDLASALSKLDSIHAQTLRVDPVNFKFLNHCFLVVVAIHQPSVLTPEVHVSLDKFLSAVGTVLTSKYR
 4)   SHGKKVMLALGDAVNHIDDIATALSALSDKHAHILRVDPVNFKLLSHCLLVVVARHHPTLFTPDVHVSLDKFMGTVSTVLTSKYR
 5)   AHGKKVVNAMGEAVNHLDDMAGALLKLSDLHAQKLRVDPVNFKLLAQCFLVVLGVHHPAALTPEVHASLDKFLCAVGLVLTAKYR
 6)   AHGKKVVDALTQAVNNLDDIPDALAKLADLHAEKLRVDPVNFGLLGHCILVTIAAHNHGPLKADVALSMDKFLTKVAKTLVAHYR
 7)   AHGKKVANALNQAVAHLDDIKGTLSKLSELHAQQLRVDPVNFGFLRHCLEVSIAAHLHDHLKASVIVSLDKFLEEVCKDLVSKYR
 8)   AHGKKVANALNQAVNHLDDIGGTLSKLSDLHAQQLRVDPVNFGFLRHCLEVSIAANLHDHLKASLIVSLDKFLEEVCKVLVSKYR
 9)   VHGKKVIDALTEAVNNLDDVAGALSKLSDLHAQKLRVDPDNFQFLGLCLEVTIAAHSGGPLKPEVLLSVDKFLGQISKVLASRYR
10)   AHGKKVIDAITEAVNNLDDVAGTLSKLSDLHAQKLRVDPVNFKLLAHCLLVTIAAHNGGVLKPEVIVSLDKFLGDLSKDLVSKYR
11)   AHGKKVIDALDNAVEGLDDAVATLSKLSDLHAQKLRVDPANFKILSQCLLSTLANHRNPEFGPAVLASVDKFLCNVSEVLESKYR
12)   AHGKKVFAALHEAVNHIDDLPGALCRLSELHAHSLRVDPVNFKFLAQCVLVVVAIHHPGSLTPEVHASLDKFLCAVSSVLTSKYR
13)   AHGKKVFSALHEAVNHIDDLPGALCRLSELHAHSLRVDPVNFKFLAHCVLVVFAIHHPSALSPEIHASLDKFLCAVSAVLTSKYR
14)   GHGKKVFAALHDAVNHIDDLAGALCRLSDLHAHNLRVDPVNFKFLSQCILVVFGVHHPCSLTPEVHASLDKFLCAVSAMLTSKYR

15)   AHGQKVLTSFGEAVKNLDNIKKTFAQLSELHCEKLHVDPENFKLLGNILIIVLATHFPKEFTPASQAAWTKLVNAVAHALALGYH
16)   AHGKKVLTSFGEAVKNLDNIKKTFAQLSELHCEKLHVDPENFKLLGNILIIVLATHFPKEFTPASQAAWTKLVNAVAHALALGYH
17)   AHGKKVLTSFGEAVKNLDHIKQTFATLSKLHCEKLHVDPENFKLLGNVLIIVLASHFTKEFTPACQAAWQKLVSAVAHALALGYH
18)   EHGKKVLNSFGSAVKNMDHIKETFADLSKLHCETLHVDPENFKLLGSILIIVLAMHFGKEPTPTWQAAWQKLVSAVAHALTLQYH
19)   AHGKKVFTMFGEALKHLDNLKETFASLSELHCDKLHVDTENFKLLGNLVIVVLAARLHDSFTPAAQAAFHKLAYSVAHALARRYH
20)   AHGKKVFTSFGEAVKNLDNIKATYAKLSELHCEKLHVDPQNFNLLGDIFIIVLAAHFGKDFTPACQAAWQKLVRVVAHALAYHYH
21)   AHGKKVLTSFGDAVKNLDNIKDTFAKLSELHCDKLHVDPVNFRLLGNVMITRLAAHFGKDFTPACHAAFQKLTGAVAHALARRYH
22)   AHGKKVLTSFGDAIKNLDNIKDTFAKLSELHCDKLHVDPTNFKLLGNVLVIVLADHHGKEFTPAHHAAYQKLVNVVSHSLARRYH
23)   AHGKKVLTSFGEAIKHLDSIKETFAKLSELHCEKLHVDPENFRLLGNILIIVLAGHHGKEFTPSTHAAFQKLVRAVAHSLARVYH
24)   AHGKKVLTSFGEALKHLDNVKETFAKLSELHFDKLHVDPENFKLLGNVLIIVLAGHHGKEFTPSTHASFQKLVNVVAHALARRYH
25)   AHGKKVFTAFGDAVKNPEGVKDTFAKLSELHCDKLHVDPVNFKLLGQILITVLAAHFGKDFTPNVQAAYQKLVSVVAHALAHQYH
26)   AHGKKVLASFGEAVCHLDGIRAHFANLSKLHCEKLHVDPENFKLLGDIIIIVLAAHYPKDFGLECHAAYQKLVRQVAAALAAEYH
27)   EHGKKVLASFGEAVKHLDNIKGHFANLSKLHCEKFHVDPENFKLLGDIIIIVLAAHHPEDFSVECHAAFQKLVRQVAAALAAEYH
28)   EHGKKVLASFGEAVKHLDNIKGHFAHLSKLHFEKFHVDCENFKLLGDIIIVVLGMHHPKDFTLQTHAAFQKLVRHVAAALSAEYH
       ** **      *              *   *     **  **  *                              *        *   *
```

ZOOLOGICAL SCIENCE 19: 197–206 (2002)

The Primary Structure of Hemoglobin D from the Aldabra Giant Tortoise, *Geochelone gigantea*

Fumio Shishikura*

Department of Biology, Nihon University School of Medicine, Oyaguchi-kamimachi, Itabashi-ku, Tokyo 173-8610, Japan

ABSTRACT—The complete primary structures of α^D-2- and β-globin of hemoglobin D (Hb D) from the Aldabra giant tortoise, *Geochelone gigantea,* have been constructed by amino acid sequencing analysis in assistance with nucleotide sequencing analysis of PCR fragments amplified using degenerate oligonucleotide primers. Using computer-assisted sequence comparisons, the α^D-2-globin shared a 92.0% sequence identity *versus* α^D-globin of *Geochelone carbonaria*, a 75.2% *versus* α^D-globin of Aves (*Rhea americana*) and a 62.4% *versus* α^A-globin of Hb A expressed in adult red blood cells of *Geochelone gigantea*. Additionally, judging from their primary structures, an identical β-globin was common to the two hemoglobin components, Hb A and Hb D. The α^D-2- and β-globin genes contained the three-exon and two-intron configurations and showed the characteristic of all functional vertebrate hemoglobin genes except an abnormal GC dinucleotide instead of the invariant GT at the 5' end of the second intron sequence. The introns of α^D-2-globin gene were both small (224-bp/first intron, 227-bp/second intron) such that they were quite similar to those of adult α-type globins; the β-globin gene has one small intron (approximately 130-bp) and one large intron (approximately 1590-bp).

A phylogenetic tree constructed on primary structures of 7 α^D-globins from Reptilia (4 species of turtles, 2 species of squamates, and 1 species of sphenodontids) and two embryonic α-like globins from Aves (*Gullus gullus*) and Mammals (*Homo sapiens*) showed the following results: (1) α^D-globins except those of squamates were clustered, in which *Sphenodon punctatus* was a closer species to birds than turtles; (2) separation of the α^A- and α^D-globin genes occurred approximately 250 million years ago after the embryonic α-type globin-genes (π' and ζ) first split off from the ancestor of α-type globin gene family.

Key words: PCR, degenerate primer, nucleotide sequence, intron, exon

INTRODUCTION

Amniota (reptiles, birds and mammals), in general, have two or more hemoglobin components (Brown and Ingram, 1974; Moss and Hamilton, 1974; Lawn *et al.*, 1978; Efstratiadis *et al.*, 1980; Bunn and Forget, 1986; Fushitani *et al.*, 1996; Gorr *et al*, 1998) that are expressed according to the demands of different physiological conditions. Among them, hemoglobin D (Hb D) was first found in birds as a minor component of the embryonic and adult definitive erythrocytes (Hagopian and Ingram, 1971; Brown and Ingram, 1974). Based on functional studies of Hb D, the presence of α^D-globin raises the oxygen affinity and might be one such adaptation of insufficient oxygen supply as observed in the embryonic stages (Dodgson *et al.*, 1981; Chapman *et al.*, 1982) or extreme hypoxic and even anoxic conditions (Rücknagel and Braunitzer, 1988). On the other hand, the primary structure of α^D-globin of Hb D shows closely resemblance with embryonic hemoglobins (Chapman *et al.*, 1982) and thus, the Hb D is of interest for the study of the molecular evolution of Amniota globins because the distribution of the α^D-globin, to date, has been restricted in Aves and Reptilia (Rücknagel *et al.*, 1984; Abbasi *et al.*, 1988; Rücknagel *et al.*, 1988; Matsuura *et al.*, 1989; Fushitani *et al.*, 1996; Gorr *et al.*, 1998; Accession No. AF304335 in GenBank; Shishikura and Takami, 2001), except for Crocodilia (Leclercq *et al.*, 1981; Leclercq *et al.*, 1982). Most of the studies on globin gene structures have been carried out on birds and mammals (Bunn and Forget, 1986; Kleinschmidt and Sgouros, 1987), however, only one study has been conducted on reptilian α^D-globin cDNA structure from globin mRNA isolated from the red blood cells present in the adult *Geochelone carbonaria* (Accession No. AF304335 in GenBank). In addition to adult α^D-type globins, there are many genes related to α-globins such as embryonic α-like globins termed π'-globin (Chapman *et al.*, 1980) for birds and ζ-globin (Aschauer *et al.*, 1981) for mammals, all of

* Corresponding author: Tel. + 81-3-3972-8111(ext. 2291); FAX. +81-3-3972-0027. E-mail: fshishi@med.nihon-u.ac.jp

which are important clues for understanding the molecular evolution of α- and α-related globins.

This study describes the primary structures of α^D-2- and β^D-globin of *G. gigantea* Hb D (hereafter the author uses β instead of β^D because the primary structure of β^D-globin prepared from Hb D was definitively shown to be identical when compared with that of β-globin prepared from the *G. gigantea* Hb A) in assistance with nucleotide sequences of the two globin genes of *G. gigantea*, and constructs a phylogenetic tree concerning the molecular evolution of α^D-type globins. The tree also shows the relationships of α- and embryonic α-related globins, π'- and ζ-globin, as well as a few representatives of α^A-type globins from vertebrates. This study first describes the genomic structures of globins amplified by PCR with degenerate primers, and then, the nucleotide sequences, to ascertain the amino acid sequences of α^D-2- and β-globin. During the course of this study, it was also demonstrated that an identical β-globin was shared in both Hb A and Hb D as predicted in the previous study (Shishikura and Takami, 2001).

MATERIALS AND METHODS

Materials

Hb D from the Aldabra giant tortoise, *G. gigantea*, was prepared as described in the previous study (Shishikura and Takami, 2001).

Acetonitrile, ammonium sulfate, ammonium bicarbonate, *tri-n*-butyl phosphine, *4*-vinyl pyridine and *V8* protease (from *Staphylococcus aureus* strain *V8*) were purchased from Nakalai Tesque, Inc. (Kyoto, Japan). Separation columns, Alkyl Superose column HR5/5 and Resource column (prepackaged with 3 ml source 15 RPC gel matrix), were purchased and placed in a fast protein liquid chromatography (FPLC) system (Amersham Pharmacia Biotech, Upsala, Sweden). Lysyl endopeptidase (*Achromobactor* protease I) was purchased from Wako Pure Chemical Industries, Ltd. (Tokyo, Japan).

Taq DNA polymerase and GenElute Agarose Spin Columns were obtained from Sigma-Aldrich Co. (St. Louis, MO, USA). DNA molecular standard markers, pHY Marker (Takara Shuzo Co., Ltd., Shiga, Japan) and 100-bp DNA Ladder (New England Biolabs Inc., MA, USA) were used. Sequencing primers, M13 forward 17-mer (5'-GTA AAA CGA CGG CCA GT-3') and PUC/M13 reverse 17-mer (5'-CAG GAA ACA GCT ATG AC-3'), were obtained from Sigma-Aldorich Co. and Promega Co. (Madison, WI, USA), respectively. A BigDye Terminator Cycle Sequencing Ready Reaction Kit was purchased from Perkin-Elmer Japan Co. Ltd (Tokyo, Japan).

All other chemicals and solvents used were the most purified grade commercially available.

Globin-chain separation

The Hb D was modified by reduction and *S*-pyridylethylation (Friedman *et al.*, 1970) and then directly applied on a reversed-phase column (Resource column), which had been equilibrated with a 0.1% TFA solution. Removal of unincorporated reagents bound on the Resource column could be achieved by washing with an excess amount of 0.1% TFA solution until the base line was below 0.05 at 280 nm. The globin-peptides were, then, eluted from the column by a linear gradient with 60% acetonitrile in 0.08% TFA. Flow rates were maintained at 0.5 ml/min. All fractions were monitored at 214 nm and 280 nm by a spectrophotometer (Model 116, Gilson).

Enzymatic digestion and peptide separation

Lysyl endopeptidase digestion was performed essentially with modifications of Jekel *et al* (1983), the details of which were previously described (Shishikura and Takami, 2001). To obtain overlapping peptides, the globin (about 10 nmoles) was digested with the *V8* protease at a ratio of 1:100 (w/w, enzyme/substrate) for 48 hr at 37°C in a 0.1M Tris-HCl solution, pH 8.5 containing 1 M urea.

All peptides derived from the parent molecules were separated using a reversed-phased column, Resource, in a linear gradient with 60% acetonitrile in 0.08% TFA. Flow rates were maintained at 0.5 ml/min. All fractions were monitored at 214 nm and 280 nm by a spectrophotometer (Model 116, Gilson). When necessary, rechromatography of selected peptides was performed as previously described (Shishikura *et al.*, 1987).

Amino acid sequencing

Sequence analysis was performed using a Shimadzu gas phase protein sequencer, PPSQ-10, equipped with a PTH-10 amino acid analyzer (Shimadzu Co., Kyoto, Japan). Phenylthiohydantoin (PTH)-derivatives from the sequencer were separated and quantified. PTH-cysteine was detected as pyridylethylated-PTH-cysteine, the elution point of which was determined as described in the manufacturer's manual.

Isolation of genomic DNA

Prior to DNA extraction, fixed-tissue samples (80–120 mg) in absolute alcohol were dissolved in 600 µl DNA extraction buffer (10 mM Tris, 10 mM EDTA, 150 mM NaCl, pH 8.0) in a micro-centrifuge tube to obtain wet forms. Samples were treated with SDS (final concentration: 0.4%) and proteinase K (final concentration: 20 mg/ml), mixed well, and incubated for 60 min at 55°C, followed by overnight incubation at 37°C. The extraction of DNA was performed by the procedure described by Sambrook *et al.* (1989) with minor alterations: two rounds of precipitation with ethanol and spooling the precipitate purified DNA. DNA was then resuspended in 1 ml of TE buffer (10 mM Tris/HCl buffer containing 1 mM EDTA, pH 8.0) and stored at 4°C: about 0.1 mg/ml of high-molecular-weight genomic DNA was obtained, as evaluated by the absorption spectrum and by 0.8% agarose gel electrophoresis.

Primers design

Degenerate primers were designed based on the amino acid sequences of lysyl endopeptidase digested fragments of parent molecules. In order to sequence the PCR amplified fragments with a BigDye Terminator Cycle Sequencing Ready Reaction Kit, the degenerate oligo-nucleotide primers were tailed with the M13 forward or M 13 reverse sequencing primer tail (for the tail sequences shown above). A list of degenerate primers used in PCR amplifications is shown in Table 1.

PCR conditions

The PCR amplifications were performed in a 25-µl volume containing about 100 ng of genomic DNA template, 3 to 30 pmoles of each degenerate primer, deoxynucleotide triphosphates (400 µM) and 1.25 U of *Taq* DNA polymerase in the buffer conditions recommended by the manufacturer, 2.5 mM $MgCl_2$. The reactions started with denaturation at 95°C for 3 min, followed by 45 cycles and ended with 7 min of extension at 72°C on a DNA Thermal Cycler 9700 (Perkin-Elmer, Norwalk, CT, USA). The first five cycle profile began with denaturation for 1 min at 95°C, 5-stepwise different annealing temperatures (65°C, 62.5°C, 60°C, 57.5°C and 55°C) for 10 sec each, and ended with elongation for 1min every cycle at 72°C. The thermal profile including denaturation of the first 5 cycles modified the procedures described by Sachadyn *et al.* (1998), Skantar and Carta (2000), and Don *et al.* (1991). The remaining cycles were programmed according to the method recommended by the manufacturer.

Agarose gel electrophoresis

A 1.5% agarose gel was used to examine the purity and the size range of the PCR products amplified from the *Geochelone* genomic DNA. In each lane, except lanes of DNA-markers, 10 µl of each of the amplified DNA samples were loaded. The two DNA molecular weight standard markers were used. The gel was run in TBE (Tris-Borate-EDTA) buffer at 110V for 50 min. The results were then recorded using a KODAK Electrophoresis Documentation and Analysis System 290 (EDAS 290), and analyzed by a 1D Image Analysis Software (v. 3.5.4; Eastman Kodak Co., Rochester, NY, USA).

Extraction of PCR products and nucleotide sequencing analysis

After trimming away excess agarose, the gel slices (<500 µg) containing the PCR products were placed into the GenElute Agarose Spin column and centrifuged for 10 min at 14,000 x g. The filtrate was concentrated by Microcon-100 and sequenced using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit with the following modifications to the manufacturer's recommended protocol: 3.6 picomoles of M13 sequencing primer (forward or reverse) were annealed with about 32 ng of PCR product by mixing primer and template with 8 µl of Terminator Ready Reaction Mix in a final volume of 20 µl. This mixture was placed in a GeneAmp PCR system 9700 and subjected to cycle sequencing depending on the manufacturer's recommended protocol: start with heating for 10 sec at 96°C, and then 25 cycles of 96°C for 10 sec, 50°C for 5 sec, and 60°C for 4 min and reactions allowed to end with rapid thermal ramp at 4°C. Purifying extension products and the removal of unincorporated dye terminators in sequencing reactions were subjected to Centri-Sep spin columns (Princeton Separations P/N CS-901). Sequences of the PCR fragments were determined for both strands with the BigDye Primer Cycle Sequencing Ready Reaction Kit and the samples were on an ABI PRISM 310 Genetic Analyzer (Perkin-Elmer Japan Co. Ltd. Tokyo).

Computer analysis

A multiple alignment program, Clustal W (Thompson *et al.*, 1994), was used in the alignment of reptilian and other vertebrate's globin primary structures. Pair-wise distances among the globin sequences were analyzed using a computer program PROTDIST stored in the PHYLIP package (v. 3.51C; Felsenstein, 1993) under the Kimura-formula option. Based on the pair-wise distances, Neighbor-joining/UPGMA in NEIGHBOR (Felsenstein, 1993) was used to construct the phylogenetic tree of globins. Pair-wise alignments of DNA sequences were carried out using softwares of DNASIS as well as DNA Strider (V. 1.0.1).

RESULTS AND DISCUSSION

Globin isolation

In a preceding paper (Shishikura and Takami, 2001) we have described the isolation of the two hemoglobin components of the Aldabra giant tortoise *G. gitantea*, in which the two were designated as Hb A and Hb D. The nomenclature of Hb A and Hb D was adopted in Ingram's laboratory (Hagopian and Ingram, 1971; Brown and Ingram, 1974) where the various domestic fowl hemoglobins were defined. Among them, the adult definitive erythrocytes contained the major adult hemoglobin (Hb A) and the minor definitive hemoglobin (Hb D). After establishing the complete amino acid sequences of globins as described below, the presence of Hb D in the Aldabra giant tortoise, *G. gigantea*, was completely confirmed when compared with the known primary structures of α^D-globins (Kleinshmidt and Sgouros, 1987) specific to the Hb D. The advantage of modifying the protein by reduction and *S*-pyridylethylation also applied for separation of globin-constituents from the Hb D. As the results, three major fractions, α-1, α-2 and β in the order of elution, were separated as shown in Fig. 1: the two peaks, α-1 and α-2, were identical to each other having characteristics of α^D-type globins so far sequenced until the first 20 N-terminal amino acid residues, but in contrast their chromatograms on reversed-phase column were shown as distinctly different. There might be sequence microheterogeneity of their primary structures as found in those of Hb A (Shishikura and

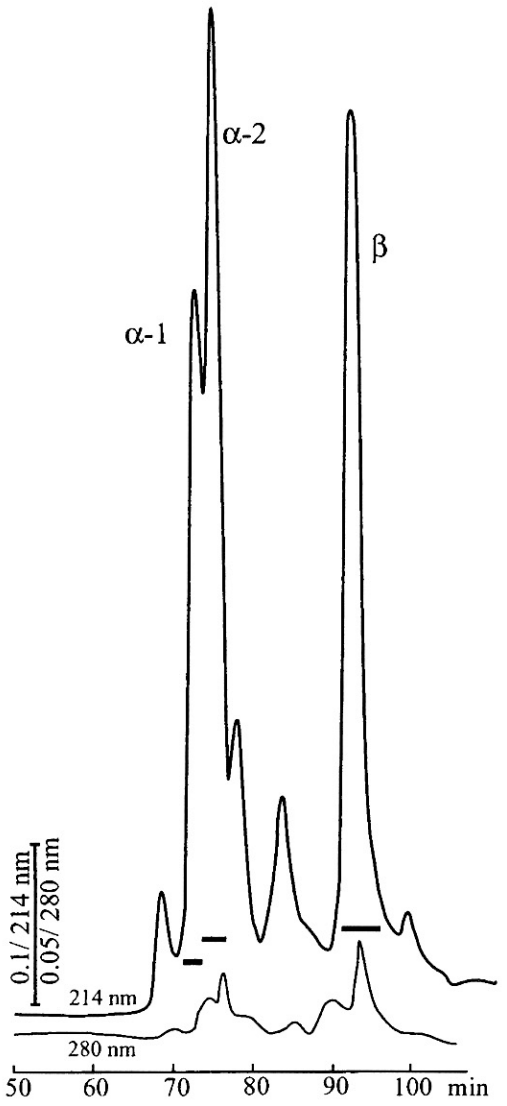


Fig. 1. Separation of globin-constituents from reduced and *S*-pyridylethylated *G. gigantea* Hb D (about 2 mg) on Resource column. A linear gradient was used between 0.1% TFA in water and 60% acetonitrile in 0.08% at a flow rate of 0.5 ml/min. Major peaks are designated as α-1, α-2 and β, respectively. Bars indicate fractions used for sequencing.

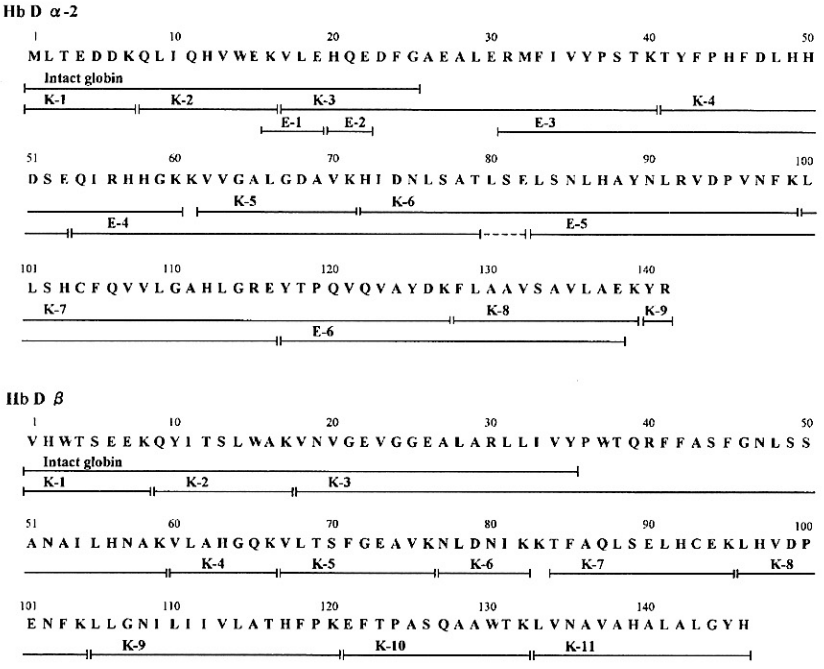


Fig. 2. Strategies and complete amino acid sequences of α-2 (*top*) and β (*bottom*) globins of Hb D from the Aldabra giant tortoise, *G. gigantea*. The complete amino acid sequences of the α-2 (*top*), and β (*bottom*) globins of *G. gigantea* Hb D have been established. Fragments generated by cleavage with lysyl endopeptidase and overlapping peptides obtained by *V8* protease are used for the amino acid sequence determination. The residues marked with continuous lines are those identified by Edman degradation method. Vertical lines represent the beginning and the end of sequencing. Dashed lines indicate the residues not determined but which might be included in the fragment. Peptide nomenclatures are as follows: lysyl endopeptidase; K, *V8* protease; E.

Takami, 2001). Hence, the author first sequenced the α-2-globin from the two α-types of globins.

Sequence strategies

For establishing complete primary structures of α-2- and β-globin, two sequencing methodologies, protein and DNA sequencing, were carried out. First, the parent molecules and their peptide fragments were sequenced and aligned tentatively with the assistance of sequence similarities toward the known primary structures of reptilian α- and β-globins, in particular, those obtained from the *G. gigantea* Hb A (Shishikura and Takami, 2001). Fig. 2 shows the results of amino acid sequence analyses of α^D-2- and β-globin. Appendix provides the data supporting the amino acid sequences in Fig. 2. The α^D-2-globin chain was composed of 141 amino acid residues and the β-globin chain was composed of 146 residues. Two lysine-lysine residues appeared in positions 60-61 of α^D-2-globin chain and 82–83 in β-globin chain were difficult to determine by analyzing the peptide fragments derived from digestions with lysyl endopeptidase. To complete the primary structure, peptide fragments containing the lysine-lysine residues generated by another enzymatic digestion such as *V8* protease are required to be sequenced. This was done in the construction of α^D-2-globin structure (Fig. 2, *top*) but required time-consuming work. To cope with time-consuming problems in

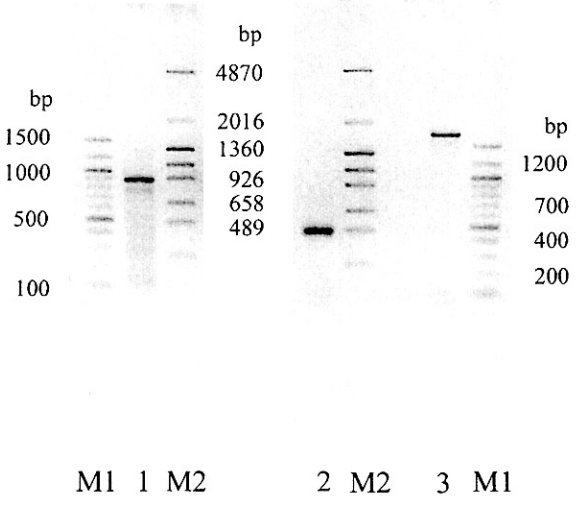


Fig. 3. Agarose gel electrophoreses of PCR products amplified from *G. gigantea* genomic DNA using degenerate primers. *Lane 1*; an 870-bp fragment amplified with degenerate PCR primers M13a-1 and M13a-2, *Lane 2*; a 480-bp fragment amplified with degenerate PCR primers M13b-1 and M13b-2, *Lane 3*; 1.75-kbp fragment amplified with degenerate PCR primers M13b-3 and M13b-4, *Lanes M1* and *M2*; DNA molecular standard markers, 100-bp DNA Ladder (M1) and pHY Marker (M2)

determining primary structures, the following methods were used: (1) based on sequencing information of both intact globins and digested fragments, degenerate oligo-nucleotide primers were synthesized with a M13 forward or M13 reverse sequencing primer tail; (2) using these primers (forward and reverse), a target gene was amplified by PCR from genomic DNA as a template; (3) the PCR fragment was purified and sequenced by cycle sequencing with the M13 forward or reverse sequencing primer. Fig. 3 shows amplified fragments on agarose gel electrophoresis: An 870-bp fragment was generated from the PCR-amplification of genomic DNA using primers M13a-1 and M13a-2, assuming amplified complete coding regions (three exons) and interviewing regions (two introns), and the remaining, 480-bp fragment and 1.75-kbp fragment, were amplified using primer-sets of M13b-1/M13b-2 and M13b-3/M13b-4, respectively; (4) nucleotide sequences of the three PCR fragments were determined with manufactured M13 sequencing primers, the BigDye Terminator Cycle Sequencing Ready Reaction Kit and the ABI PRISM 310 Genetic Analyzer; (5) finally, both protein and DNA sequencing data were complementary combined to establish complete structures of the α-2- and β-globin chains of *Geochelone* Hb D. As shown in Fig. 2 and Table 2 (2A and 2B), the two primary structures reinforced each other by the two different methods.

In comparison with the structural data of β-globin of Hb A (Shishikura and Takami, 2001), the primary structure of β-globin derived from the Hb D was completely identical, indicating that the β-globin was common in the construction of the two adult hemoglobin components, Hb A and Hb D. This finding supports the studies of Rücknagel and Braunitzer (1988) who described that the red blood cells shared the same β-globin chains in Hb A and Hb D. The sharing of identical β-globin chains has also been demonstrated in crocodiles (Leclercq *et al.*, 1981; Leclercq *et al.*, 1982), while lizards and snakes express two adult β-types of globins (Rücknagel *et al.*, 1988; Matsuura *et al.*, 1989; Abbasi and Braunitzer, 1991; Naqvi *et al.*, 1994; Fushitani *et al.*, 1996; Gorr *et al.*, 1998). In this context, adult mammals (Braunitzer *et al.*, 1961; Leclercq *et al.*, 1981) and birds (Rücknagel *et al.*, 1984; Oberthür *et al*, 1983; Oberthür *et al.*, 1986) have been reported to have one kind of β-globin, but adult frogs (Knöchel *et al.*, 1983; Patient *et al.*, 1983) contained two subtypes of β-globin chains. Due to an inconsistency in the number of subtypes of adult β-type globin-chains among amphibians, reptiles, birds and mammals, reinvestigations are needed, especially, in regards to the evolution of Tetrapoda (Benton, 1990; Hardison, 1998).

Comparison of the primary structure of α-type globins within *G. gigantea*, α^D-2-globin of *G. gigantea* differs from α^A-globin in 53 amino acid residues (62.4% identity), but when compared with homologous globin chains found in adult *Geochelone carbonaria* (a different species of tortoises) and adult *Rhea Americana* (a species of birds), only 7 (95.0% idnetity) and 35 (75.2% identity) amino acid residues were substituted, respectively.

PCR amplification of globin gene by degenerate primers

Two degenerate oligo-nucleotides (M13a-1 and M13a-2 in Table 1) which were designed from the regions of N-terminal (8 amino acid residues in length) and C-terminal (8 amino acid residues in length) of α^D-2-globin successively amplified a PCR-product with 870-bp estimated by migration distance on agarose gel electrophoresis (Fig. 3, *lane 1*). On the contrary, in the case of amplification of β-globin using M13b-1 and M13b-4 primers no product was observed on agarose gel electrophoresis, indicating that the whole coding region of β-globin gene was impossible to amplify at once using two degenerate primers designed by its N-terminal and C-terminal amino acid sequences. It seems to be difficult to amplify extremely long nucleotides such the case over 1.75-kbp PCR-fragment. Hence, several sets of sense and anti-sense degenerate primers were synthesized and used for amplification of β-globin gene in total with the genomic DNA: the two sets of sense and anti-sense primers (M13b-1/M13b-2 and M13b-3/M13b-4 shown in Table 1) produced a single fragment in each PCR, in which nucleotide-sized fragments were determined to be a 480-bp fragment and a 1.75-kbp fragment, respectively (Fig. 3, *lane 2* and *3*). Both products and the 870-bp fragment of α^D-2-globin gene were sequenced from both sides and aligned by computer-assisted programs. Table 2A and 2B show alignments of nucleotide sequences in encompassing whole

Table 1. Oligo-nucelotide primers used in this study

Gene	Primer Name		Nucleotide sequences	Reference	Degeneracy
			1 10 20		(fold)
Hb D α-2	M13a-1(Forward)	M13-	A T G Y T N A C N G A R G A Y G A Y A A R C A	N-terminal	512
	M13a-2(Reverse)	M13-	A A Y T T R T C R T A N G C N A C Y T G N A C	C-terminal	1024
			1 10 20 20		
Hb D β	M13b-1(Forward-1)	M13-	G T G C A C T G G A C Y W S N G A R G A G A A G	N-terminal	64
	M13b-2(Reverse-1)	M13-	C T T G A A G T T C T C R G G R T C C A C R T G	104-97	8
	M13b-3(Forward-2)	M13-	C A Y G T G G A Y C C Y G A G A A C T T C A A G	97-104	8
	M13b-4(Reverse-2)	M13-	G T G G T A V C C S A G R G C C A G R G C R T G	C-terminal	48

M13 forward sequence: 5'-GTA AAA CGA CGG CCA GT-3'
M13 reverse sequence: 5'-CAG GAA ACA GCT ATG AC-3'
The International Union of Pure and Applied Chemistry Symbols used to denote multiple nucleotides are as follows: N=A+G+C+T; R=A+G; S=C+G; V=A+G+C; W=A+T; Y=C+T.

Table 2A. Nucleotide sequences of three exons and exon-intron boundaries of α^D-2-globin gene

Exon-1	1	ATG	CTA	ACA	GAG	GAC	GAC	AAG	CAG	CTG	ATC	CAA	CAT	GTG	TGG	GAG
	46	AAG	GTG	CTG	GAG	CAC	CAG	GAG	GAC	TTT	GGG	GCC	GAG	GCC	CTG	GAG
	91	AG														
Intron-1	1	gta	ggg	ccc	ggg	gca	ggc	ggc	ccg	ggc	gca	ggg	tgc	agg	gag	gaa
		(46-168 not shown)														
	169	tgg	ggg	act	ttg	gga	ttc	act	gtc	tct	gac	ctc	cct	ccc	ccg	cag
Exon-2	1	G	ATG	TTC	ATC	GTC	TAC	CCC	TCC	ACC	AAG	ACC	TAC	TTC	CCC	CAC
	44	TTC	GAC	CTG	CAT	CAT	GAC	TCG	GAA	CAG	ATC	CGC	CAC	CAC	GGC	AAG
	89	AAG	GTG	GTG	GGC	GCC	CTG	GGG	GAC	GCC	GTG	AAG	CAC	ATC	GAC	AAC
	134	CTC	AGC	GCG	ACG	CTC	TCC	GAG	CTC	AGC	AAC	CTG	CAC	GCC	TAC	AAC
	179	TTG	CGC	GTG	GAC	CCG	GTC	AAC	TTC	AAG						
Intron-2	1	gca	agt	gca	ggc	tac	ggc	cag	gaa	gag	ttc	ccg	ggg	ggt	gcg	gga
		(46-181 not shown)														
	182	ggc	cga	ggg	ctg	gct	gcc	gct	gac	cca	gtg	cac	ttt	gct	ttg	cag
Exon-3	1	CTG	CTG	TCC	CAC	TGC	TTC	CAG	GTG	GTG	CTG	GGC	GCG	CAC	TTG	GGC
	46	CGC	GAG	TAC	ACC	CCG	CAG	GTG	CAA	GTC	GCC	TAT	GAC	AAG	TTC	CTG
	91	GCC	GCC	GTC	TCG	GCG	GTG	CTG	GCT	GAG	AAG	TAC	CGG			

Table 2B. Nucleotide sequences of three exons and exon-intron boundaries of β-globin gene

Exon-1	1	GTG	CAC	TGG	ACC	AGC	GAG	GAG	AAG	CAG	TWC	ATT	ACC	AGT	CTG	TGG
	46	GMC	AAG	GTC	AAC	GTG	GRG	GAA	GTG	GGT	GGC	GAA	GCC	CTG	GCC	AG
Intron-1	1	gta	ggc	tcg	agc	ctc	aca	tgg	ata	tct	gcc	tcg	cat	tgc	tcc	tct
		(approximately 40 nucleotides not shown)														
		gca	gta	acc	ctg	tgt	ctg	tct	ctg	ctc	ctg	tct	ccc	tct	ctc	tag
Exon-2	1	G	CTG	CTG	ATC	GTC	TAC	CCC	TGG	ACC	CAG	AGG	TTT	TTC	KCT	TCC
	44	TTT	GGG	AAC	CTS	TCC	AGC	SCC	AAC	GCC	ATC	MTG	SRC	AAC	GCC	AAG
	89	GTG	STT	GCC	CAT	GGC	MAG	AAA	GTG	CTG	ACC	TCS	TTT	GGG	GAA	GCT
	134	GTG	AAG	AAC	CTG	GAC	AAC	ATC	AAG	RMM	ACG	TWC	GCC	CAG	CTG	AGC
	179	GAG	CTG	CAC	TGC	SAR	AAG	CTG	CAT	GTG	GAT	CCT	GAG	AAC	TTC	AAG
Intron-2	1	gtg	agt	ccg	gct	ctg	ggt	tga	ccc	tct	tcc	cag	ccc	cct	ttc	cat
		(approximately 1590 nucleotides not shown)														
		cag	agc	ggt	gct	gac	cca	gcg	ggt	atc	ttc	ttc	ctt	ctc	ctc	cag
Exon-3	1	CTC	CTG	GGC	AAT	ATC	CTC	ATC	ATC	GTC	CTG	GCC	ACC	CAC	TTC	CCA
	46	AAG	GAG	TTC	ACT	CCT	GCC	AGT	CAG	GCC	GCC	TGG	ACA	AAG	CTC	GTC
	91	AAT	GCA	GTG	GCC	CAT	GCT	CTG	GCT	CTC	GGT	TAC	CAC			

The International Union of Pure and Applied Chemistry Symbols used to denote multiple nucleotides are as follwos: K=G or T; M=A or C; R=A or G; S=G or C; W=A or T.

exon regions of the α^D-2- and the β-globin gene and exon-intron boundaries of the two genes. Breathnach and Chambon (1981) stated that there was no exception to the GT-AG rule according to which all intron sequences start with GT and end with AG. However, Table 2A shows that a unique structural feature of the α^D-2-globin gene is a GC instead of a GT dinucleotide at the 5' end of the second intron sequence. This finding is the first exception found in reptilian hemoglobin gene and supported the previous studies on gene structures of bird's hemoglobin (Erbil and Niessing, 1983; Dodgson and Engel, 1983). Erbil and Niessing (1983) found the T to C transition at the second intron position 2 of α^D-globin gene from a duck, *Cairina moschata*. This evidence together with the unique structure of α^D-2-globin gene found in the tortoise strongly indicates that the two animals, tortoises and birds, are the closest living relatives to each other.

When compared with the intron lengths among the four α-types of globin genes (Table 3), it was clearly determined that the *Geochelone* α^D-2-globin gene structure corresponded to that of the adult chicken α^D-globin gene and not to the embryonic chicken π'-globin gene (Engel *et al.*, 1983) nor the embryonic human ζ-globin gene (Proudfoot *et al.*, 1982). On the contrary, the *G. gigantea* β-globin gene was hard to classify since the second intron length (about 1.59-kbp) was large compared with those of the adult β-globin gene (Lawn *et al.*, 1980; Dolan *et al.* 1983), including embryonic β-like globins (Efstratiadis *et al.*, 1980; Chapman *et al.*, 1981). In addition, the evolutionary relatedness of the intron sizes of the *Geochelone* globin genes to the other amniotes

Table 3. Comparison of exson and intron sizes (in bp) of α^D -2- and β-globin genes

Class	Globin-gene Name	Exon-1	1st Intron	Exon-2	2nd Intron	Exon-3	Total	Reference
Reptilia	***G. gigantea* α^D-2**	**92**	**214**	**205**	**227**	**126**	**864**	**This study**
	G. carbonaria α^D	92	n.d.	205	n.d.	126	n.d.	AF304335 in GenBank
Aves	*Gullus gullus* α^D	92	148	205	261	126	832	Dodgson and Engel, 1983
	Gullus gullus π'	92	577	205	294	126	1294	Engel *et al.*, 1983
Mammalia	*Homo sapiens* ζ	92	886	205	239	126	1548	Proudfoot *et al.*,1982
	Homo sapiens α	92	117	205	141	126	681	Liebhaber *et al.*,1980
Reptilia	***G. gigantea* β**	**89**	**130**[1)]	**223**	**1590**[1)]	**126**	**2158**[1)]	**This study**
	G. carbonaria β	89	n.d.	223	n.d.	126	n.d.	Bordin *et al.*,1997
Aves	*Gullus gullus* β	82	92	223	810	126	1333	Dolan *et al.*, 1983
Mammalia	*Homo sapiens* β	89	130	223	850	126	1418	Lawn *et al.*, 1980

n.d.; not determined. Intron sizes could not be determined as this sequence is only represented by an RT-PCR product.
1) Intron sizes were estimated by migration distances on agarose gel electorphoresis (Fig. 4).

globin genes was defined for the first time.

Reptilian phylogeny and diversity based on α^D-Globin structures

Shishikura and Takami (2001) have constructed a phylogenetic tree based on α- and β-globins of 28 reptilian Hb As, by which the molecular phylogeny of Reptilia is highly correlated at the level of orders with the traditional phylogeny established mainly upon their morphological characteristics (Carroll, 1969; Benton, 1990). To date, there have been four different types of α-globins in amniotes reported: α^A, α^D, π' and ζ. The former two are adult α-type globins and the remaining are embryonic α-like globins. Fig. 4. shows a molecular tree of reptilian evolution constructed mainly by α^D-globins of 7 reptiles as well as relatedness among representatives of adult and embryonic α-type globins. The tree also strongly supports the previous molecular studies (Goodman *et al.*, 1975; Fushitani *et al.*, 1996; Gorr *et al.*, 1998; Shishikura and Takami, 2001), however, it is reasonable to note the following two points: (1) the two kinds of embryonic globins, π' and ζ, first split off from the ancestor of the α-type of globins and formed a cluster; (2) the ancestor of squamates (snakes; *L. miliaris*, lizards; *Varanus komodoensis*) occupied unusual positions since α^D-globins of squamates began to diverge approximately 335 million years ago, much earlier than the separation of the three other clusters of α-type globin families.

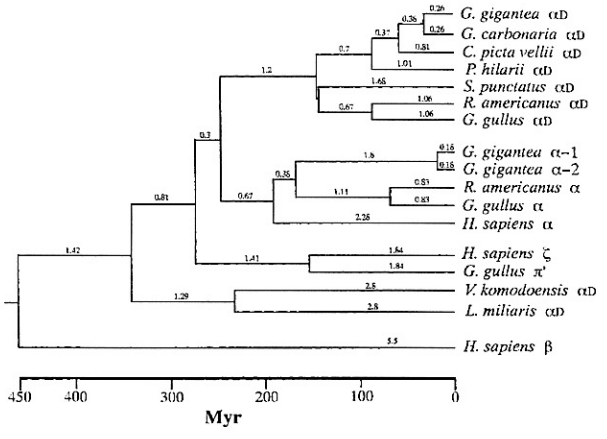


Fig. 4. A phylogenetic Tree based on primary structures of α^D-types of globins including some representative Amniota embryonic α-like globins and *Homo sapiens* β-globin as an outer group. Branch lengths are proportional to protein distances(×1/10) and shown on the individual branches of the tree. The abscissa is a time scale in Myr (million years) ago based on the separations of the α- and β-globin chains described by Goodman *et al.*, (1975). The references of primary structures of globins used in the present analysis are as follows: α^D-globins: *G. gigantea* (this study), *G. carbonaria* (Accession No. AF304335 in GenBank), *C. picta vellii* (Rüchnagel *et al.*, 1984), *S. punctatus* (Abassi *et al.*, 1998), *R. americanus* (Oberthür *et al.*, 1986), *G. gullus* (Takei *et al.*, 1975), *V. komodoensis* (Fushitani *et al.*, 1996), *L. miliaris* (Matsuura *et al.*, 1989). α^A-globins: *G. giantea* α-1 and α-2 (Shishikura and Takami, 2001), *R. americanus* (Oberthür *et al.*, 1983), *G. gullus* (Knöchel *et al.*, 1982), *H. sapiens* (Braunitzer *et al.*, 1961). Embryonic α-like globins: *H. sapiens* ζ (Aschauer *et al.*, 1981), *G. gullus* π' (Chapman *et al.*, 1980; 1982). *H. sapiens* β-globin (Braunitzer *et al.*, 1961).

ACKNOWLEDGMENTS

The author thanks the staff of veterinary surgeons of Animal Hospital of Osaka Municipal Tennoji Zoo for their cooperation in collecting blood from the Aldabra giant tortoise, *G. gigantea*. The author also thanks Dr. M. Esumi and Dr Y-H. Zhou, Department of Pathology, Nihon University School of Medicine, for her generosity (Dr. Esumi) in helping to prepare the equipment for analyzing amino acid sequences of globins and his instruction (Dr. Zhou) for analyses of nucleotide sequences under the ABI PRISM 310 Genetic Analyzer.

REFERENCES

Abbasi A, Braunitzer G (1991) Primary structure of hemoglobin from Monitor Lizard (*Varanus exanthematicus albigularis*-Squamata). Biol Chem Hoppe-Seyler 372: 473–479

Abbasi A, Wells RMG, Brittain T, Braunitzer G (1988) Primary structure of the hemoglobins from Sphenodon (*Spehnodon punctatus*, Tuatara, Rhyncocephalia). Biol Chem Hoppe-Seyler 369: 755–764

Aschauer H, Sanguansermsri T, Braunitzer G (1981) Embryonale Hamoglobine des Menschen: Die Primärstruktur der ζ-Ketten. Hoppe-Seyler's Z Physiol Chem 362: 1159–1162

Benton MJ (1990) Phylogeny of the major tetrapod groups: Morpho-

logical data and divergence dates. J Mol Evol 30: 409–424
Bordin S, Meza AN, Saad STO, Ogo SH, Costa FF (1997) cDNA-derived amino-acid sequence of a land turtle (*Geocheone carbonaria*) β-chain hemoglobin. Biochem Mol Biol Int 42: 255–260
Braunitzer G, Gehring-Muller R, Hilschmann H, Hilse K, Hobom G, Rudloff V, Wittmann-Liebold B (1961) The constitution of normal adult human haemoglobin. Hoppe-Seyler's Z Physiol Chem 325: 283–286
Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. Ann Rev Biochem 50: 349–383
Brown JL, Ingram VM (1974) Structural studies on chick embryonic hemoglobins. J Biol Chem 249: 3960–3972
Bunn HF, Forget BG (1986) Hemoglobin function, In: Bunn HF, Forget BG (Eds) Hemoglobin: Molecular, genetic and clinical aspects. Saunders WB Co, Philadelphia, Chap 3, pp 37–60
Carroll RL (1969) Origin of reptiles, In: Gans C *et al*, (Eds), Biology of Reptilia, Vol. 1 Academic Press, London & New York, pp 1–44
Chapman BS, Tobin AJ, Hood LE (1980) Complete amino acid sequences of the major early embryonic β-like globins of the chicken. J Biol Chem 255: 9051–9059
Chapman BS, Tobin AJ, Hood LE (1981) Complete amino acid sequence of the major early embryonic α-like globin in chickens. J Biol Chem 256: 5524–5531
Chapman BS, Hood LE, Tobin AJ (1982) Minor early embryonic chick hemoglobin M. Amino acid sequences of the ε- and α^D chains. J Biol Chem 257: 651–658
Dodgson JB, McCune KC, Rusling DJ, Krust A, Engel JD (1981) Adult chicken α-globin genes, a^A and α^D: No anemic shock α-globin exists in domestic chickens. Proc Natl Acad Sci USA 78: 5999–6002
Dodgson JB, Engel JD (1983) The nucleotide sequence of the adult chicken α-globin genes. J Biol Chem 258: 4623–4629
Dolan M, Dodgson JB, Engel JD (1983) Analysis of the adult chicken β-globin gene. Nucleotide sequence of the locus, microheterogeneity at the 5'-end of β-globin mRNA, and aberrant nuclear RNA species. J Biol Chem 258: 3983–3990
Don PH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) "Touchdown" PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res 19: 4008
Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissman SM, Slightom JL, Blechl AE, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human β-globin family. Cell 21: 653–668
Engel JD, Rusling DJ, McCune KC, Dodgson JB (1983) Unusual structure of the chicken embryonic α-globin gene, π'. Proc Natl Acad Sci USA 80: 1392–1396
Erbil C, Niessing J (1983) The primary structure of the duck α^D-globin gene: an unusual 5' splice junction sequence. EMBO J 2: 1339–1343
Felsenstein J (1993) PHYLIP: Phylogeny Inference Package, Version 3.5c. Department of Genetics SK-50, University of Washington, Seattle, WA 98195, USA
Friedman M, Krull LH, Cavins JF (1970) The chromatographic determination of cystine and cysteine residues in proteins as *S*-β -(4-pyridylethyl) cysteine. J Biol Chem 245: 3568–3871
Fushitani K, Higashiyama K, Moriyama EN, Imai K, Hosokawa K (1996) The amino acid sequences of two α chains of hemoglobins from Komodo dragon *Varanus komodoensis* and phylogenetic relationship of amniotes. Mol Biol Evol 13: 1039–1043
Goodman M, Moore GW, Matsuda G (1975) Darwinian evolution in the genealogy haemoglobin. Nature 253: 603–608
Gorr TA, Mable BK, Kleinschmidt T (1998) Phylogenetic analysis of reptilian hemoglobins: Trees, rates, and divergences. J Mol Evol 47: 471–485
Hagopian HK, Ingram VM (1971) Developmental changes of erythropoiesis in cultured chick blastoderms. J Cell Biol 51: 440–451
Hardison R (1998) Hemoglobins from bacteria to man: evolution of different patterns of gene expression. J Exp Biol 201: 1099–1117
Jekel PA, Weijer WJ, Beintema JJ (1983) Use of endoproteinase Lys-C from *Lysobacter enzymogenes* in protein sequence analysis. Anal Biochem 134: 347–354
Kleinschmidt T, Sgouros JG (1987) Hemoglobin sequences. Biol Chem Hoppe-Seyler 368: 579–615
Knöchel W, Wittig B, Wittig S, John ME, Grundmann U, Oberthür W, Godovac J, Braunitzer G (1982) No evidence for 'stress' α-globin genes in chicken. Nature 295: 710–712
Knöchel W, Meyerhof W, Hummel S, Grundmann U (1983) Molecular cloning and sequencing of mRNAs coding for minor adult globin polypeptides of *Xenopus laevis*. Nucleic Acids Res 11: 1543–1553
Lawn RM, Fritsch EF, Parker RC, Blake G, Maniatis T (1978) The isolation and characterization of linked δ- and β-globin genes from a cloned library of human DNA. Cell 15: 1157–1174
Lawn RM, Efstratiadis A, O'Connell C, Maniatis T (1980) The nucleotide sequence of the human β-globin gene. Cell 21: 647–651
Leclercq F, Schnek AG, Braunizer G, Stangl A, Schrank B (1981) Direct reciprocal allosteric interaction of oxygen and hydorgen carbonate sequence of the haemoglobins of the caiman (*Caiman crocodylus*), the Nile crocodile (*Crocodylus niloticus*) and the Mississippi crocodile (*Alligator mississippiensis*). Hoppe-Seyler's Z Physiol Chem 362: 1151–1158
Leclercq F, Stangl A, Schrank B, Vandecasserie C, Paul C, Schnek AG, Leonis J, Braunitzer G (1982) *Caiman crocodylus* hemoglobin. Complete primary structures of α and β chains; phylogenic and regulatory aspects. Int J Peptide Protein Res 20: 337–350
Liebhaber SA, Goossens MJ, Kan YW (1980) Cloning and complete nucleotide sequence of human 5'-α-globin gene. Pro Natl Acad Sci USA 77: 7054–7958
Matsuura MSA, Fushitani K, Riggs AF (1989) The amino acid sequences of the α and β chains of hemoglobin from the snake, *Liophis miliaris*. J Biol Chem 264: 5515–5521
Moss BA, Hamilton EA (1974) Chicken definitive erythorcyte haemoglobins. Biochim Biophys Acta 371: 379–391
Naqvi S, Abbasi A, Zaidi ZH (1994) Primary structure of hemoglobin from cobra *Naja naja naja*. J Prot Chem 13: 669–679
Oberthür W, Braunitzer G, Baumann R, Wright PG (1983) Die Primärstruktur der α- und β-Ketten der Hauptokomponenten der Hämoglobine des Straußes (*Struthio camelus*) und des Nandus (*Rhea americana*) (Struthioformes). Aspekte zur Atmungsphysiologie und Systematik. Hoppe-Seyler's Z Physiol Chem 364: 119–134
Oberthür W, Godovac-Zimmermann J, Braunitzer G (1986) The expression of an α^D-chains in the hemoglobin of adult ostrich (*Struthio camelus*) and American rhea (*Rhea americana*). The different evolution of adult bird α^A-, α^D- and β-chains. Biol Chem Hoppe-Seyler 367: 507–514
Patient RK, Harris R, Walmsley ME, Williams JG (1983) The complete nucleotide sequence of the major adult β globin gene of *Xenopus laevis*. J Biol Chem 258: 8521–8523
Proudfoot NJ, Gil A, Maniatis T (1982) The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. Cell 31: 553–563
Rücknagel KP, Reischl E, Braunitzer G (1984) Hämoglobine der Reptilian. Expression von α^D-Genen bei Schildkroten, *Chrysemys picta bellii* und *Phrynops hilarii* (Testudines). Hoppe-Seyler's Z Physiol Chem 365: 1163–1171
Rücknagel KP, Braunitzer G (1988) Hemoglobins of reptiles. 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

Rücknagel KP, Braunitzer G, Wiesner H (1988) Hemoglobins of reptiles. The primary structures of the α^{I}- and β^{I}-chains of common iguana (*Iguana iguana*). Biol Chem Hoppe-Seyler 369: 1143–1150

Sachadyn P, Sobiewska G, Kur J (1998) Thermal profile with alternately raised and lowered annealing temperature improves the PCR amplification using highly degenerate primers. Acta Biochim Pol 45: 691–694

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual. 2nd ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Shishikura F, Snow JW, Gotoh T, Vinogradov SN, Walz DA (1987) Amion 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 sequecences of the α- and β-globin chains of hemoglobin from the Aldabra giant tortoises, *Geochelone gigantea*. Zool Sci 18: 515–526

Skantar AM, Carta LK (2000) Amplification of Hsp 90 homologs from plant-parasitic nematodes using degenerate primers and ramped annealing PCR. BioTechniques 29: 1182–1185

Takei H, Ota Y, Wu KC, Kiyohara T, Matsuda G (1975) Amino acid sequence of the α chain AI hemoglobin. J Biochem 77: 1345–1347

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequences weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

(Received September 20, 2001 / Accepted November 2, 2001)

Appendix. Sequence analysis of peptides obtained by cleavage with lysyl endopeptidase and *V8* protease

Step	Hb D α-2 Amino Acid (*p* moles)				Hb D β Amino Acid (*p* moles)		
	Intact	**K-1**			**Intact**	**K-1**	
1	M (254)	M (6160)			V (687)	V (3918)	
2	L (297)	L (5178)			H (272)	H (884)	
3	T (117)	T (2209)			W (383)	W (104)	
4	E (291)	E (4423)			T (265)	T (1340)	
5	D (178)	D (2857)			S (91)	S (322)	
6	D (229)	D (3461)			E (436)	E (1164)	
7	K (232)	K (1208)	**K-2**		E (498)	E (1472)	
8	Q (127)		Q (2241)		K (458)	K (354)	**K-2**
9	L (151)		L (2614)		Q (315)		Q (2687)
10	I (117)		I (2241)		Y (247)		Y (2053)
11	Q (111)		Q (1782)		I (294)		I (2608)
12	H (59)		H (459)		T (150)		T (1210)
13	V (122)		V (2006)		S (49)		S (347)
14	W (56)		W (1206)		L (247)		L (1855)
15	E (104)		E (1360)	**E-1**	W (85)		W (645)
16	K (101)	**K-3**	K (989)	K (542)	A (256)		A (1722)
17	V (103)	V (1448)		V (1039)	K (178)	**K-3**	K (787)
18	L (94)	L (1357)		L (800)	V (205)	V (2667)	
19	E (94)	E (1219)	**E-2**	E (295)	N (186)	N (2406)	
20	H (42)	H (263)	H (368)		V (204)	V (2489)	
21	Q (60)	Q (870)	Q (871)		G (145)	G (1667)	
22	E (83)	E (1075)	E (831)		E (136)	E (2434)	
23	D (68)	D (532)			V (338)	V (1899)	
24	F (68)	F (814)			G (140)	G (1323)	
25	G (47)	G (535)			G (163)	G (1354)	
26		A (730)			E (386)	E (1726)	
27		E (727)			A (235)	A (1896)	
28		A (695)			L (211)	L (1716)	
29		L (581)			A (252)	A (1665)	
30		E (592)		**E-3**	R (123)	R (840)	
31		R (400)		N.D.	L (350)	L (1430)	
32		M (435)		M (785)	L (316)	L (1600)	
33		F (389)		F (761)	I (200)	I (1171)	
34		I (337)		I (1979)	V (185)	V (1140)	
35		V (314)		V (762)	Y (122)	Y (934)	
36		Y (265)		Y (487)	P (104)	P (776)	
37		P (270)		P (587)		W (332)	
38		S (44)		S (107)		T (439)	
39		T (134)		T (295)		Q (640)	
40	**K-4**	K (58)		K (803)		R (412)	
41	T (4005)			T (268)		F (650)	
42	Y (6603)			Y (279)		F (757)	
43	F (7549)			F (492)		A (541)	
44	P (4645)			P (256)		S (93)	
45	H (2838)			H (109)		F (413)	
46	F (5030)			F (308)		G (232)	
47	D (4682)			D (256)		N (334)	
48	L (4563)			L (216)		L (291)	
49	H (2133)			H (107)		S (58)	
50	H (2659)			H (142)		S (63)	
51	D (3410)			D (167)		A (254)	
52	S (570)			S (15)		N (175)	
53	E (2847)		**E-4**	E (68)		A (225)	
54	Q (2077)		Q (336)			I (136)	
55	I (2060)		I (480)			L (156)	
56	R (1581)		N.D.			H (65)	
57	H (947)		N.D.			N (119)	
58	H (1391)		H (95)			A (137)	
59	G (1365)		G (108)			K (42)	**K-4**
60	K (1344)		K (228)				V (3698)
61		**K-5**	K (218)				L (3684)
62		V (7177)	V (296)				A (3476)
63		V (6907)	V (170)				H (1136)
64		G (4722)	G (97)				G (1916)
65		A (6187)	A (138)				Q (2164)
66		L (5590)	L (200)			**K-5**	K (1299)
67		G (3871)	G (135)			V (1036)	
68		D (3735)	D (130)			L (1090)	
69		A (4395)	A (187)			T (949)	
70		V (3403)	V (169)			S (182)	
71	**K-6**	K (2660)	K (122)			F (719)	
72	H (917)		H (48)			G (421)	
73	I (2647)		I (74)			E (1007)	

Step	Hb D α-2 Amino Acid (*p* moles) Continued				Hb D β Amino Acid (*p* moles) Continued	
74	D (1667)		D (167)		A (790)	
75	N (1990)		N (59)		V (397)	
76	L (2351)		L (85)		K (319)	**K-6**
77	S (360)		S (13)			N (3797)
78	A (2010)		A (78)			L (3545)
79	T (798)		T (97)			D (2082)
80	L (1669)					N (2410)
81	S (235)					I (1914)
82	E (1070)		**E-5**			K (741)
83	L (1085)		L (896)		**K-7**	
84	S (164)		S (145)		T (1846)	
85	N (702)		N (535)		F (3200)	
86	L (862)		L (664)		A (3206)	
87	H (259)		H (139)		Q (2612)	
88	A (742)		A (604)		L (2743)	
89	Y (477)		Y (412)		S (421)	
90	N (582)		N (504)		E (1800)	
91	L (615)		L (364)		L (1788)	
92	R (625)		R (182)		H (860)	
93	V (483)		V (465)		C/pe-cys	
94	D (N.D)		D (293)		E (1547)	
95	P (259)		P (314)		K (831)	**K-8**
96	V (254)		V (359)			L (559)
97	N (218)		N (291)			H (174)
98	F (177)		F (267)			V (400)
99	K (117)	**K-7**	K (372)			D (315)
100		L (4545)	L (284)			P (280)
101		L (4673)	L (423)			E (519)
102		S (793)	S (43)			N (258)
103		H (1551)	H (114)			F (279)
104		C/pe-cys	C/pe-cys		**K-9**	K (259)
105		F (2856)	F (196)		L (2391)	
106		Q (2256)	Q (174)		L (1757)	
107		V (2454)	V (181)		G (674)	
108		V (2869)	V (217)		N (790)	
109		L (2279)	L (155)		I (425)	
110		G (1458)	G (117)		L (247)	
111		A (2129)	A (153)		I (252)	
112		H (656)	H (69)		I (416)	
113		L (1709)	L (142)		V (141)	
114		G (1272)	G (114)		L (174)	
115		R (1150)	R (97)		A (185)	
116		E (1564)	E (52)	**E-6**	T (128)	
117		Y (1098)		Y (3742)	H (53)	
118		T (601)		T (259)	F (127)	
119		P (974)		P (286)	P (50)	
120		Q (956)		Q (314)	K (34)	**K-10**
121		V (989)		V (306)		E (3554)
122		Q (878)		Q (249)		F (3406)
123		V (919)		V (244)		T (849)
124		A (889)		A (272)		P (2086)
125		Y (626)		Y (200)		A (2568)
126		D (535)		D (216)		S (521)
127	**K-8**	K (387)		K (316)		Q (1504)
128	F (3065)			F (212)		A 1090)
129	L (2272)			L (238)		A (1885)
130	A (2211)			A (213)		W (683)
131	A (2732)			A (219)		T (559)
132	V (1917)			V (202)	**K-11**	K (731)
133	S (273)			S (33)	L (1518)	
134	A (1525)			A (174)	V (1388)	
135	V (1329)			V (159)	N (1116)	
136	L (1186)			L (189)	A (1411)	
137	A (1268)			A (151)	V (1157)	
138	E (930)			E (100)	A (1267)	
139	K (740)	**K-9**			H (253)	
140		Y (2427)			A (1722)	
141		R (620)			L (787)	
142					A (1019)	
143					L (804)	
144	N.D.; not determined				G (523)	
145					Y (545)	
146					H (41)	

総　説

爬虫類のヘモグロビン：ゾウガメ (*Geochelone gigantea* と *Geochelone nigra*) グロビンの分子進化

宍　倉　文　夫
日本大学医学部自然科学系生物学

Reptilian Hemoglobin: Globin Evolution of the Two Giant Tortoises, *Geochelone gigantea* and *Geochelone nigra*

Fumio SHISHIKURA

Department of Biology, Nihon University School of Medicine

Amniota (reptiles, birds, and mammals), in general, have two or more hemoglobin components that are expressed according to the demands of different physiological conditions. Among them, hemoglobin A (Hb A) and hemoglobin D (Hb D) were detected in the extant giant tortoises (*Geochelone gigantea* and *Geochelone nigra*), and the latter was first found in birds as a minor component of the embryonic and adult definitive erythrocytes. Using computer-assisted analysis, a molecular tree was constructed on primary structures of 53 globins from Amniota including 38 reptilian α- and non α-globins. The divergence time between the two giant tortoises was estimated at 21–15 million years (myr) ago, which represents a significant lapse following the break up of Gondwana (formed by Africa, South America, and the Atlantic Ocean) at 65 myr ago. Hence, it is difficult to determine the place of origin of the two species. However, it is conceivable that the two extant giant species had diversified less than 21–15 myr ago from a common ancestor and thereafter, reached oceanic islands, one being volcanic (the Galapagos archipelagos) and the other being an atoll (the Aldabra atoll), by rafting, this is the most reliable mode of migration for terrestrial animals.

Key words: hemoglobin, Amniota, reptile, giant tortoise, molecular evolution
ヘモグロビン，羊膜類，爬虫類，ゾウガメ，分子進化

(J. Nihon Univ. Med. Ass., 2002; 61 (8): 263–276)

はじめに

近年，分子進化学の発展によって，生物の系統をタンパク質レベル・遺伝子レベルで検証できるようになった[1]．筆者は，環形動物 (Annelida) ミミズの巨大ヘモグロビン[2~6]と爬虫類 (Reptilia) ゾウガメ (海洋島に棲息する大形のリクガメの呼称) のヘモグロビン[7,8]を分子指標に，生物の多様化にともないグロビンの構造がどのように精緻化し，その構造に基づいてヘモグロビンの生理的機能がどのように多様化しているのかを明らかにしたい．また，同時に，分子進化学的証拠から生物の多様化の歴史を明らかにしたいと考えている．

ヘモグロビンは，動物・植物・原生生物・かび・細菌など生物界に広く分布している主要な呼吸色素タンパク質である[9~13]．ヘモグロビンは結合した酸素を各組織へ運搬・供給する重要な生理的機能を担っている[14,15]．生物がどのようなニッチ (niche；生態的地位) に生活しているかによって，その生物がもつヘモグロビンの構造と生理機能にはニッチの生物学的・非生物学的影響が顕著に現われている[16~18]．ヘモグロビンの構造と機能の解析は，ヒトの場合，疾病の解析という医学的必要性から詳細に研究され，一次構造 (アミノ酸配列) や遺伝子構造をとりあげても，膨大な情報が蓄積されている[14,15,19~21]．爬虫類の場合，17 種の爬虫類から 38 個のグロビンの一次構造が決定された．この中には海産のカメ[22]と淡水産のカメ[23,24]と陸生のカメ[7,8,25]からの報告が含まれている．本稿の主題であるゾウガメのグロビンについても，筆者らの研究[7,8]があり，比較的豊富なデータが揃っている．しかしながら，グロビン遺伝子の構造に関しては今のところ 3 種類のカメ[8,25]で解析されただけであり，この分野における爬虫類のデータは脊椎動物の中でも極めて少ない．

受付：2002 年 9 月 3 日，　受理：2002 年 9 月 13 日

本稿では，ゾウガメのグロビンの一次構造および最近明らかにされたグロビン遺伝子の構造を手がかりに爬虫類とその他の有羊膜類 (Amniota：爬虫類・鳥類・哺乳類を含む) グロビンの分子進化について，現在までの知見を総括し，現在2種類[26,27]しか生息していないゾウガメの種分化について解説したい．

I. 爬虫類の系統分類学的位置と地理的分布

系統分類学的位置

多様な脊椎動物群は，発生の途上で羊膜 (および漿膜・尿膜) を生ずるグループとそれをもたないグループとに2段階に分けることができる．羊膜をもつグループを有羊膜類 (Amniota) という．このグループは一生の生活環を通じてほとんど陸上 (二次的に水中にもどった種もいる) で生活し，肺で呼吸することが特色である．有羊膜類には爬虫類・鳥類・哺乳類が含まれ，無羊膜類 (Anamnia) と対比される．魚類・両生類は無羊膜類の動物である[28,29]．また，脊椎動物群の多様化の段階を四足 (肢) 類 (Tetrapoda；3節からなる四肢をもつグループ) とその他のグループに分ける場合もある．四足類には両生類・爬虫類・鳥類・哺乳類が含まれ，この場合対比される脊椎動物は魚類である．いずれにしろ，おもに比較形態学ならび比較発生学などの証拠から生物の分類がおこなわれてきた[29~30]．

爬虫類の古い化石は，ノバスコシア (北アメリカ) の石炭紀後期の地層から発見されている[31~33]．従って，最初の爬虫類は石炭紀後期に両生類中の迷歯類の一部 (炭竜類) から多様化したと推定される[32,34]．爬虫類は，中生代全般にわたって陸上・水中・空中のすべての環境に著しく適応放散して繁栄し，さまざまなニッチを占めたが，中生代末に急激に衰退し，白亜紀の終わり (6500万年前) にそのほとんどが絶滅してしまった[32,34] (恐竜は絶滅した爬虫類として有名な例である．恐竜の骨組織からヘモグロビン様タンパク質も検出されている[35])．現生の爬虫類 (Reptilia) は，カメ類 (Chelonia)，トカゲ・ヘビ類 (有鱗類：Squamata)，ワニ類 (Crocodylia) とムカシトカゲ (喙頭類 Rhynchocephalia の唯一の生き残り種) の4目 (order) が残存しているに過ぎない[29,31,34,36]．しかし，脊椎動物の多様化を調べるとき，現生の爬虫類は重要な系統分類上の位置にいる．例えば，鳥類は爬虫類双弓亜綱主竜類 (ワニ類と恐竜類とを含む) から，哺乳類は爬虫類の単弓類から派生したものであることが古典的な研究方法で明らかにされている[32,34]．

カメ類は頚部を曲げる習性をもっている．これには2つのタイプが見られ，頚部を横に曲げて甲羅の中に隠すタイプと頚部をまっすぐにしたまま甲羅の中に引き入れて隠すタイプである．前者のグループを曲頸類 (Pleurodira) といい，後者のグループを潜頸類 (Cryptodira) という[31,34]．いずれも白亜紀に出現し，北アメリカを中心として多くの種に分化し，世界中に分布するようになった．現生の曲頸類はすべて水中生活をおこない，南米・アフリカ・セーシェル諸島・オーストラリア・ニューギニアに分布している[32]．一方，ゾウガメ (giant tortoise) は潜頸類のリクガメ科 *Geochelone* 属の種 (species) である．*Geochelone* 属はリクガメ科 (Testudinoidae：10属41種) の中で最も大きなグループでリクガメ科の約半数の現生種を含む．*Geochelone* 属は始新世 (5200～3800万年前) に現れ，化石は北米・南米・アジア・ヨーロッパ・アフリカなど広範囲の地域から知られている[33,34]．ゾウガメは，マダガスカルやインド北部，アジアにも生息していたが，とくにインド洋の海洋島に生息していたゾウガメは19世紀まで生き続けていた．しかし，ヒトの活動が地球的規模で拡大したためこれらのゾウガメは消滅してしまった[37~39]．従って，現生のゾウガメはアルダブラゾウガメ (*Geochelone gigantea: Dipsochelys dussumieri* は synonym) とガラパゴスゾウガメ (*Geochelone nigra*) のみである[26,27]．Fig. 1 はチャールス・ダーウィン研究所 (Charles Darwin Research Station) で飼育されている Lonesome George の飼育舎の案内板に見つけた挿し絵で，その右に George は1971年 (M. Cruz 氏によれば，1972年) にピンタ (Pinta) 島で捕獲された最後の雄で，同じ種族 (race) の雌の個体が発見されない限りこの種族が永遠に途絶えてしまうことなどが書き留められていた．

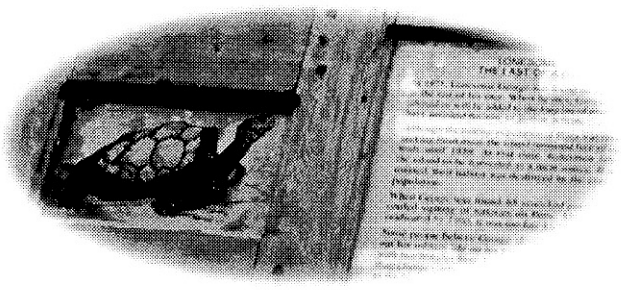


Fig. 1 Lonesome George, the last of a dying race. Lonesome George is the most famous giant tortoise of the Charles Darwin Research Station at Puerto Ayora, Santa Cruz. The wooden information tells us a story of Lonesome George discovered in 1971.

地理的分布

リクガメ科の現生種の地理的分布は Fig. 2 の通りである．オーストラリアを除く全世界の亜熱帯，熱帯地域に分布している．ゾウガメの分布は Fig. 2 に白抜きの矢印で示した．Fig. 2 でみる通り，ゾウガメの生息地は2大陸 (南アメリカとアフリカ) と大西洋で遠く隔たれた2つの海域に限られ，陸にすむ動物が「いつ」「どこから」「どのように」現在の生息地に渡ってきたのか．ヘモグロビンの構造を調べることによってゾウガメの歴史を明らかにしたい．

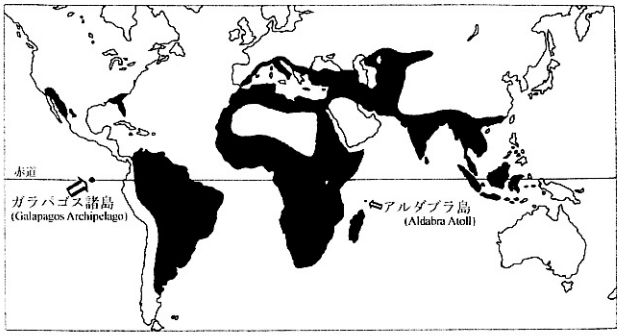


Fig. 2 The geologic and geographic distribution of land tortoises known as living species[33,40]. Open allows indicate the habitat of *Geochelone gigantea* (the larger one) and that of *Geochelone nigra* (the smaller one), respectively. The habitats of the two giant tortoises are remote and isolated. After the permission of the publisher, Nakayama Shoten Co., Ltd.

II. 爬虫類のヘモグロビン：Hb A と Hb D

爬虫類の赤血球は有核細胞で，血液量の29～35%を占める[41,42]．赤血球中のヘモグロビン量は7.8～12 g/dl[41,42]あり，その中に2種類のヘモグロビンが存在する．2種類のヘモグロビンは，哺乳類の成体ヘモグロビン (Adult hemoglobin) と鳥類の成体と胚胎期の赤血球中で副次的分画として最初に発見されたヘモグロビン D[43] (Definitive hemoglobin) の名称にちなみ，それぞれ Hb A と Hb D とに命名された[24,44,45]．爬虫類の Hb A は脊椎動物の成体に一般的に見られものと同様に，2種類のグロビンサブユニットからなる4量体 ($\alpha\beta$ プロトマー protomer の2量体ともいう) として存在する．ゾウガメの2種類のヘモグロビン，Hb A と Hb D，は一次構造 (アミノ酸配列) が異なる α-グロビンと一次構造 (アミノ酸配列) が同じ β-グロビンの $\alpha\beta$ プロトマーの2量体である．従って，後述するように，α-グロビンには α^A と α^D の2種類を認めるが，β-グロビンは1種類しかないことになる．当然，Hb A と Hb D の4量体はそれぞれ $\alpha^A{}_2\beta_2$ と $\alpha^D{}_2\beta_2$ となる．爬虫類の他の目 (Rhynchocephalia や Squamata) には，アミノ酸配列が異なる β-グロビンからなる $\alpha\beta$ プロトマーの2量体も見られる．この場合の β-グロビンは，β-I, β-II とよばれている[46～51]．いずれにしろ，爬虫類のおもな目 (order) の種には2種類のヘモグロビンが赤血球中に存在し[7,23,24,46～53]，著者らによれば，それらの赤血球中での存在比は約 5:3 (Hb A : Hb D) となっている[7]．

2種類のヘモグロビンが血球中に存在していることは，どのような生物学的意義があるのだろうか．生物の環境への適応とタンパク質の分子進化という興味深い課題として研究されている．生物の環境への適応に関しては，鳥類の Hb D で調べられている．鳥類は海中から高度数千メートルに至る広いニッチに生活しているので，ニッチの環境要因が鳥類のヘモグロビンの生理機能やグロビンの分子構造に反映しているものと推測できる．事実，ニワトリの Hb D は胚胎期の低酸素条件でも高い酸素親和性を示した[54,55]．また，アンデスコンドル[56]やペンギン[57]あるいは潜水性のカメ[22,58]のヘモグロビンやその他の爬虫類のヘモグロビンについて生理的機能に関する研究報告がある[59～62]．特異な環境に生息する生物がもつ生体のタンパク質分子の構造と機能解析は，ヒトの血液生理学に新しい展開と貴重な情報をもたらすであろう．現実には，有羊膜類 (Amniota) ヘモグロビンの構造・機能の多様化および分子進化の解析に手がかりを与えるものと期待されている．

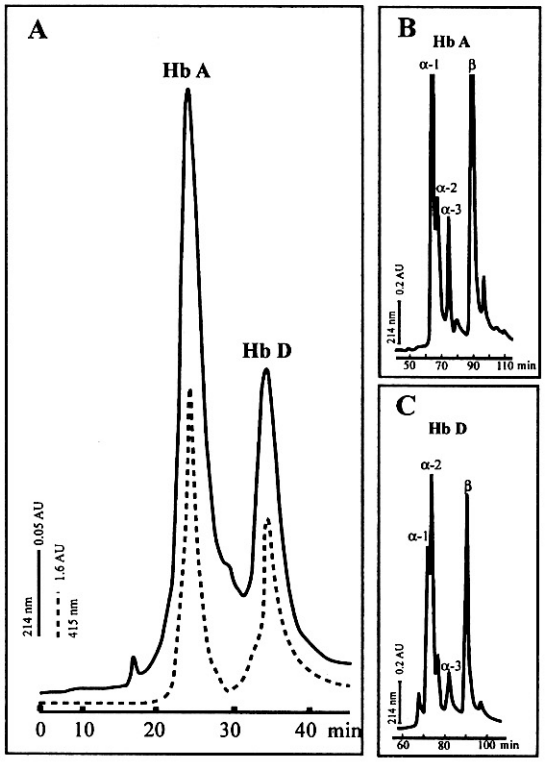


Fig. 3 Separation and purification of the *G. gigantea* globin-chains[7,8]. *Panel* A: Separation of Hb A and Hb D from *G. gigantea* red blood cells on an Alkyl Superose HR5/5 column. *Panels* B and C: Separation of globin-chains from reduced and *S*-pyridylethylated Hb A (*Panel* B: about 2 mg globin) and Hb D (*Panel* C: about 2 mg globin) on a Resource column. Major peaks are designated as α-1, α-2, α-3 and β, respectively.

III. グロビンの分離・精製

多くのタンパク質の生理的機能は，複数のポリペプチド (サブユニット) が非共有性に結合した四次構造を構築して発現される．上述したように，脊椎動物のヘモグロビンは2種類のグロビンサブユニットがそれぞれ2個集まり4量体 ($\alpha_2\beta_2$) の四次構造をつくる．

まず，2種類のヘモグロビン，Hb A と Hb D，の精製が試みられた[24,47,59～61]．Bonila らの方法[61]では無変性

```
The Galapagos Giant Tortoise
  G. nigra  αA-1      -VLTAGDKANVKTVWSKVGSHLEEYGSETLERLFIVYPSTKTYFPHFDLHH------DSAQVRAHGRKVL
  G. nigra  αD-2      -MLTEDDKQLIQHVWETVLEHQEDFGAEALERMFTVYPSTKTYFPHFDLHH------GSEQIRHHGKKVV
  G. nigra  β         VHWTPEEKQYITSLWAKVN--VEEVGGEALARLLIVYPWTQRFFSSFGNLSSASAILHNAKVLAHGKKVL
The Aldabra Giant Tortoise
  G. gigantea αA-1    -VLTAGDKANVKTVWSKVGSHLEEYGSETLERLFVVYPSTKTYFPHFDLHH------DSPQVRAHGKKVL
  G. gigantea αD-2    -MLTEDDKQLIQHVWEKVLEHQEDFGAEALERMFIVYPSTKTYFPHFDLHH------DSEQIRHHGKKVV
  G. gigantea β       VHWTSEEKQYITSLWAKVN--VGEVGGEALARLLIVYPWTQRFFASFGNLSSANAILHNAKVLAHGQKVL
                        *   *      *  *      * * * *   *** *   *  *                  ** **

SALGEAVNHIDDIPGALSKLSDLHAQTLRVDPVNFKLLNLCFVVVVGRHHPTILTPEVHVSLDKFLSAVATALTSKYR
GALGDAVRHIDDLSATLSELSNLHAYNLRVDPVNFKLLSHCFQVVLGAHLGREYTPQVQVAYDKFLAAVSAVLAEKYR
TSFGDAVKNLDNIKKTFAQLSELHCEKLHVDPENFKLLGNILIIVLATRFPKEFTPASQAAWTKLVNAVAHALALGYH

SALGEAVNHIDDIPGALSKLSDLHAQNLRVDPVNFKLLNLCFVVVSGTHHPTILTPEVHVSLDKFLSAVATALTSKYR
GALGDAVKHIDNLSATLSELSNLHAYNLRVDPVNFKLLSHCFQVVLGAHLGREYTPQVQVAYDKFLAAVSAVLAEKYR
TSFGEAVKNLDNIKKTFAQLSELHCEKLHVDPENFKLLGNILIIVLATHFPKEFTPASQAAWTKLVNAVAHALALGYH
   * **   *         ** **   * *** *****      *          **       *   **   *   *
```

Fig. 4 An alignment of primary structures of three kinds of globins (α^A, α^D, and β) from the two giant tortoises[7,8]. Asterisks indicate the invariant amino acid residues. Dashed lines are gaps automatically inserted to maximize the sequence alignment.

のヘモグロビンを精製するために，調整用等電点電気泳動法 (アガロースゲルにキャリアーとしてアンフォライン Ampholline を加えた) で分けたタンパク質バンドをアガロースゲルから切り出してタンパク質を抽出する．彼らはヘビ (Riotropical Racer: *Mastigodryas bifossatus*) から 2 種類のヘモグロビンを精製した．その等電点電気泳動パターンは，2 本の近接したバンド (等電点：8.02 と 8.07) を示すもので，バンドを切り出すときにコンタミの不安を払拭できない．近年，疎水性カラム (Alkyl Superose HR 5/5) を用いることにより，ゾウガメ赤血球のリゼート (赤血球抽出液) から，無変性の状態で，ヘモグロビンを精製することができるようになった (Fig. 3A). 筆者らは，この方法で精製したヘモグロビンから結晶を作成することに成功している．この方法は，Hb A および Hb D の立体構造を構築するために重要な精製方法になるものと思われる．

次いで，2 種類のヘモグロビン (Hb A と Hb D) からそれぞれの $\alpha\beta$ プロトマーを構成するグロビン (α^A, α^D, β) を単離するため，カラムクロマトグラフィーをおこなった．とくに，ヘモグロビンをピリジルエチル化[63]することで 4 量体を解離させ，簡単に精製できた．Hb A と Hb D の α-グロビンと β-グロビンの分離・精製パターンを Fig. 3 (B と C) に示す．α-グロビン分画には 3 つのピークが認められた．そのうち，α-1 と α-2 のアミノ酸配列を決定[7]して比較したところ，数残基のアミノ酸の置換があり，これらのアミノ酸の違いが複数の α-グロビンピークの原因であることが示唆された．アミノ酸の置換 (microheterogeneity) が $\alpha\beta$ プロトマー 2 量体の形成にどのように影響しているかまだ解決されていない．β-グロビンは共通に使われているとしても，α-グロビンとの組合せがさらに複数存在し，それぞれの組合せによって異なる生理的機能を発現しているのだろうか．ともかく，このカラム法で $\alpha\beta$ プロトマー 2 量体 ($\alpha_2\beta_2$) のそれぞれのサブユニット，α-グロビンと β-グロビン，を分離・精製できた．なお，グロビンの精製方法については，多種多様な方法[64,65]が報告されているが，本法は爬虫類ゾウガメのグロビンを精製し，グロビンの一次構造から四次構造までの解析をおこなう目的に最もよい方法と思われる．

IV. グロビンの分子進化—ゾウガメを中心として

ゾウガメの α^A, α^D, β：一次構造

2 種類のゾウガメから精製した 3 種類のグロビン (α^A, α^D, β) の一次構造を決定し，それらに基づいてアライメントを作成した (Fig. 4)．α-グロビンと β-グロビンの類似度を比較すると，1) *G. gigantea* α^A 対 *G. nigra* α^A は 95.8%の相同；*G. gigantea* α^D 対 *G. nigra* α^D は 96.5 %の相同；*G. gigantea* β 対 *G. nigra* β は 95.2%の相同であった．また，2) 同種内および異種内の α^A 対 α^D の比較では，60～62.5%の相同であった．ところが，3) α-グロビン (α^A または α^D) と β-グロビンとを比較するとホモロジーは低い値 (35～40%) になった．同じタイプのグロビンの比較では極めて高いホモロジーを示し，2 種類のゾウガメは互いに近縁な生物であることが示唆されたが，異なったタイプのグロビンの比較 (α^A：β, α^D：β) で得られた低い値は，α (α^A と α^D) と β がゾウガメの種分化以前にすでに分岐していたこと，すなわち，2 つのグロビングループ (以下では α-グロビンタイプを α-like globin と β-グロビンタイプを non α-like globin という) は非常

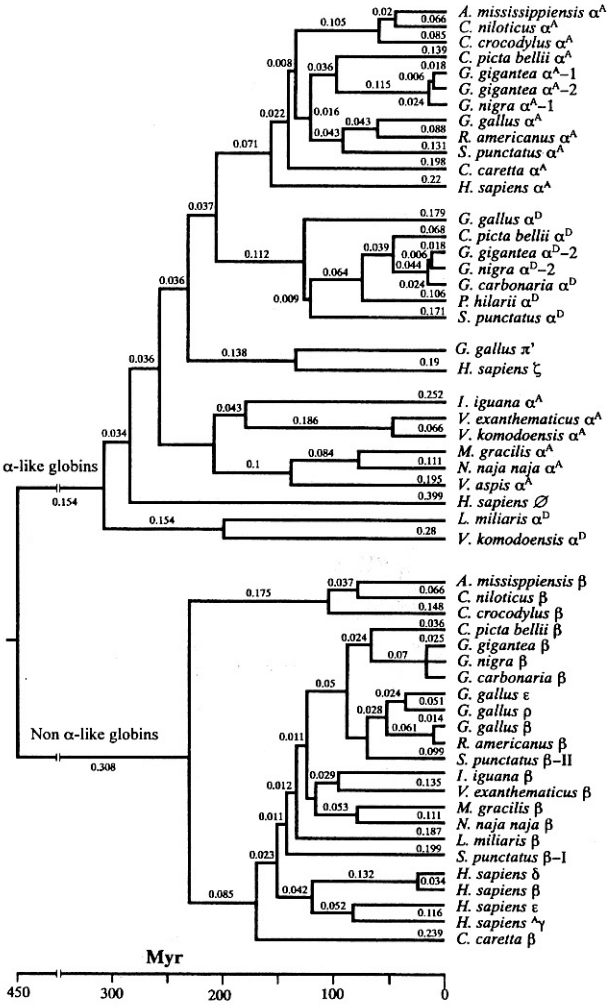


Fig. 5 The phylogenetic tree of 38 reptilian globins and the other representative Amniota globins. Branch lengths are proportional to protein distances and shown on the individual branches of the tree. The abscissa is a time scale in Myr (million years) ago based on the separations of the α- and β-globin described by Goodman et al[66]. The references of primary structures of globins used in the present analysis are as follows: α^A-globins: *A. mississippiensis*[69], *C. niloticus*[69], *C. crocodylus*[69,70], *C. picta bellii*[23], *G. giantea* α-1[7] and α-2[7], *G. nigra* α^A-1 (Accession Nos. P83131, P83132 and P83135 of Swiss-Prot Data Bank), *G. gullus*[71], *R. americanus*[72], *S. punctatus*[47], *C. caretta*[22], *H. sapiens*[73], *I. iguana*[49], *V. exanthematicus*[50], *V. komodoensis*[51], *M. gracili*[74], *N. naja naja*[75], *V. aspis*[46]. α^D-globins: *G. gullus*[76], *C. picta vellii*[24], *G. gigantea* α^D-2[8], *G. nigra* α^D-2[8] (Accession No. P83124 of Swiss-Prot Data Bank), *G. carbonaria* (Accession No. AF304335 of GeneBank), *P. hilarii*[24], *S. punctatus*[47], *L. miliaris*[53], *V. komodoensis*[51]. α-like embryonic globins: *G gallus* π'[77], *H. sapiens* ζ[78], *Homo sapiens* θ[79]. Non α-globin: *A. mississippiensis* β[69], *C. niloticus* β[69], *C. crocodylus* β[69,70], *C. picta bellii* β[24], *G. gigantea* β[7], *G. nigra* β (Accession No. P83123 of Swiss-Prot Data Bank), *G. carbonaria* β[25], *G. gallus* ε[80], *G. gallus* ρ[81], *G. gallus* β[82], *R. americanus* β[72], *S. punctatus* β-II[47], *I. iguana* β[49], *V. exanthematicus* β[50], *M. gracilis* β[74], *N. naja naja* β[75], *L. miliaris* β[53], *S. punctatus* β-I[47], *H. sapiens* δ[83], *H. sapiens* β[84], *H. sapiens* ε[85], *H. sapiens* $^A\gamma$[86], *C. caretta* β[22].

に古い時代に別れたタンパク質ファミリーであることが示唆された．定説では，αとβは4億5000万年前～5億年前に分かれたと考えられている．また，α^A対α^Dの比較によって得られたホモロジー値(35～40%)は，α^Aとα^Dのようにα-likeglobinを2つの亜系統に区分けできる程度にそれらが多様化していることを示唆している(詳しくは，次の項参照)．

有羊膜類グロビンの近縁関係

グロビンはタンパク質化学的情報が豊かな生体分子である．爬虫類のグロビンも例にもれず，一次構造(アミノ酸配列)は4目すべてから複数の報告がある．しかも，上述したように，近縁な種同士でも異なったタイプのグロビンは極めて低いホモロジー値を示した．従って，タンパク質に翻訳されそれ自身重要な生理的機能を発現している生体分子を指標に生物の分子系統関係を解析する場合，グロビン・ファミリーは生物の多様化の歴史を追跡できる好適なタンパク質といえる[21,66]．この場合，グロビンはα-like globinとnon α-like globinのどちらを採用しても貴重な情報が得られるはずだ．

脊椎動物有羊膜類のα-like globinはα^A-, α^D-, β-, π-, θ-, ζ-グロビン，non α-like globinはβ-, δ-, ε-, ρ-, γ-グロビンなどが知られている．哺乳類では，これらのグロビンタンパク質をコードしているグロビン遺伝子がどの(何番の)染色体にどのような順序で配列しているかについても解析が進んでいる[19,20]．鳥類と爬虫類の赤血球は脊椎動物のグロビンの中で特異なα^D-グロビンをもっているが，爬虫類からはカメ類・トカゲ・ヘビ類・喙頭類から報告[51,52]されているだけで，今のところ，ワニ類からの報告はない．

まず，爬虫類の38種類のグロビンと他の有羊膜類の代表的グロビンを加えて総数53種類のグロビンのアライメントをCLUSTAL W[67]で作成した．次に，PROTDIST[68]でグロビン間の遺伝的距離を計算し，NEIGHBOUR[68]で分子系統樹を作成した(Fig. 5)．爬虫類以外の有羊膜類のグロビンと比較するため，この系統樹の中には哺乳類からヒト(*Homo sapiens*)のグロビン，鳥類からニワトリ(*Gallus gallus*)とレア(*Rhea americanus*)のグロビンをouter groupsとして採用した．なお，脊椎動物の魚類および両生類は有羊膜類に分類されていないので本稿から削除した．

この分子系統樹を見ると次の①～⑨が示唆された．A) ①グロビン系統樹には2つの分枝，すなわち，α-like globinとnon α-like globinが認められた(既成の事実の再確認)．B) α-like globinの枝に関して；B-1) ②α^A-グロビンとα^D-グロビンと胚胎グロビン(π', ζ)はそれぞれクラスターを形成した．B-2) ③有羊膜類の中で，トカゲ・ヘビ類は別の分枝になった．B-3) ④ヘビ(*L. miliaris*)とコモドオオトカゲ(*V. komodoensis*)のα^D-グロビンはこ

Table 1 Comparison of exson and intron sizes (in bp) among representative amniotes (reptiles, birds and mammals)-globin genes

Class	Globin-gene Name	Exon-1	1st Intron	Exon-2	2nd Intron	Exon-3	Ref.
Reptilia	***G. gigantea* α^{A}-2**	**92**	**74**	**205**	**340**[1)]	**126**	unpub.
	***G. gigantea* α^{D}-2**	**92**	**214**	**205**	**227**	**126**	8
Aves	*Gallus gallus* α^{D}	92	148	205	261	126	87
	Gallus gallus π'	92	577	205	294	126	88
	Gallus gallus α^{A}	92	131	205	109	126	87
Mammalia	*Homo sapiens* ζ	92	886	205	239	126	89
	Homo sapiens α	92	117	205	141	126	90
Reptilia	**G. gigantea β**	**89**	**130**[1)]	**223**	**1590**[1)]	**126**	8
Aves	*Gallus gallus* β	89	92	223	810	126	91
	Gallus gallus ρ	92	108	223	541	126	92
	Gallus gallus ε	92	108	223	973	126	80
Mammalia	*Homo sapiens* ε	92	122	223	854	126	85
	Homo sapiens $^{A}\gamma$	92	122	223	885	126	86
	Homo sapiens δ	89	128	223	886	126	93
	Homo sapiens β	89	130	223	850	126	84

[1)] Intron sizes were estimated by migration distances of PCR-amplified fragments on agarose gel electrophoresis (Fig. 6).

のα-like globin の枝から最も初期 (3 億年前) に分かれた．また，B-4) ⑤ヒトの θ-グロビンは約 3 億年前に他のグロビン系から分岐した (今，考察する資料を持ち合わせていない)．C) non α-like globin の枝に関して；C-1) ⑥各グロビン ($\beta, \varepsilon, \rho, \delta, \gamma$など) は種 (species) または綱 (class) ごとにクラスターを形成した．C-2) ⑦爬虫類のワニは他の有羊膜類から約 2 億 3000 万年前に分岐した．D) その他；D-1) ⑧古典的系統類縁関係と比較するとグロビンの一次構造に基づく分子系統関係は，従来の分類体系をおおよそ支持する生化学的根拠を与える結果になった．D-2) α-like globin と non α-like globin の分岐年代を 4 億 5000 万年前[66]とすると，⑨2 種類のゾウガメは 2100～1500 万年前に種分化したことが推定された (個々の有羊膜類の分岐年代については Fig. 5 を参照)．

V. グロビン DNA の塩基配列

爬虫類のグロビン遺伝子の構造

グロビン遺伝子の構造研究は，おもに，哺乳類と鳥類で進められている[14,15]．爬虫類のグロビン遺伝子の cDNA またはゲノム DNA の塩基配列が決定された例は以下に述べるように最近のことである．例えば，1997 年，アカアシガメ (*Geochelone carbonaria*) の赤血球から β-グロビン mRNA を精製し，その cDNA 塩基配列を決定した報告[25]が爬虫類で最初の論文となった．また，同じ研究グループから同じ種 (*G. carbonaria*) の α^{D}-グロビン cDNA の塩基配列が決定され，データバンクに登録 (Accession No. AF304335 GeneBank) されている．ゾウガメの α^{A}, α^{D}, βのゲノム DNA の塩基配列は筆者によって 2002 年に報告[8]された．このように，爬虫類のグロビン mRNA およびゲノム DNA の構造解析は端緒についたばかりといえる．ここでは，筆者の研究をもとに，ゾウガメのグロビン遺伝子の構造を他の有羊膜類のものと比較して解説したい．

2 イントロン 3 エキソン構造

1970 年代のはじめ，組換え DNA 技術 (recombinant DNA technology) が導入されるとただちにグロビンの遺伝子構造が解析され，ヒトのグロビンは 2 イントロン 3 エキソンであることが明らかにされた．Table 1 に有羊膜類のグロビン遺伝子のイントロンとエキソンに含まれる塩基数を示す．現在では例外も知られている[94,95]が，2 エキソン 3 イントロンの枠組みはほとんどの生物のグロビンに共通の遺伝子構造と認められている[96,97]．筆者は 2 種類のゾウガメの赤血球からゲノム DNA を抽出・精製し，グロビンの遺伝子の塩基配列 (エキソン構造はもちろんイントロンの構造を含む) を決定した．Table 2 にゾウガメのグロビン遺伝子の塩基配列を示す (*G. nigra* のグロビン遺伝子の塩基配列は未発表のため省略)．

塩基配列を決定するとき，筆者が用いた方法の概略は次の通りである (この方法は，他の生物のグロビン遺伝子の研究に汎用できるので略記したい)．1) 一次構造の研究から明らかにされたアミノ酸配列からプライマー (Table 3) を作成し，ゲノム DNA を鋳型にして PCR 法で目的 DNA 断片を増幅する．この原法は，MOPAC (mixed oligonucleotides primed amplification of cDNA[98]) 法といわれている．次いで．2) アガロースゲル電気泳動法で目的 DNA 断片を分離し，切り出す．次に，スピンカラム法でアガロースゲルから目的 DNA 断片を抽出す

Table 2-A Nucleotide sequences of three exons and exon-intron boundaries of *Geochelone gigantea* α^{A}-2-globin gene

Exon-1	***1***	GTG	CTA	ACA	GCA	GGC	GAC	AAG	GCC	AAC	GTG	AAG	ACC	GTG	TGG	AGC
	46	AAG	GTG	GGC	AGC	CAC	CTG	GAG	GAC	TAT	GGC	TCC	GAG	ACC	CTG	GAG
	91	AG														
Intron-1	***1***	gta	aca	gcc	gct	gcg	ccc	tgc	ccc	gcc	tgc	gcc	cgg	cca	ggc	ccc
	46	cgc	tta	acc	cgc	ctt	gtt	ctt	tgc	ag						
Exon-2	***1***	G	CTG	TTC	GTC	GTC	TAC	CCC	TCC	ACC	AAG	ACC	TAC	TTC	CCC	CAC
	44	TTC	GAC	CTG	CAC	CAC	GAC	TCC	CCC	CAG	GTC	CGG	GCC	CAC	GGC	AAG
	89	AAG	GTG	CTG	AGC	GCC	CTG	GGG	GAA	GCC	GTG	AAC	CAC	ATC	GAT	GAC
	134	ATC	CCC	GGG	GCT	CTC	AGC	AAA	CTG	AGC	GAC	CTG	CAC	GCC	CAQ	AAC
	179	CTG	CGC	GTG	GAT	CCC	GTC	AAC	TTC	AAA						
Intron-2	***1***	gtg	agt	gac	cgg	cag	aaa	tgc	tcc	ctc	cgg	gtc	cga	gcc	cgc	tgc
		(approximately 250 nucleoitdes not shown)														
		gca	gcc	ggt	tcc	cag	ctt	ttc	tca	ctc	gcg	cgc	ctt	gtc	ttc	cag
Exon-3	***1***	CTG	CTG	AAC	CTG	TGC	TTC	GTG	GTG	GTC	GTG	GGC	CGC	CAC	CAC	CCC
	46	ACC	ATC	CTC	ACC	CCC	GAG	GTC	CAC	GTG	TCC	CTG	GAC	AAG	TTC	CTG
	91	AGC	GCT	GTG	GCC	ACC	GCG	CTS	ACT	AGT	AAG	TAC	CGG			

Table 2-B Nucleotide sequences of three exons and exon-intron boundaries of *Geochelone gigantea* α^{D}-2-globin gene[8)]

Exon-1	***1***	ATG	CTA	ACA	GAG	GAC	GAC	AAG	CAG	CTG	ATC	CAA	CAT	GTG	TGG	GAG
	46	AAG	GTG	CTG	GAG	CAC	CAG	GAG	GAC	TTT	GGG	GCC	GAG	GCC	CTG	GAG
	91	AG														
Intron-1	***1***	gta	ggg	ccc	ggg	gca	ggc	ggc	ccg	ggc	gca	ggg	tgc	agg	gag	gaa
	46	(46–168 not shown)														
		tgg	ggg	act	ttg	gga	ttc	act	gtc	tct	gac	ctc	cct	ccc	ccg	cag
Exon-2	***1***	G	ATG	TTC	ATC	GTC	TAC	CCC	TCC	ACC	AAG	ACC	TAC	TTC	CCC	CAC
	44	TTC	GAC	CTG	CAT	CAT	GAC	TCG	GAA	CAG	ATC	CGC	CAC	CAC	GGC	AAG
	89	AAG	GTG	GTG	GGC	GCC	CTG	GGG	GAC	GCC	GTG	AAG	CAC	ATC	GAC	AAC
	134	CTC	AGC	GCG	ACG	CTC	TCC	GAG	CTC	AGC	AAC	CTG	CAC	GCC	TAC	AAC
	179	TTG	CGC	GTG	GAC	CCG	GTC	AAC	TTC	AAG						
Intron-2	***1***	gc[1)]a	agt	gca	ggc	tac	ggc	cag	gaa	gag	ttc	ccg	ggg	ggt	gcg	gga
		(46–181 not shown)														
	182	ggc	cga	ggg	ctg	gct	gcc	gct	gac	cca	gtg	cac	ttt	gct	ttg	cag
Exon-3	***1***	CTG	CTG	TCC	CAC	TGC	TTC	CAG	GTG	GTG	CTG	GGC	GCG	CAC	TTG	GGC
	46	CGC	GAG	TAC	ACC	CCG	CAG	GTG	CAA	GTC	GCC	TAT	GAC	AAG	TTC	CTG
	91	GCC	GCC	GTC	TCG	GCG	GTG	CTG	GCT	GAG	AAG	TAC	CGG			

Table 2-C Nucleotide sequences of three exons and exon-intron boundaries of *Geochelone gigantea* β-globin gene[8)]

Exon-1	***1***	GTG	CAC	TGG	ACC	AGC	GAG	GAG	AAG	CAG	TWC	ATT	ACC	AGT	CTG	TGG
	46	GMC	AAG	GTC	AAC	GTG	GRG	GAA	GTG	GGT	GGC	GAA	GCC	CTG	GCC	AG
Intron-1	***1***	gta	ggc	tcg	agc	ctc	aca	tg g	ata	tct	gcc	tcg	cat	tgc	t cc	tct
		(approximately 40 nucleotides not shown)														
		gca	gta	acc	ctg	tgt	ctg	tct	ctg	ctc	ctg	tct	ccc	tct	ctc	tag
Exon-2	***1***	G	CTG	CTG	ATC	GTC	TAC	CCC	TGG	ACC	CAG	AGG	TTT	TTC	KCT	TCC
	44	TTT	GGG	AAC	CTS	TCC	AGC	SCC	AAC	GCC	ATC	MTG	SRC	AAC	GCC	AAG
	89	GTG	STT	GCC	CAT	GGC	MAG	AAA	GTG	CTG	ACC	TCS	TTT	GGG	GAA	GCT
	134	GTG	AAG	AAC	CTG	GAC	AAC	ATC	AAG	RMM	ACG	TWC	GCC	CAG	CTG	AGC
	179	GAG	CTG	CAC	TGC	SAR	AAG	CTG	CAT	GTG	GAT	CCT	GAG	AAC	TTC	AAG
Intron-2	***1***	gtg	agt	ccg	gct	ctg	ggt	tga	ccc	tct	tcc	cag	ccc	cct	ttc	cat
		(approximately 1590 nucleotides not shown)														
		cag	agc	ggt	gct	gac	cca	gcg	ggt	atc	ttc	ttc	ctt	ctc	ctc	cag
Exon-3	***1***	CTC	CTG	GGC	AAT	ATC	CTC	ATC	ATC	GTC	CTG	GCC	ACC	CAC	TTC	CCA
	46	AAG	GAG	TTC	ACT	CCT	GCC	AGT	CAG	GCC	GCC	TGG	ACA	AAG	CTC	GTC
	91	AAT	GCA	GTG	GCC	CAT	GCT	CTG	GCT	CTC	GGT	TAC	CAC			

The International Union of Pure and Applied Chemistry Symbols used to denote multiple nucleotides are as follwos: M=A or C, R=A or G, S=G or C, W=A or T.

1) An abnormal gc dinucleotide instead of the invariant gt at the 5' end of the second intron sedquence.

Table 3 Oligo-nucleotide primers used for the amplification of *Geochelone gigantea* globin-genes[8]

Gene	Primer Name (Type)	Nucleotide sequences	Reference	Degeneracy
		1 10 20		(fold)
Hb α^A-2	M13F-1 (Forward)	M13-GTGCTRACNGCNGGNGAYAAGGC	N-terminal	256
	M13F-2 (Forward)	M13-AARACNTAYTTYCCNCAYTTYGA	40–47	512
	M13R-1 (Reverse)	M13-YCGGTACTTNSWRGTBAGVGC	C-terminal	576
Hb α^D-2	M13F-3 (Forward)	M13-ATGYTNACNGARGAYGAYAARCA	N-terminal	512
	M13R-2 (Reverse)	M13-AAYTTRTCRTANGCNACYTGNAC	C-terminal	1024
Hb β	M13F-4 (Forward)	M13-GTGCACTGGACYWSNGARGAGAAG	N-terminal	64
	M13R-3 (Reverse)	M13-CTTGAAGTTCTCRGGRTCCACRTG	104–97	8
	M13F-5 (Forward)	M13-CAYGTGGAYCCYGAGAACTTCAAG	97–104	8
	M13R-4 (Reverse)	M13-GTGGTAVCCSAGRGCCAGRGCRTG	C-terminal	48

M13 forward sequence: 5'-TGT AAA ACG ACG GCC AGT-3', M13 reverse sequence: 5'-CAG GAA ACA GCT ATG ACC-3'
The International Union of Pure and Applied Chemistry Symbols used to denote multiple nucleotides are as follows: B = C, G or T not A; N = A, G, C or T; R = A or G; S = C or G; V = A, G or C not T; W = A or T; Y = C or T.

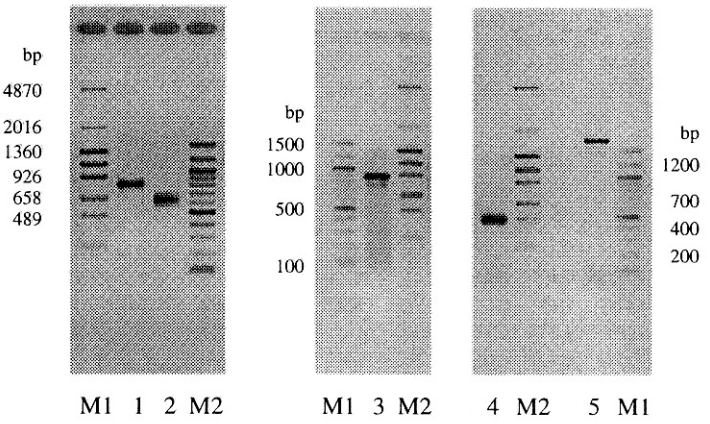


Fig. 6 Agarose gel electrophoreses of PCR products amplified from *G. gigantea* genomic DNA using degenerate primers (see Table 3)[8]. *Lane 1*; an 834-bp fragment amplified with M13F-1 and M13R-1, *Lanes 2*; a 645-bp fragment amplified with M13F-2 and M13R-1, *Lane 3*; a 870-bp fragment amplified with M13F-3 and M13R-2, *Lane 4*; a 480-bp fragment amplified with M13F-4 and M13R-3, *Lane 5*; a 1.75-kbp fragment amplified with M13F-5 and M13R-4, Lanes M1 and M2; DNA molecular standard markers, 100-bp DNA Ladder (M1) and pHY Marker (M2).

る．Fig. 6 は *G. gigantea* グロビン遺伝子の PCR 増幅パターンである (*G. nigra* の増幅パターンは未発表のため省略)．3) 市販の M13-forward プライマーまたは M13-reverse プライマーを用いて塩基配列をサイクルシーケンス (Dye Terminator) 法で決定する．本法により，エドマン法で決定したタンパク質のアミノ酸配列の信憑性と，また，遺伝子の塩基配列の信憑性とが著しく高まった．ただし，グロビンの mRNA から cDNA を作成し塩基配列を決定した場合と異なり，解析された遺伝子の塩基配列は翻訳領域とイントロン領域の構造のみとなっている．これは，翻訳されたタンパク質のアミノ末端とカルボキシル末端の配列情報を利用しているので止むをえない．翻訳領域の 5' 上流 (または，3' 下流) の情報を必要とする場合は，別の方法を加えて解析する必要がある．

Table 2 にエキソン部分の塩基配列とスプライシングされる配列部分のイントロン／エキソンの境界領域のみを示した．興味深いことは，通常のイントロンの両端に見られる共通配列 (consensus sequence)，すなわち 5'-GU/AG-3' (教科書的にはこの配列は 100%保存されている) に塩基の置換がみられた．α^D-グロビン遺伝子の第 2 イントロン 5'-端にある共通配列 5'-GU の U が C に置換されている．この現象は，鳥類の α^D-グロビン遺伝子の第 2 イントロン 5'-端にも認められ[99]，鳥類とカメ類の近縁関係を生化学的に支持する貴重な証拠資料となるだろう．ゾウガメ以外の爬虫類からグロビン遺伝子の構造が明らかにされていないので，この α^D-グロビン遺伝子の第 2 イントロン 5'-GC が爬虫類全体に共通配列として認められるのかどうか現時点では定かでない．

VI. ゾウガメの種分化

ゾウガメの生息地と大陸移動説

現生の爬虫類の分布は現在の大陸の地理的分布の状態

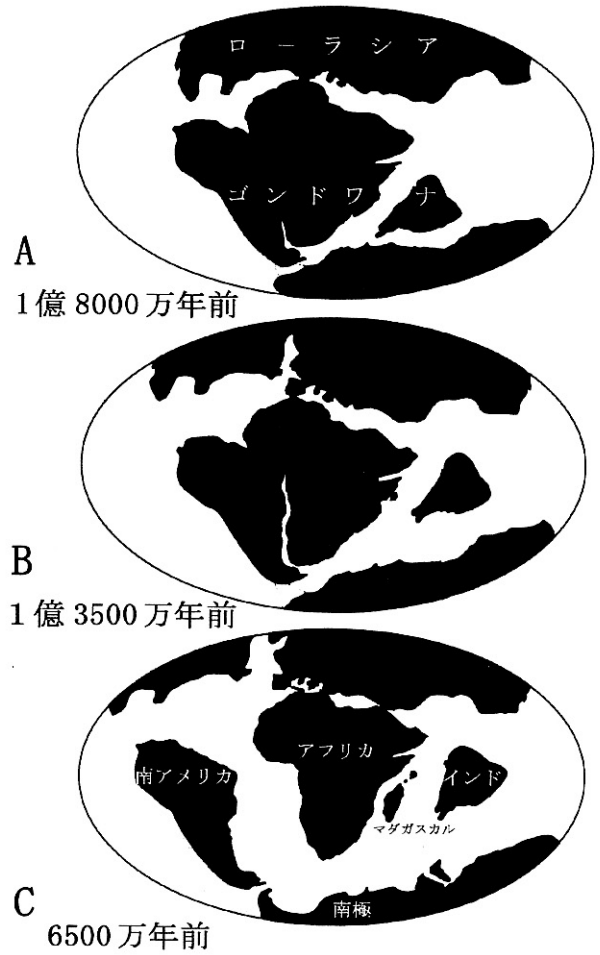


Fig. 7 The breakup of land mass, Gondwana[32, 104]. A: At the end of the Triassic period (180 myr ago), the southern group, known as Gondwana, has begun to isolate the Africa-South America land mass from Antarctica-Australia. B: At the end of the Jurassic period (135 myr ago), the North Atlantic and the Indian Ocean have opened considerably. A rift has initiated the birth of the South Atlantic. C: At the end of the Cretaceous period (65 myr ago), the South Atlantic has widened into a major ocean. A new rift has carved Madagascar away from Africa. After the permission of the publisher, Nakayama Shoten Co., Ltd.

から説明できない分布を示す例が多い[32]．アルダブラ島に生息するアルダブラゾウガメとガラパゴス諸島に生息するガラパゴスゾウガメも，両種が近縁とすると現在のそれぞれの生息場所から推測して，ゾウガメの地理的分布がなぜそのようになっているのか不思議に思う (Fig. 2)．

中央に一つの礁湖 (lagoon：約 30 km × 8 km) をもちその周囲を大小 4 つの珊瑚礁がリング状に取り囲むアルダブラ島は，アフリカ東海岸から約 640 km，マダガスカルの北西約 420 km の西インド洋に位置している．Braithwaite らの報告[100] によると，アルダブラ島は 125,000 年前頃に一度完全に海上から姿を消し，8 万年前に現在の珊瑚礁が出現した．その頃の島の大きさは 50 km^2 ほどでカメ，トカゲ，ヘビ，鳥などが移住していた．その後，もう一度沈没し，それまでの生物相が消滅したが，今から 13,000 年前に再び出現した珊瑚礁は 155 km^2 の広さであった．一方，ガラパゴス諸島はエクアドルから海上約 1000 km の赤道に位置し，海底火山の爆発によって出現した 14 の大きな島 (4.9 km^2～4,588 km^2) と多数の小島からなる火山群島である．その出現は，1100～500 万年前と推定されている[101～103]．1995 年 1 月フェルナンディナ (Fernandina) 島の噴火は私の今までの考え方を一蹴した．その翌々年，訪れたときはまだ白煙を上げている低い丘を船上から遠目に確認できた．ガラパゴスは，生物の研究者にとって，今も変化を続けている魅惑的な島であることにかわりがない．

「大陸はいつも固定した状態でなく移動している」という説は，ウェゲナー (A. Wegener) が 20 世紀初頭に提案した仮説であったが，今日では事実として受け入れられている[104]．2 種類の陸生のゾウガメが「いつ」,「どこで」,「どんな祖先種から」分岐して，「どのような方法」で現在の海洋島に生息するようになったのか．大陸移動説がヒントを与えてくれるかも知れない．

ウェゲナーの大陸移動説によれば，ゴンドワナ大陸はジュラ期の終わり (約 1 億 3500 万年前) に南方から入った亀裂 (南アメリカとアフリカ大陸との境をつくる) が拡大し，その大陸の北端が南アメリカとアフリカとを分けるように西方から侵入した湾と合一して，南アメリカを完全にアフリカから分離するようになった．この両者を分けた海峡はしだいに拡大し，やがてここに大西洋がつくられた (Fig. 7)．従って，白亜紀の始めに (1 億 3500 万年前) ゴンドワナ大陸の陸塊が分裂し始めてから少なくとも白亜紀の終わり (6500 万年前) 頃にほぼ現在の輪郭に近い構図ができたと考えられる．マダガスカルがアフリカ大陸から分離した時期もこのころであった[105]．しかし，前述したように，アルダブラ島で現在のような生物相が見られるようになるのはずーと後の 13,000 年前頃からのことである[100]．また，ガラパゴス諸島が海底火山の爆発で出現するのは，今から 1100 万年前～500 万年前[101～103] で，新生代第三紀中新世から鮮新世に至る頃である．

ゾウガメの種分化—そして，いつ海をわたったのか

ゾウガメがいつ海を渡って離れ小島の現在の生息地に移ったかは，いつアルダブラ島が形成され，いつガラパゴス諸島が海底火山の爆発で出現したかによる．それ以前ではありえない．グロビンの分子進化によれば，それは 2100～1500 万年前以降のことである．問題の所在は，2 種のゾウガメが種分化した分岐年代にはゴンドワナがすでに 2 大陸に分離した後であったこと．ゾウガメの分岐年代とこれらの海洋島の出現にかなりのスパンがみられること．ゾウガメの生息地がそれぞれアフリカの東と南アメリカの西に位置する海洋の小島で，陸生の動

物にとって，ゴンドワナ大陸が分裂する前と異なり，お互いの生息地が大西洋で大きく隔てられていることなどである．

南極大陸は，ゴンドワナ大陸が2大陸に分離された後も，2大陸(南アメリカとアフリカ)に今よりもっと接近していたようだ．南アメリカの南端は南極の西側北端近くに位置し続け，6500万年前以降の新生代に入っても海岸線が後退したときには，南極を経由して動物相の交換が可能であったようである[106]．2大陸間の動物相の交流があったと考えられる資料として，1)現在，南アメリカに生息する53種にもおよぶ多様な新世界ザルは，すべて，およそ4000万年前に，アフリカから(南極を経由して)海を渡って南アメリカに辿り着いたサルの子孫であると考えられている[107]．2)アルダブラゾウガメと南アメリカ(パナマ・コロンビア・ヴェネズエラ・ブラジル・パラグアイ・アルゼンチンなど)に生息している同属のアカアシガメ(*Geochelone carbonaria*)のβグロビンに基づく分岐年代は，1700万年前頃と推定されている[7]．3)ミトコンドリアのDNAの研究から，マダガスカルのカメ類(*Geochelone* 属を含む)の分岐年代は，2200～1400万年前と推定され，それらのリクガメはアフリカ西海岸からマダガスカルへ移住したことが推測されている[108]．以上の資料は，陸生の動物がアフリカと南アメリカ間を移動していたこと，また，筆者のデータを加えると *Geochelone* 属の種分化はおおよそ2200万年前～1400万年前におこったことなどを物語っている．

一方，大陸からガラパゴス海域へ向かう海流には，南赤道海流 (south equatorial current) や南アメリカの稜線に平行して南極から赤道へ北上するフンボルト海流 (Humboldt current) があり，ゾウガメ，アシカ，ガラパゴスペンギンなどは海流に乗ってガラパゴスへやって来たとされている[102]．

ゾウガメは海を渡れるか

陸に棲む生物が1000 kmも大陸から離れている海洋島に辿り着くには，何かの拍子(暴風など)に海に流され，流木または流れものかなどに乗って漂流しながら偶然目的地に辿り着くことがもっとも可能性の高い方法であろう．実際に，爬虫類のゾウガメではこの可能性が示唆されている．両生類や陸生の哺乳類と異なり，ゾウガメは比較的少量の淡水で生き延び，その皮膚は海水に耐える性質を持ち合わせている[37]．少なくとも7日間洋上を無傷で漂うことができたとの報告もある[27]．また，ハリケーンに翻弄されて32 km以上も漂流した記録が残っている[109]．この海域の海流と風の速度から見積もって，2週間あればガラパゴスに漂着できると推定される．ゾウガメは餌を探して海浜をさまよう[37]ことも知られているので，何かの折に海岸から洋上へ押しながされ漂流しながら移住するチャンスは大いにあった．

大型化は孤島に生息するカメの特徴；近縁関係とは無関係か

大陸から隔離された海洋島に到達したカメは3グループ知られている[39]．それらに最も特徴的なのは，大型化した背甲であろう．背甲長は1.7 m(ガラパゴスゾウガメ)や1.4 m(アルダブラゾウガメ)に達する個体の記載がある[110]．リクガメの他の種と比較すると，最大13倍になる．化石としては，2.5 mにもなる個体(*Geochelone atlas*)も発見されているのでまさにゾウの大きさである．ゾウガメガが住む環境は大陸や大きな島とどのように異なっているのだろうか．ゾウガメが住む海洋島に共通している特色は，1)大型の捕食動物がいない．2)草食性の競争者がいない．3)固有種は比較的単純な社会を構成している．4)気候の変化，とくに雨期と乾季の変動があるなどである[37]．爬虫類が漂着すると，その島の最も大型の捕食者になり，競争動物のいない島の環境に適応して巨大化したものと考えられる(コモド島のコモドオオトカゲもその一例である)．カメを捕食する競争動物がいない限り，巨大になることが海洋の小島で生息するにもっとも都合のよいことであったと思われる．MacFarlandら[111]によれば，ガラパゴスに人間の移住ともに導入されたイヌとブタ(大型の補食動物)が少なくとも背甲長55～40 cmの若いゾウガメの個体(もちろん，卵も)を餌食にしていることが観察されている．大型化は競争的捕食動物がいない海洋島で爬虫類がとった多様化の一つの類型かも知れない．

今，筆者が考えているゾウガメの種分化のシナリオは次の通りである．

ゴンドワナ大陸がアフリカと南アメリカに分断された後(13,500万年前～6500万年前)，今から2100～1500万年前に，ゴンドワナ大陸の子孫種がいずれか一方の大陸で種分化をおこない(種分化の地点はアフリカか南アメリカのどこかで起きたであろうが，今のところ場所を特定できる資料を持ち合わせていない)，ガラパゴスに放散する種は，ガラパゴス諸島の出現後(1100～500万年前)に，南アメリカから漂流により移住(transmarine immigration)したものと考えられる．また，アルダブラへ放散した種は，アフリカのグループの後裔が，マダガスカルを経由しアルダブラへ漂着したと考えられる．2種類のゾウガメの漂流は，2100～1500万年前から幾度も繰返され，珊瑚礁の形成や海底火山による陸地の形成などにより今の生息地ができてから，条件が偶然に揃い，移住が成功したと推測できる．ゾウガメの2大陸間の移動はどちらが起点になったかは不明であるが，両大陸間の移動は大西洋を直接漂流するのでなく，南極を経由したものと思われる．最近のミトコンドリアDNAの研究[112]によれば，ガラパゴスゾウガメは南アメリカ原産のカメの中でチャコリクガメ(*Geochelone chilensis*)に

最も近縁で，その分岐年代は1200～600万年前と算出された．従って，ミトコンドリアDNAのデータからこの頃 *Geochelone* 属のカメが南アメリカに生息していたことになる．ガラパゴスやアルダブラに移住する前に，それぞれの祖先種が南アメリカとアフリカでさらに種分化をおこない新たな子孫種を形成していたかどうかは不詳である．Caccone らの報告[112]はチャコリクガメがガラパゴスゾウガメの直接の祖先だと示唆しているが，そのような直接の祖先種は消滅していると考えている研究者もいる[113,114]．しかしながら，グロビンの分子内時計と大陸移動説から，ゾウガメの祖先種がどちらかの大陸で種分化をしたことは確実であろう．その時は新生代第三紀中新世と推定している．

おわりに：今後の課題

ダーウィンの「種の起原」で有名なガラパゴスゾウガメは，生物の多様性を考える上で貴重な動物である．大陸から遠く隔離された海洋の島々は，進化の実験室といわれている(日本の小笠原諸島はそのうちの一つと考えられている[115])．筆者は，現存しているアルダブラゾウガメの血液を手に入れるチャンス (いつも希望していた) に巡り会い，グロビンの一次構造と遺伝子構造の成果を研究論文として発表した[7,8]．次に，本丸と目していたガラパゴスゾウガメの血液をいただく交渉のために，東京都上野恩賜動物園を訪れた．さいわい，爬虫類館の責任者は同窓の先輩の一人だった．こうして，調べてみたかったガラパゴスゾウガメの血液 (0.3 m*l*) も手に入れることができた．貴重なガラパゴスゾウガメのグロビンのアミノ酸配列とゲノム遺伝子配列を読み取った情報はさっそくデータバンクに登録した (論文は未発表)．材料を確認するため，上野動物園でガラパゴスゾウガメの飼育記録のコピーをいただいて驚いた．このゾウガメは筆者の大学院時代の指導教官であり終生の恩師と深く尊敬している関口晃一先生が海鷹丸でガラパゴス諸島を調査探検したとき持ち帰った生きた標本で，記録によれば，関口先生が1960年上野動物園に寄贈されたゾウガメであった．

上述したように，アルダブラゾウガメとガラパゴスゾウガメの分岐年は2100～1500万年前と推定できる．この分岐年代と遺伝子構造を比較して分かったことを基盤に，今後研究したい課題は次のようになる．1) 遺伝子のイントロン部分で最も気にかかるヌクレオチド配列が見られた．この10 merのヌクレオチドを手がかりに，*Geochelone* 属 (南アメリカ現生種とアフリカ現生種) のイントロンを探索して，直接の祖先種を探し当てる (先に述べたように，ゾウガメの直接の祖先種はすでに消滅しているとの考え方もある[113,114])．2) α^D 遺伝子の第2イントロン5'-GCは，今のところ，鳥類 (ニワトリ) とカメ類 (ゾウガメ) に特徴的な共通配列である．従って，爬虫類で原始的な形質をもつムカシトカゲ (*Sphenodon*) と特異な分子系統関係を示したコモドオオトカゲ α^D-グロビン (Fig. 5参照) を解析する．3) 現生の曲頸類 (ヨコクビガメの類) はすべて淡水生のカメである．淡水生にもかかわらず，それらの生息地は南アメリカと西アフリカに分布している．それらのグロビンの分岐年代を推定し，ゾウガメと比較する．4) Hb Dをもたない爬虫類のグループ[51,52]が知られているので，そのグループのグロビン遺伝子の構造を検証する (ワニ類は，爬虫類の中で最も特異な系統的位置にいる動物かも知れない)．4) Hb Dの結晶構造を明らかにして，立体構造からHb Dの構造と機能の多様性を探る．

謝辞　アルダブラゾウガメの血液を提供していただいた大阪市天王寺動植物公園の高見一利獣医ならびスタッフの方々，また，ガラパゴスゾウガメの血液を提供していただいた東京都立上野恩賜動物園の飼育課の皆さんのご協力に感謝を申し上げる．本稿に使用した筆者のデータは，日本大学量子科学研究所からの研究費による成果の一部である．

文　献

1) 長谷川政美，岸野洋久．分子系統学．岩波書店，東京，1996.
2) Shishikura F, Snow JW, Gotoh T, et al. Amino acid sequence of the monomer subunit of the extracellular hemoglobin of *Lumbricus terrestris*. *J Biol Chem* 1987; **262**: 3123–3131.
3) Gotoh T, Shishikura F, Snow JW, et al. Two globin strains in the giant extracellular haemoglobins. *Biochem J* 1987; **241**: 441–445.
4) Shishikura F. Amino acid sequence of the monomer subunit of the extracellular hemoglobin of the earthworm, *Pheretima hilgendorfi*. *Zoolog Sci* 1996; **13**: 551–558.
5) Shishikura F, Nakamura M. A comparative study on earthworm hemoglobins: An amino acid sequence comparison of monomer globin chains of two species, *Pontodrilus matsushimensis* and *Pheretima communissima* that belong to the family Megascolecidae. *Zoolog Sci* 1996; **13**: 849–856.
6) Shishikura F, Ochiai T, Yamanaka I. Leech extracellular hemoglobin: Two globin strains that are akin to vertebrate hemoglobin a and a chains. *Zoolog Sci* 1997; **14**: 923–930.
7) Shishikura F, Takami K. The amino acid sequences of the α- and β-globin chains of hemoglobin from the Aldabra giant tortoises, *Geochelone gigantea*. *Zoolog Sci* 2001; **18**: 515–526.
8) Shishikura F. The primary structure of hemoglobin D from the Aldabra giant tortoise, *Geochelone gigantea*. *Zoolog Sci* 2002; **19**: 197–206.
9) Keilin D. Distribution and diversity of haemoglobin. *Acta Biochim Polon* 1956; **3**: 439–457.
10) Kleinschmidt T, Sgouros JG. Hemoglobin sequences. *Biol Chem Hoppe-Seyler* 1987; **368**: 579–615.
11) Hardison R. Hemoglobins from bacteria to man: evolution of different patterns of gene expression. *J Exp Biol* 1998; **201**: 1099–1117.
12) Weber RE, Vinogradov SN. Nonvertebrate hemoglobins: Functions and molecular adaptations. *Physiol Rev* 2001; **81**:

569–628.
13) 宍倉文夫，桑沢清明．血液・血リンパ液 (無脊椎動物)．生物学データ大百科事典 (石原勝敏，金井龍二，河野重行，野村哲郎　編集代表)，東京：朝倉書店，2002; 1019–1093.
14) Dickerson RF, Geis I. Hemoglobin: Structure, Function, Evolution, and Pathology. RedWood City, CA: The Benamin/Cummings Publishing Co, 1982.
15) Bunn HF, Forget BG. Hemoglobin Function. In: Bunn HF, Forget BG, editors. Hemoglobin: Molecular, genetic and clinical aspects. Philadelphia: Saunders WB Co, 1986: 37–60.
16) de Prisco G, Condo SG, Tamburrini M, et al. Oxygen transport in extreme environments. *Trends Biochem Sci* 1991; **16**: 471–474.
17) Jessen TH, Weber RE, Fermi G, et al. Adaptation of bird hemoglobins to high altitudes: Demonstration of molecular mechanism by protein engineering. *Proc Natl Acad Sci USA* 1991; **88**: 6519–6522.
18) Maina JM. What it takes to fly: The structural and functional respiratory refinements in birds and bats. *J Exp Biol* 2000; **203**: 3045–3064.
19) Efstratiadis A, Posakony JW, Maniatis T, et al. The structure and evolution of the human β-globin family. *Cell* 1980; **21**: 653–668.
20) Lauer J, Shen CKJ, Maniatis T. The chromosomal arrangement of human a-like globin genes: sequence homology and α-globin gene deletions. *Cell* 1980; **20**: 119–130.
21) Czelusniak J, Goodman M, Hewett-Emmett D, et al. Phylogenetic origins and adaptive evolution of avian and mammalian haemoglobin genes. *Nature* 1982: **298**: 297–300.
22) Petruzzelli R, Aureli G, Lania A, et al. Diving behaviour and haemoglobin function: the primary structure of the α- and β-chains of the sea turtle (Caretta caretta) and its functional implications. *Biochem J* 1996; **316**: 959–965.
23) Rucknagel KP, Braunitzer G. Hemoglobins of reptiles. The primary structure of the major and minor hemoglobin component of adult Western Painted Turtle (*Chrysemys picta bellii*). *Biol Chem Hoppe-Seyler* 1988; **369**: 123–131.
24) Rucknagel KP, Reischl E, Braunitzer G. Hamoglobine der Reptilian. Expression von α^D-Genen bei Schildkroten, *Chrysemys picta bellii* und *Phrynops hilarii* (Testudines). *Hoppe-Seyler's Z Physiol Chem* 1984; **365**: 1163–1171.
25) Bordin S, Meza AN, Saad STO, et al. cDNA-derived amino-acid sequence of a land turtle (*Geocheone carbonaria*) β-chain hemoglobin. *Biochem Mol Biol Int* 1997; **42**: 255–260.
26) Grubb P. The growth, ecology and population structure of giant tortoises on Aldabra. *Phil Trans Roy Soc Lond B* 1971; **260**: 327–372.
27) Van Denburgh J. The gigantic land tortoises on the Galapagos archipelago. *Proc Calif Acad Sci* 1914; **2**: 203–374.
28) 八杉龍一，小関治男，古谷雅樹，他．岩波生物学辞典．第4版，東京：岩波書店，1996.
29) 内田　亨．動物系統分類の基礎．東京：北隆館．1968: 249–290.
30) Benton MJ. Phylogeny of the major tetrapod groups: Morphological data and divergence dates. *J Mol Evol* 1990; **30**: 409–424.
31) 疋田　努．系統と分類 (その 1)．動物系統分類学　9 (下 B1) 脊椎動物 (IIb1) 爬虫類 I (内田　亨，山田真弓監修)．東京：中山書店，1988; 223–229, 230–269.
32) 中村健治，松井正文．爬虫類の進化と分布の概要．動物系統分類学　9 (下 B1) 脊椎動物 (IIb1) 爬虫類 I (内田　亨，山田真弓監修)．中山書店，東京，1988: 208–221.
33) Auffenberg W. Checklist of fossil land tortoises (Testudinidae). *Bull Florida State Mus* 1974; **18**: 121–251.
34) 疋田　努．爬虫類の進化．東京，東京大学出版会，2002.
35) Schweitzer MH, Marshall M, Carron K, et al. Heme compounds in dinosaur trabecular bone. *Proc Natl Acad Sci USA* 1997; **94**: 6291–6296.
36) Carroll RL. Origin of reptiles. In: Gans C et al, editor. Biology of Reptilia. London & New York: Academic Press, 1969; **1**: 1–44.
37) Arnold EN. Indian Ocean giant tortoises: their systematics and island adaptations. *Phil Trans R Soc Lond B* 1979; **286**: 127–145.
38) Boulenger GA. On remains of an extinct gigantic tortoise from Madagascar (Testudo grandidieri, Vaillant). *Trans Zool Soc Lond* 1894; **13**: 305–311.
39) Stoddart DR, Peake JF. Historical records of Indian Ocean giant tortoise populations. *Phil Trans R Soc Lond B* 1979; **286**: 147–161.
40) Goin CJ, Goin OB, Zug GR. Introduction to herpetology. San Francisco, Freeman & Company, 1978.
41) Mauro NA, Isaacks RE. Examination of reptilian ery-thorcytes as models of the progenitor of mammalian red blood cells. *Comp Biochem Physiol* 1997; **116A**: 323–327.
42) Hawkey CM, Bennett PM, Gascoyne SC, et al. Erythrocyte size, number and haemoglobin content in vertebrates. *British J Haematol* 1991; **77**: 392–397.
43) Moss BA, Hamilton EA. Chicken definitive erythrocyte haemoglobins. *Biochim Biophys Acta* 1974; **371**: 379–391.
44) Hagopian HK, Ingram VM. Developmental changes of erythropoiesis in cultured chick blastoderms. *J Cell Biol* 1971; **51**: 440–451.
45) Brown JL, Ingram VM. Structural studies on chick embryonic hemoglobins. *J Biol Chem* 1974; **249**: 3960–3972.
46) Duguet M, Chauvet J-P, Acher R. Phylogeny of hemoglobins: The complete amino acid sequence of an α-chain of viper (Vipera aspis) hemoglobin. *FEBS Let* 1974; **47**: 333–337.
47) Abbasi A, Wells RMG, Brittain T, et al. Primary structure of the hemoglobins from Sphenodon (*Spehnodon punctatus*, Tuatara, Rhynchocephalia). *Biol Chem Hoppeseyler* 1988; **369**: 755–764.
48) Naqvi S, Zaidi ZH, von Bahr-Lindström H, et al. Characterization of hemoglobin from the lizard *Uromastix hardwickii*. *FEBS Lett* 1983; **162**: 290–295.
49) Rücknagel KP, Braunitzer G, Wiesner H. Hemoglobins of reptiles. The primary structures of the αI- and βI-chains of common iguana (*Iguana iguana*) hemoglobin. *Biol Chem Hoppe-Seyler* 1988; **369**: 1143–1150.
50) Abbasi A, Braunitzer G. Primary structure of hemoglobin from Monitor Lizard (*Varanus exanthematicus albigularis*-Squamata). *Biol Chem Hoppe-Seyler* 1991; **372**: 473–479.
51) Fushitani K, Higashiyama K, Moriyama EN, et al. The amino acid sequences of two a chains of hemoglobins from Komodo dragon *Varanus komodoensis* and phylogenetic relationships of amniotes. *Mol Biol Evol* 1996; **13**: 1039–1043.
52) Gorr TA, Mable BK, Kleinschmidt T. Phylogenetic analysis of reptilian hemoglobins: Trees, rates, and divergences. *J Mol Evol* 1998; **47**: 471–485.
53) Matsuura MSA, Fushitani K, Riggs AF. The amino acid sequences of the α and β chains of hemoglobin from the snake, *Liophis miliaris*. *J Biol Chem* 1989; **264**: 5515–5521.
54) Cobb JA, Manning D, Kolatkar PR, et al. Deoxygenation-

linked association of a tetrameric component of chicken hemoglobin. *J Biol Chem* 1992; **267**: 1183–1189.
55) Chapman BS, Hood LE, Tobin AJ. Minor early embryonic chick hemoglobin M. Amino acid sequences of the ε- and α^{D} chains. *J Biol Chem* 1982; **257**: 651–658.
56) Bauer H, Braunitzer G, Oberthür W, et al. Das Hämoglobin des adulten Andeskondors (*Vultur gryphus, Cathartiformes*). *Biol Chem Hoppe-Seyler* 1985; **366**: 1141–1148.
57) Tamburrini M, Condo SG, di Prisco G, et al. Adaptation to extreme environments: Structure-function relationships in emperor penguin haemoglobin. *J Mol Biol* 1994; **237**: 615–621.
58) Lenfant C, Johansen K, Petersen JA, et al. Respiration in the fresh water turtle, Chelys fimbriata. *Resp Physiol* 1970; **8**: 261–275.
59) Weber RE, Kleinschmidt T, Abbasi A, et al. Allosteric transition in hemoglobin ($\alpha^{A}{}_2\beta'_2$) from the Rhynchocephalian reptile relict *Sphenodon punctatus*. *Hemoglobin* 1989; **13**: 625–636.
60) Brittain T. An investigation of the functioning of the two major haemoglobins of the Sphenodon using fast reaction kinetic methods. *Biochem J* 1988; **251**: 771–776.
61) Bonilla GO, Focesi A Jr, Bonaventura C, et al. Functional properties of the hemoglobin from the South American snake *Mastigodryas bifossatus*. *Comp Biochem Physiol* 1994; **109A**: 1085–1095.
62) Torsoni MA, Ogo SH. Oxygenation properties of hemoglobin from the turtle *Geochelone carbonaria*. *Brazilian J Med Biol Res* 1995; **28**: 1129–1131.
63) Friedman M, Krull LH, Cavins JF. The chromatographic determination of cystine and cysteine residues in proteins as S-β-(4-pyridylethyl) cysteine. *J Biol Chem* 1970; **245**: 3568–3871.
64) Riggs A. Preparation of blood hemogloibns of vertebratge. In: Antonini E, Rossi-Bernardi L, Chiancone E, editors. New York: Academic Press, 1981; **76**: 5–29.
65) Bonaventura J, Bonaventura C. Preparation of high molecular weight invertebrate hemoglobins. In: Antonini E, Rossi-Bernardi L, Chiancone E, editors. New York: Academic Press, 1981; **76**: 43–54.
66) Goodman M, Moore GW, Matsuda G. Darwinian evolution in the genealogy haemoglobin. *Nature* 1975; **253**: 603–608.
67) Thompson JD, Higgins DG, Gibson TJ. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequences weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–4680.
68) Felsenstein J. PHYLIP: Phylogeny Inference Package, Version 3.5c. Department of Genetics SK-50, University of Washington, Seattle, WA 98195, USA, 1993.
69) Leclercq F, Schnek AG, Braunizer G, et al. Direct reciprocal allosteric interaction of oxygen and hydrogen carbonate sequence of the haemoglobins of the caiman (*Caiman crocodylus*), the Nile crocodile (*Crocodylus niloticus*) and the Mississippi crocodile (*Alligator mississippiensis*). *Hoppe-Seyler's Z Physiol Chem* 1981; **362**: 1151–1158.
70) Leclercq F, Stangl A, Schrank B, et al. Caiman crocodylus hemoglobin. Complete primary structures of α and β chains; phylogenic and regulatory aspects. *Int J Peptide Protein Res* 1982; **20**: 337–350.
71) Dodgson JB, McCune KC, Rusling DJ, et al. Adult chicken a-globin genes, α^{A} and α^{D}: No anemic shock α-globin exists in domestic chicken. *Proc Natl Acad Sci USA* 1981; **78**: 5998–6002.
72) Oberthür W, Godovac-Zimmermann J, Braunitzer G. The expression of a α^{D}-chains in the hemoglobin of adult ostrich (*Struthio camelus*) and American rhea (*Rhea americana*). The different evolution of adult bird α^{A}-, α^{D}- and β-chains. *Biol Chem. Hoppe-Seyler* 1986; **367**: 507–514.
73) Braunitzer G, Gehring-Muller R, Hilschmann H, et al. The constitution of normal adult human haemoglobin. *Hoppe-Seyler's Z Physiol Chem* 1961; **325**: 283–286.
74) Islam A, Persson B, Zaidi ZH, et al. Sea snake (*Microcephalophis gracilis*) hemoglobin: Primary structure and relationships to other forms. *J Prot Chem* 1990; **9**: 533–541.
75) Naqvi S, Abbasi A, Zaidi ZH. Primary structure of hemoglobin from cobra *Naja naja naja*. *J Prot Chem* 1994; **13**: 669–679.
76) Takei H, Ota Y, Wu KC, et al. Amino acid sequence of the α chain of chicken AI hemoglobin. *J Biochem* 1975; **77**: 1345–1347.
77) Chapman BS, Tobin AJ, Hood LE. Complete amino acid sequences of the major early embryonic a-like globins of the chicken. *J Biol Chem* 1980; **255**: 9051–9059.
78) Aschauer H, Sanguansermsri T, Braunitzer G. Embryonale Hämoglobine des Menschen: Die Primärstruktur der ζ-Ketten. *Hoppe-Seyler's Z. Physiol Chem* 1981; **362**: 1159–1162.
79) Hsu SL, Marks J, Shaw JP, et al. Structure and expression of the human θ 1 globin gene. *Nature* 1988; **331**: 94–96.
80) Chapman BS, Hood LE, Tobin AJ. Amino acid sequences of the ε and α^{E} globins of Hb E, a minor early embryonic hemoglobin of the chicken. *J Biol Chem* 1982; **257**: 643–650.
81) Chapman BS, Tobin AJ, Hood LE. Complete amino acid sequence of the major early embryonic β-like globin in chickens. *J Biol Chem* 1981; **256**: 5524-5531.
82) Richards RI, Shine J, Ullrich A, et al. Molecular cloning and sequence analysis of adult chicken β globin cDNA. *Nucleic Acids Res* 1979; **7**: 1137–1146.
83) Braunizer G, Schrank B, Stangl A, et al. Hemogloibns, XXIII. Note on the sequence of the delta-chains of human hemoglobin. *Hoppe-Seyler's Z Physiol Chem* 1978; **395**: 777–783.
84) Lawn RM, Efstratiadis A, O'Connell C, et al. The nucleotide sequence of the human β-globin gene. *Cell* 1980; **21**: 647–651.
85) Baralle FE, Shoulders CC, Proudfoot NJ. The primary structure of the human ε-globin gene. *Cell* 1980; **21**: 621–626.
86) Slightom JL, Blechl AE, Smithies O. Human fetal. $^{G}\gamma$- and $^{A}\gamma$-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* 1980; **21**: 627–638.
87) Dodgson JB, Engel JD. The nucleotide sequence of the adult chicken α-globin genes. *J Biol Chem* 1983; **258**: 4623–4629.
88) Engel JD, Rusling DJ, McCune KC, et al. Unusual structure of the chicken embryonic α-globin gene, π'. *Proc Natl Acad Sci USA* 1983; **80**: 1392–1396.
89) Proudfoot NJ, Gil A, Maniatis T. The structure of the human zeta-globin gene and a closely linked, nearly identical pseudogene. *Cell* 1982; **31**: 553–563.
90) Liebhaber SA, Goossens MJ, Kan YW. Cloning and complete nucleotide sequence of human 5'-α-globin gene. *Pro Natl Acad Sci USA* 1980; **77**: 7054–7058.
91) Dolan M, Dodgson JB, Engel JD. Analysis of the adult chicken β-globin gene. Nucleotide sequence of the locus, microheterogeneity at the 5'-end of β-globin mRNA, and aberrant nuclear RNA species. *J Biol Chem* 1983; **258**: 3983–3990.
92) Dodgson JB, Stadt SJ, Choi OR, et al. The nucleotide sequence of the embryonic chicken β-type globin genes. *J Biol Chem* 1983; **258**: 12685–12692.
93) Spritz RA, DeRiel JK, Forget BG, et al. Complete nucleotide

sequence of the human δ-globin gene. *Cell* 1980; **21**: 639–646.
94) Takagi T, Iwaasa H, Yuasa H, et al. Primary structure of *Tetrahymena* hemoglobins. *Biochim Biophys Acta* 1993; **1173**: 75–78.
95) Suzuki T, Ohta S. The hemoglobin gene of the deep-sea clam *Calyptogena soyoae* has a novel intron in A-helix. *Int J Biochem Cell Biol* 2000; **32**: 1205–1211.
96) Vinogradov SN, Walz DA, Pohajdak B. Organization of non-vertebrate globin genes. *Comp Biochem Physiol B* 1992; **103**: 759–773.
97) Jellie AM, Tate WP, Trotman CN. Evolutionary history of introns in a multidomain globin gene. *J Mol Evol* 1996; **42**: 641–647.
98) Lee CC, Caskey CT. デジェネレートプライマーを用いたcDNAクローニング. PCR実験マニュアル(斉藤 隆監訳). 東京：HBJ出版局, 1991; 41–46.
99) Erbil C, Niessing J. The primary structure of the duck α^D-globin gene: an unusual 5' splice junction sequence. *EMBO J* 1983; **2**: 1339–1343.
100) Braithwaite CJR, Taylor JD, Kennedy WJ. The evolution of an atoll: The depositional and erosional history of Aldabra. *Phil Trans R Soc Lond B* 1973; **266**: 307–340.
101) Christie DM, Duncan RA, McBirney AR, et al. Drowned islands downstream from the Galapagos hotspot imply extended speciation times. *Nature* 1992; **365**: 246–248.
102) Jackson MH. Galapagos, a natural history, Canada, Alberta: University of Calgary Press, 1984.
103) White WM, McBirney AR, Duncan RA. Petrology and geochemistry of the Galapagos Islands: Portrait of a pathological mantle plume. *J Geophys Res* 1993; **98**: 19533–19563.
104) Dietz RS, Holden JC. The breakup of pangaea. *Sci Am* 1970; **223**: 34–41.
105) Ravinowitz PD, Coffin Mf, Falvey D. The separation of Madagascar and Africa. *Science* 1983; **220**: 67–69.
106) 長谷川政美, 曹 纓. 大陸の移動と哺乳類の進化. 科学, 1999; **69**: 440–448.
107) Ciochon RL, Chiarelli AB. Evolutionary Biology of the New World Monkeys and Continental Drift. Plenum Press, 1980.
108) Caccone A, Amato G, Gratry OC, et al. A molecular phylogeny of four endangered Madagascar tortoises based on mtDNA sequences. *Mol Phylogenet Evol* 1999: **12**: 1–9.
109) Tounsend CH. Two giant tortoises were swept twenty miles by hurricane. *Bull NY zool Soc* 1936; **39**: 119–120.
110) 中村健治. 体長・寿命. 動物系統分類学 9(下 B1) 脊椎動物(IIb1) 爬虫類 I(内田 亨, 山田真弓 監修). 東京：中山書店, 1988; 204–206.
111) MacFarland GG, Villa J, Toro B. The Galapagos giant tortoises (Geochelone elephantopus). Part 1. Status of the surviving population. *Biol Conserv* 1974; **6**: 118–133.
112) Caccone A, Gibbs JP, Ketmaier V, et al. Origin and evolutionary relationships of giant Galapagos tortoises. *Proc Natl Acad Sci USA* 1999; **96**: 13223–13228.
113) Marlow RW, Patton JL. Biochemical relationships of the Galapagos giant tortoises (*Geochelone elephantopus*). *J Zool Lond* 1981; **195**: 413–422.
114) Patton JL. Genetical processes in the Galapagos. *Biol J Linnean Soc* 1984; **21**: 97–111.
115) 中村方子. ミミズに魅せられて半世紀. 東京：新日本出版社, 2001.

ELSEVIER

Micron 35 (2004) 123–125

micron

www.elsevier.com/locate/micron

Leech hemoglobin: primary structures of four kinds of globins from *Haemadipsa zeylanica* var. *japonica*

Fumio Shishikura*

Biology Department, Nihon University School of Medicine, Tokyo 173-8610, Japan

1. Introduction

The extracellular hemoglobin of the land leech *Haemadipsa zeylanica* var. *japonica* is basically composed of three constituent subunits, a dimer (D1-globin and D2-globin) and two monomers (M1-globin and M2-globin) (Shishikura et al., 1997). They were separated into two distinct groups, A and B, originally proposed by Gotoh et al. (1987) for the multisubunit globins of annelids. Our previous report on the amino (N)-terminal amino acid sequences of the leech globins has shown that D1-globin and M1-globin belong to group A, and D2-globin and M2-globin to group B (Shishikura et al., 1997).

As for the complete primary structure of hemoglobin of annelids, many of studies have been performed on oligochaetes and polychaetes. However, to date, there are few studies on leech hemoglobin. To establish the primary structures of the leech globins is therefore required, and their structures should be useful in clarifying the molecular evolution of hemoglobin in annelids and other invertebrates, as well as in vertebrates. Here the author reports the complete primary structures of the four globins from the land leech *H. zeylanica* by nucleotide and peptide sequencing.

2. Materials and methods

2.1. Preparation of four kinds of globins and globin-chain separation

Globins of *H. zeylanica* were prepared by the method described previously (Shishikura et al., 1997). One dimer subunit and two monomer subunits were separated by gel-filtration on a Superdex 75 column. To separate the D1-globin and D2-globin from the dimer, the disulfide-bonds of the dimer subunit were cleaved by reduction and *S*-pyridylethylation (Friedman et al., 1970), followed by isolation of each globin-chain on a Resource RPC column.

2.2. Protein sequencing

All four globin-molecules modified by *S*-pyridylethylation were digested, separately, with Lysyl endopeptidase (Wako Pure Chemicals Co., Tokyo). Peptide fragments derived from each of the parent molecules were separated using a reversed-phase column, Resource. Sequence analyses of these fragments were performed using a gas phase protein sequencer, PPSQ-10 (Shimadzu Co., Kyoto), equipped with a class LC-10 amino acid analyser. Phenylthiohydantoin (PTH)-derivatives from the sequencer were separated and quantified. The peptides were then aligned with the assistance of sequence similarities of known globin structures of annelids.

2.3. Nucleotide sequencing

Total RNA from *H. zeylanica* was extracted from the whole body by acid guanidium thiocyanate–phenol–chloroform method (Chomczynski and Sacchi, 1987), and mRNAs including the four kinds of globins were isolated with a TaKaRa Oligotex™-dT30 <Super> mRNA Purification Kit (TaKaRa Bio Inc., Shiga). The single-stranded cDNAs were synthesized with a TaKaRa RNA PCR Kit (v. 2.1) using the Oligo dT-Adaptor Primer (M13 primer M4, 17-mer).

For PCR amplification of the 3′ region of the cDNA, the primers used were the adaptor and a redundant oligomer based on N-terminal amino acid sequence of each globin. The second PCR-amplification was conducted with a nested PCR primer (redundant oligomer) and the adaptor. One major fragment was detected on agarose-gel electrophoresis. Then, the fragment was purified and sequenced directly with a BigDye Terminator v1.1 Cycle Sequencing Kit

* Fax: +81-2-3972-0027.
E-mail address: fshishi@med.nihon-u.ac.jp (F. Shishikura).

0968-4328/$ - see front matter
doi:10.1016/j.micron.2003.10.040

```
                  1                                                       57
M1-globin         ------DPHQCGLLEKFKFYKQWTEVFGLG--EQRIEFGLKVFAKLFHDHPDARKLFSNV
M2-globin         DVHVEDHDELCSGGDGNIVVEDWNQLWEGSDSSFRIAFAKEVLLEVVNAHPEAKELFHAV
D1-globin         -------THVCPELSAIKVQTQWREAYADS--SDRVALAQLVYRHLFKMAPESANLFHRV
D2-globin         -------DYHCSIEDIRDIQHDWQFTWGDASLDARIVFGQAVFKKLIELDSSVVEPLKGV
                             *          *          *     *                *
                               73          88
M1-globin         NGENIYSHEFKAHVKRVLSSLDLNAILLSRNDLLEDQLAHLKGQHDSRG-VDWSYVQAFK
M2-globin         NIEDPNSGEFEAHSLRIINTFDLLVNLLQDRHALHEASLHLGHQHAARPGVVAKYFKTFG
D1-globin         NSEEPDSAEFIAFSLRVLNGLDVVITLLDQEKALFAQIEHLHSQHIERH-IPPKYASAFV
D2-globin         HVEDPNSLTFKNHVLRVLNGLDNLINLFDEQGVLVSQLNHLSQQHKERAGVNAAHFKAFA
                    *   *  *     *     *    *      *     **  **  *          *
                                139
M1-globin         QAMLEVLPEYLGVFVCYESWDGCLEHILTGIFKGH--
M2-globin         QELIKAL-AHLIDDFHFIAWKGCFKTLTKEIVGSIPE
D1-globin         EALHHVLPSVIGHCYDEHAWSQCLNSIAKKILS----
D2-globin         RAFIDVL-EVSGNCPNLDAWKGCLAALGHRISLQLKK
                        *           * *       *
```

Fig. 1. Alignment of primary structures of the four globins from *H. zeylanica* hemoglobin. Accession numbers of gene banks ((DDBJ/EMBL/Genebank) are: M1-globin, AB119122; M2-globin, AB119123; D1-globin, AB119124; D2-globin, AB119125. * indicates positions which have a single, fully conserved residue.

(Applied Biosystems, Foster City, CA). The rest of the 3′ end was afterwards confirmed by 3′RACE with the adaptor and a gene-specific primer. All the forward and reverse primers, except for the Oligo dT, were tagged with pUC/M13 sequencing forward 17-mer (Sigma-Aldrich Japan, Tokyo) and reverse 17-mer (Promega, Tokyo), respectively.

For PCR amplification of the 5′ region of the cDNA, gene-specific primers with or without 5′mophosphate were designed in order to extend the sequence in the 5′ end using a TaKaRa 5′-Full RACE Core Set (Maruyama et al., 1995).

2.4. Computer analysis

A multiple alignment program, Clustal X (Jeanmougin and Thompson, 1998), was used in the alignment of the four *H. zeylanica* globin chains, including those of representative oligochaete and polychaete species. A phylogenetic tree was also constructed by neighbor-joining method stored in the program.

3. Results and discussion

Two sequencing methods, protein and nucleotide sequencing, provided sufficient information to establish the complete primary structures of the four *H. zeylanica* globins. Fig. 1 shows the complete primary structures of these four globin chains, whose sequences are reinforced by the different method (nucleotide sequences are not shown here). The mature globin-molecules are composed of 146 amino acid residues for M1-globin, 156 for M2-globin, 143 for D1-globin, and 149 for D2-globin. There are 22 invariant amino acids in the alignment. It is noticeable that the amino acid replacements occur at positions 57 (phenylalanine to leucine in D2-globin), 73 (histidine to phenylalanine in D1-globin), 88 (leucine to phenylalanine in D2-globin), and 139 (alanine to serine in M1-globin), as shown in Fig. 1. When compared sequence similarities among the four globins, 26–33% identities are found.

Fig. 2 shows a phylogenetic tree based on globin primary structures of Hirudinea, Oligochaeta, and Polychaeta. Three clusters can be seen: strain A, strain B, and others including M2-globin (B type globin) and *Lumbricus* III globin (B type globin). The present tree strongly indicates that M2-globin is a type of ancestral globins in annelids.

In conclusion, this study has provided the primary structures of the four globins from the land leech *H. zeylanica*, thereby allowing the construction of a comparative molecular phylogenetic tree of the globins of

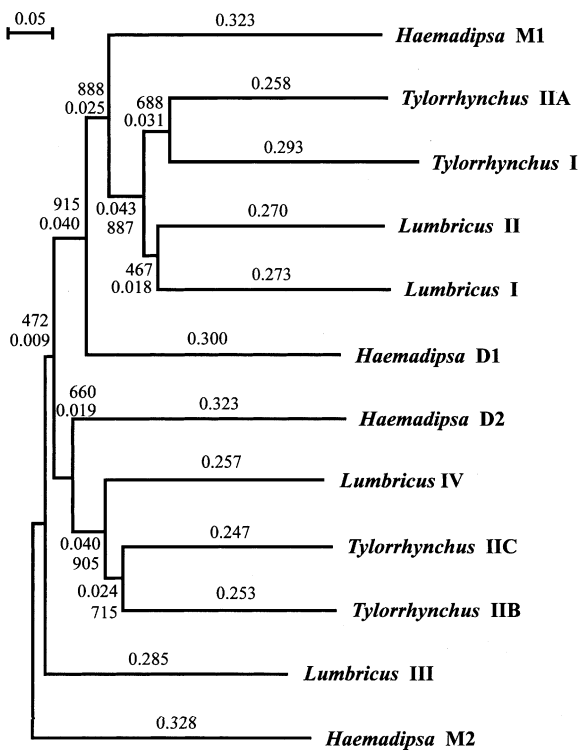


Fig. 2. Phylogenetic tree, based on globin-primary structures of representative species of Annelida, constructed from 1000 bootstrap replications by the neighbor-joining method. Primary structures of *Tylorrhynchus* I, IIA, IIB, and IIC were quoted from Suzuki and Gotoh (1986). Primary structures of *Lumbricus* I and the other three globins of *Lumbricus* (II, III, and IV) were taken from Shishikura et al. (1987) and Fushitani et al. (1988), respectively.

representative species belonging to three orders of Annnelida (Hirudinea, Oligochaeta, and Polychaeta). From the structure of the land leech globins, Hirudinea have ancestral characteristics among annelids.

Acknowledgements

The author thanks Dr S. Tani and Mr Y. Shishikura for collecting and providing *H. zeylanica*. This work was supported in part by a Grant from the Ministry of Education, Culture, Sports, Science and Technology to promote multidisciplinary research projects.

References

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. Anal. Biochem. 162, 156–159.

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.

Fushitani, K., Matsuura, M.S.A., Riggs, A.F., 1988. The amino acid sequences of chains *a*, *b* and *c* that form the trimer subunit of the extracellular hemoglobin from *Lumbricus terrestris*. J. Biol. Chem. 263, 6502–6517.

Gotoh, T., Shishikura, F., Snow, J.W., Ereifej, K.I., Vinogradov, S.N., Walz, D.A., 1987. Two globin strains in the giant annelid extracellular haemoglobins. Biochem. J. 241, 441–445.

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

Maruyama, I.N., Rakow, T.L., Maruyama, H.I., 1995. cRACE: a simple method for identification of the 5′ end of mRNAs. Nucleic Acids Res. 23, 3796–3797.

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., Ochiai, T., Yamanaka, I., 1997. Leech extracellular hemoglobin: two globin strains that are akin to vertebrate hemoglobin α and β chains. Zool. Sci. 14, 923–930.

Suzuki, T., Gotoh, T., 1986. The complete amino acid sequence of giant multisubunit hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. J. Biol. Chem. 261, 9257–9267.

ELSEVIER

Comparative Biochemistry and Physiology Part B 137 (2004) 443–454

CBP

www.elsevier.com/locate/cbpb

The complete amino acid sequences of four globins from the land leech *Haemadipsa zeylanica* var. *japonica*

Fumio Shishikura*

Biology Department, Nihon University School of Medicine, 30-1, Oyaguchi-kami-machi, Itabash, Tokyo 173-8610, Japan

Received 17 October 2003; received in revised form 15 January 2004; accepted 17 January 2004

Abstract

The amino acid sequences of four globins from the land leech, *Haemadipsa zeylanica* var. *japonica*, were determined using nucleotide sequencing and protein sequencing. The mature globin-molecules were composed of 146 amino acid residues for M-1 globin, 156 for M-2 globin, 143 for D-1 globin, and 149 for D-2 globin. Alignment of the four kinds of globins by Clustal X revealed 22 invariant amino acids. The four globins were 26–33% identical. A striking feature of amino acid alteration was: the replacement of the E7 distal-His of D-1 globin by phenylalanine because histidine is conserved among the rest of the globins of *H. zeylanica*, those of other representative species (*Lumbricus* and *Tylorrhynchus*) of Annelida and most other hemoglobins. A phylogenetic tree constructed of 18 globin structures including two species of leeches, *H. zeylanica* (a land leech) and *Macrobdella decora* (a freshwater leech), *T. heterochaetus* (a representative species of polychaetes), *L. terrestris* (a representative species of oligochaetes), and human α and β globins strongly indicated that the leech globins first separated from globin lineage of annelids.

Keywords: Annelida; c-DNA; Evolution; Globin; *Haemadipsa zeylanica*; Nucleotide sequence; Primary structure; RACE

1. Introduction

Although leeches occur in habitats ranging from aquatic (both freshwater and marine) to terrestrial ecosystems and are found on all continents, the systematics of leeches is poorly understood (Mann, 1962; Apakupakul et al., 1999). The jawed Japanese land leech (*Haemadipsa zeylanica* var. *japonica*) is known to have a sanguivorous habit and belongs to the Hirudiniformes, which includes the medicinal leech family Hirudinidae and the terrestrial Haemadipsidae. Recently, the salivary components of sanguivorous leeches have been investigated for pharmaceutical and clinical uses in the prevention of blood clot formation (Walsmann and Markwardt, 1985; Lent, 1986).

The extracellular hemoglobin of the land leech, *H. zeylanica* var. *japonica*, is basically composed of three constituent subunits, a dimer (D-1 and D-2 globins) and two monomers (M-1 and M-2 globins) (Shishikura et al., 1997). The multiple subunit globins of the annelids were separated into two distinct groups, strain A and strain B (Gotoh et al., 1987). Our previous report on the first 30 amino (*N*)-terminal amino acid sequences of the leech globins showed that D-1 globin and M-1 globin belong to strain A, and D-2 and M-2 globins belong to strain B (Shishikura et al., 1997).

Many studies of the primary structure of hemoglobin of the annelids have been performed in oligochaetes and polychaetes; however, to date,

*Corresponding author. Tel.: +81-3-3972-8111x2291; fax: +81-3-3972-0027.

E-mail address: fshishi@med.nihon-u.ac.jp (F. Shishikura).

1096-4959/04/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.cbpc.2004.01.003

few studies on leech hemoglobin have been reported (Kapp et al., 1990; Shishikura et al., 1997). Determination of the primary structures of leech globins, therefore, is required, and their structures may be useful tools for clarifying the molecular evolution of the globin super-families of annelids in particular as well as invertebrates and vertebrates in general.

Here I report the complete primary structures of the four kinds of globins from the land leech, *H. zeylanica*, by nucleotide sequencing and peptide sequencing. A phylogenetic tree based on globin structures strongly indicates that the globins of Hirudinea first diverged from the lineage of globins of Annelida.

2. Material and methods

2.1. Preparation of four globins and globin chain separation

Globins of *H. zeylanica* were prepared by the method described previously (Shishikura et al., 1997). The one dimer subunit and two monomer subunits were separated by gel-filtration on a Superdex 75 column. To separate D-1 globin and D-2 globin, the disulfide-bonds of the dimer subunit were irreversibly cleaved by reduction and *S*-pyridylethylation (Friedman et al., 1970), followed by isolation of each globin chain on a Resource RPC column (Amersham Biosciences, Tokyo). Globin chains were ascertained in accordance with previously determined *N*-terminal amino acid sequences of the four kinds of globins (Shishikura et al., 1997).

2.2. Protein sequencing

All four kinds of globin molecules modified by reduction and *S*-pyridylethylation were separately digested with Lysyl endopeptidase (Wako Pure Chemicals, Tokyo) 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). Peptide fragments derived from each of the parent molecules were separated using a reversed-phase column (Resource RPC) in a 0.1% trifluoroacetic acid (TFA) buffered gradient 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). Sequence analyses of these fragments were performed using a gas phase protein sequencer, PPSQ-10 (Shimadzu, Kyoto, Japan), equipped with a class LC-10 amino acid analyser. Phenylthiohydantoin (PTH)-derivatives from the sequencer were separated and quantified. The peptides then were aligned with the assistance of sequence similarities toward the known globin structures of annelids.

2.3. Nucleotide sequencing

Total RNA from *H. zeylanica* was extracted from three adults (approx. 500 mg in total mass; two to three individuals) by the acid guanidium thiocyanate–phenol–chloroform method (Chomczynski and Sacchi, 1987), and mRNAs (approx. 450 ng in total) including the four kinds of globin mRNAs were isolated with a Takara Oligotex™-dT30 〈Super〉 mRNA Purification Kit (Takara Bio, Shiga, Japan). The single-stranded cDNAs were synthesized with a Takara RNA PCR Kit (v. 2.1) using the Oligo dT-Adaptor Primer (M13 primer M4, 17-mer: 5′-GTTTCCCAGTCACGACT_{15}-3′), according to the manufacturer's instructions.

For PCR amplification of the 3′ region of the cDNA, the primers used were the adaptor and a redundant oligomer based on the *N*-terminal amino acid sequence of each globin (see Appendix 2-A). The second PCR-amplification was conducted with a nested PCR primer (a redundant oligomer) and the adaptor (Appendix 2-B). One major fragment was detected on agarose–gel electrophoresis in each PCR. Then, the fragment was purified by a GenElute™ Agarose Spin Column (Sigma-Aldrich, St. Louis, MO, USA), and sequenced directly with a BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA). 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 2-C).

For PCR amplification of the 5′ region of cDNAs, gene-specific primers with or without 5′ monophosphate as listed in Appendix 2-D were designed in order to extend the sequences in the 5′ ends using a Takara 5′-Full RACE Core Set (Maruyama et al., 1995), according to the manufacturer's instructions. All forward and reverse primers, except for the oligo dT, listed in Appendix 2, were tagged with pUC/M13 sequencing primers (forward 17-mer: 5′-GTAAAACGACGGCCAGT-

3′, Sigma-Aldrich Japan, Tokyo, and reverse 17-mer: 5′-CAGGAAACAGCTATGAC-3′, Promega, Tokyo).

2.4. Computer analysis

A multiple alignment program, Clustal X (Jeanmougin and Thompson, 1998) as well as Clustal W (Thompson et al., 1994), was used to align the four kinds of leech globins, with those of representative species from oligochaetes and polychaetes. Pairwise distances among 18 globin sequences were calculated using the computer program Protdist under the Dayhoff PAM matrix option of the PHYLIP package (Felsenstein, 1993). A phylogenetic tree was constructed by the neighbor-joining method (Saitou and Nei, 1987), and the tree was drawn by the NJ-prot program.

3. Results and discussion

Fig. 1 shows four chromatograms of separation of peptide fragments generated from the parent molecules, M-1 globin, M-2 globin, D-1 globin, and D-2 globin. Fragments were selected and sequenced completely except for k-4 of D-1 globin and k-6 of D-2 globin. As summarized in Appendix 1, the k-peptides from the four kinds of globins were aligned with the assistance of sequence similarities towards known sequences of globins of annelids stored in the SwissProt data bank (http://us.expacy.org/cgi-bin/).

Using primers listed in Appendix 2, cDNA-fragments amplified by PCR, whose major bands were extracted from the agarose gels and sequenced, are shown in Fig. 2. Results of nucleotide sequencing of 12 cDNA-fragments provided enough information to determine complete cDNA sequences. Entire coding regions of the four kinds of *H. zeylanica* globins have been stored in GenBank/DDBJ/EMBL (accession nos. AB119122 for M-1 globin, AB119123 for M-2 globin, AB119124 for D-1 globin, and AB119125 for D-2 globin). The nucleotide sequences of mature proteins, as shown in Appendix 1, are comprised of 146 amino acid residues for M-1 globin, 156 for M-2 globin, 143 for D-1 globin, and 149 for D-2 globin. The amino acid sequences deduced from the nucleotide sequences were identical with those determined by protein analyses described above. Their sequences reinforce each other. In addition, protein analyses demonstrated

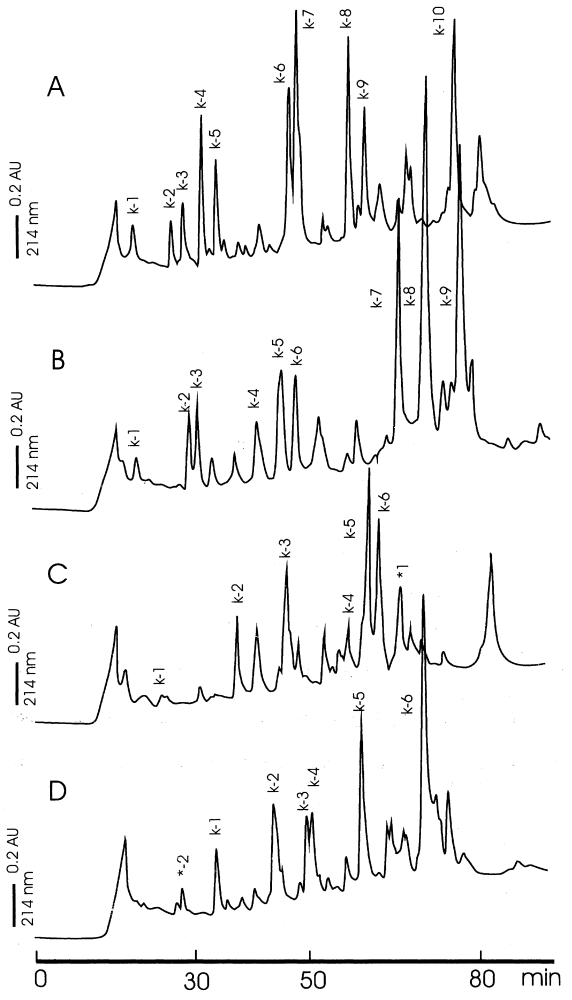


Fig. 1. Separation profiles of peptide fragments by reversed-phase column chromatography in FPLC. (A) M-1 globin, (B) M-2 globin, (C) D-1 globin, (D) D-2 globin. After digestion of the parent molecules (approx. 200 μg) with lysyl endopeptidase, each sample was applied to a Resource RPC column (3 ml packed with Source 15 RPC). All peaks with numbers were sequenced, and their amino acid sequences are shown in Appendix 1.

two kinds of k-peptides (asterisks in Fig. 1), in each of which a micro heterogeneity of amino acids was detected, while by nucleotide analyses no trace of nucleotides indicating micro heterogeneity was found.

Fig. 3 shows an alignment of the four globins, in which there are 22 invariant amino acids (asterisks). Among them, 11 amino acids are common (dots) when compared with *Lumbricus* globins and *Tylorrhynchus* globins (sequences not shown). When they were aligned with human β globin,

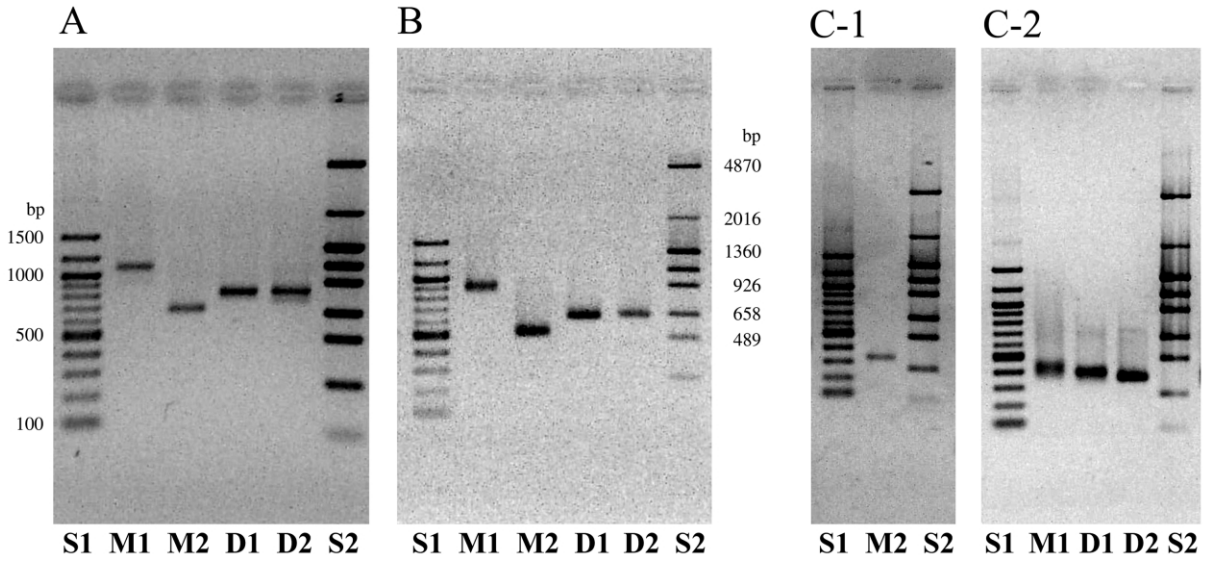


Fig. 2. Agarose gel electrophoreses (1.5% gels) of the PCR products amplified from four kinds of cDNAs using primers as listed in Appendix 2. (A) Nested PCR, (B) 3′-RACE, (C-1) 1st PCR of 5′-RACE, (C-2) 2nd PCR of 5′-RACE. The major fragment in each lane, except for the S1 and S2 lanes, was extracted from the gel and sequenced. PCR conditions: 30 cycles each consisting of 30 s at 94 °C for denaturation, 30 s at 50 °C for annealing, and 1 min at 72 °C for primer extension. S1 (100 bp ladder) and S2 (pHY) are DNA markers.

```
Residue No.      1          .          .                                   57  .
M1               ------DPHQCGLLEKFKFYKQWTEVFGLGEQ--RIEFGLKVFAKLFHDHPDARKLFSNV
M2               DVHVEDHDELCSGGDGNIVVEDWNQLWEGSDSSFRIAFAKEVLLEVVNAHPEAKELFHAV
D1               -------THVCPELSAIKVQTQWREAYADSSD--RVALAQAVYRTLFKMAPESANLFHRV
D2               -------DYHCSIEDIRDIQHDWQFTWGDASLDARIVFGQAVFKKLIELDSSVVEP LKGV
                       *          *          *          *                    *

Helix Position         NA1 A1               B1               C1      CD1       D1
Human β          -------VHLTPE-EKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPD
                                    o                   o

Residue No.                 73  .      .     88      .       .    .
M1               NGENIYSHEFKAHVKRVLSSLDLNAILLSRNDLLEDQLAHLKGQHDSR-GVDWSYVQAFK
M2               NIEDPNSGEFEAHSLRIINTFDLLVNLLQDRHALHEASLHLGHQHAARPGVVAKYFKTFG
D1               NSEEPDSAEFIAFSLRVLNGLDVVITLLDQEKALFAQIEHLHSQHIER-HIPPKYASAFV
D2               HVEDPNSLTFKNHVLRVLNGLDNLINL FDEQGVLVSQLNHLSQQHKERAGVNAAHFKAFA
                   *   *  *      *      *     *       *      **  **  *          *

Helix Position         E1                   EF1           F1        FG1    G1
Human β          AVMG--NPKVKAHGKKVLGAFSDGLAHLDN---LKGTFATLSELHCDKLHVDPENFRLLG
                                                 o       o   o

Residue No.                          139.  .       .
M1               QAMLEVLPEYLGVFVC---YE WDGCLEHILTGIFKGH--
M2               QELIKAL-AHLIDDFH---FIAWKGCFKTLTKEIVGSIPE
D1               EALHHVLPSVIGHCYD---EHAWSQCLNSIAKKILS----
D2               RAFIDVL-EVSGNCPN---LDAWKGCLAALGHRISLQLKK
                       *                *  *        *

Helix Position            Gh1    H1                  HC1
Human β          NVLVCVLAHHFGKEFTPPVQAAYQKVVAGVANALAHKYH-
                        o
```

Fig. 3. Alignment of the amino acid sequences of the four kinds of globins from *H. zeylanica*. The 22 amino acid residues conserved in the four kinds of globins are shown by asterisks. Residue No. is an arbitrary residue 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).

only six amino acids were invariant (shown by open circles): A12-Trp, B16-Val, EF5-Leu, F4-Leu, F8-His and G16-Leu.

Four amino acid substitutions, underlined in Fig. 3, are noticeable because they occurred in only one of the four kinds of globin-chains of *H. zeylanica* compared with *Lumbricus* globins and *Tylorrhynchus* globins: a Leu at position 57 of the D-2 globin, a Phe at position 73 of the D-1 globin, a Phe at position 88 of the D-2 globin, and a Ser at position 139 of the M-1 globin. Among them, the alteration of the Phe at position 73 of the D-1 globin is, in particular, noteworthy because the corresponding His at helix position E7 (E7-His), known as the distal His, is conserved widely in vertebrates and invertebrates. Phillips and Schoenborn (1981) reported that the function of substitute residues (E7-Gln or E7-Leu) are not clear, although they do present steric hindrance to linear ligands, such as carbon monoxide, and favor 'bent' ones, such as O_2. Nagai et al. (1987) produced a mutant human hemoglobin with E7-Gln, Val or Gly using protein engineering and showed that the steric hindrance of ligand binding by the E11 residue and the polarity of the E7 residue in the β subunit were critical for fine-tuning ligand affinity. They also showed that E7-His and E7-Gln are both capable of donating a hydrogen bond to the oxygen molecules in a similar manner and the E7-His to Gln substitution is therefore a structurally and functionally conservative change. A Leu substitution at an E7-His position was also found in a *Glycera* globin (Imamura et al., 1972). In this study I found a Phe substitution at the E7-His position in one of the four kinds of globins of *Haemadipsa* (D-1 globin). Most recently, Suzuki and Vinogradov have reported a Phe substitution at position E7 in a fresh water leech globin (*Macrobdella* IIA, GenBank accession no. AB118638). Suzuki et al. (1989b) found that hemoglobin of the deep sea clam, *Calyptogena soyoae*, (E7-Gln, Suzuki et al., 1989a) was autoxidized 1300 times faster than human hemoglobin under the same conditions. Since the pioneering work of Perutz (1970) who first studied geometries of distal residues in the oxygen binding sites of myoglobin and hemoglobin, E7 substitutions of the conserved His residue remain an interesting

Table 1
Percent identities and pair wise distances between *Haemadipsa* globins, *Macrobdella* globins, *Lumbircus* globins, *Tylorrhynchus* globins, and the human α and β-globin

Globin	*H. zeylanica*				*M. decora*					
	M1	M2	D1	D2	IIA[1]	IIB[2]	B[3]	C[4]		
M1		26.11 (2.089)	32.88 (1.783)	30.46 (1.985)	29.63 (1.782)	30.54 (1.967)	**73.17** (0.218)	27.81 (1.884)		
M2			30.57 (1.567)	30.77 (1.687)	29.94 (1.608)	28.31 (1.852)	26.04 (2.272)	**52.98** (0.627)		
D1				31.33 (1.596)	**75.00** (0.266)	30.54 (1.492)	30.49 (1.690)	30.18 (1.609)		
D2					28.74 (1.636)	**82.88** (0.219)	27.81 (1.879)	32.14 (1.535)		
Globin	*L. terrestris*				*T. heterochaetus*				Homo sapiens	
	I (*d*)[5]	II(*b*)[6]	III(*c*)[6]	IV(*a*)[6]	I[7]	IIA[7]	IIB[7]	IIC[7]	α[8]	β[8]
M1	36.99 (1.420)	35.62 (1.514)	31.17 (1.975)	29.61 (1.850)	35.37 (1.450)	31.51 (1.720)	29.33 (1.885)	27.15 (2.205)	17.22 (3.245)	16.13 (3.335)
M2	24.84 (2.003)	22.29 (2.178)	36.94 (1.233)	32.69 (1.477)	26.92 (1.723)	23.57 (2.456)	33.33 (1.331)	32.05 (1.738)	15.72 (2.980)	16.15 (3.096)
D1	34.03 (1.520)	39.31 (1.186)	34.42 (1.426)	32.90 (1.380)	28.47 (1.690)	34.93 (1.426)	32.90 (1.467)	31.33 (1.544)	16.45 (3.513)	17.45 (3.353)
D2	26.00 (1.910)	26.00 (1.926)	35.95 (1.472)	38.16 (1.303)	24.50 (1.905)	27.81 (1.814)	32.89 (1.335)	37.58 (1.240)	15.79 (3.651)	14.84 (3.477)

Values in parentheses are genetic distances calculated by the program Protdist uder the Dayhoff PAM matrix option (Felsenstein, 1993). (1) DDBJ accession no.: AB118638, (2) DDBJ accession no.: AB 118639, (3) DDBJ accession no.: AB118640, (4) DDBJ accession no.: AB118641, (5) Shishikura et al., 1987, (6) Fushitani et al., 1988, (7) Suzuki and Gotoh, 1986, (8) Braunitzer et al., 1961.

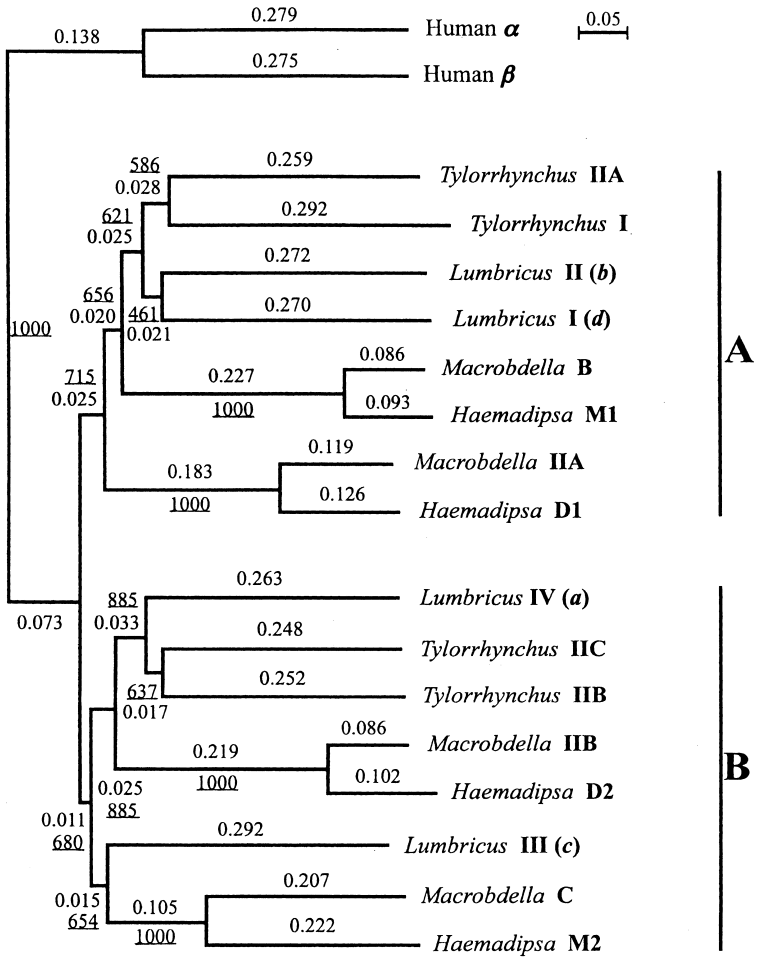


Fig. 4. Phylogenetic tree based on primary structures of 16 globins of Annelida, constructed from 1000 bootstrap replications by the neighbor-joining method (Saitou and Nei, 1987). The scale shown in the upper segment of the tree represents the evolutionary distances, given as the average number of substitutions per site. The scores (with underlines) shown at each node represent bootstrap values.

objective. We are currently making crystals from the leech hemoglobin as well as from *C. soyoae* hemoglobin.

It is well known that homologous subunits sharing an orthologous relationship in macromolecules comprised of multiple subunits, namely, hemoglobin and hemocyanin, are useful tools for analyzing the genetic relationships of invertebrates and vertebrates. I compared the primary structures of two sets of four kinds of globins from *Haemadipsa* and *Macrobdella* (GenBank accession nos. AB118638, AB118639, AB118640, AB118641). Table 1 lists the percentage identities and genetic distances between the complete amino acid sequences of *Haemadipsa*, and the scores were compared with those of *Macrobdella*, *Lumbricus*, and *Tylorrhynchus* as well as human α and β globins. Low similarities (22–38%) were found when leech globins were compared with *Lumbricus* globins and *Tylorrhynchus* globins, as well as human α globin (16–17%) and β globin (15–18%). However, very high scores (53–83%) were found between two globins of the leeches such as the M-1 globin vs. *Macrobdella* B globin (73.2%), M-2 globin vs. *Macrobdella* C globin (53.0%), D-1 globin vs. *Macrobdella* IIA globin (75.0%), and D-2 globin vs. *Macrobdella* IIB globin (82.9%), suggesting each of these two globins to be in orthologous relationships. Traditionally, *H. zeylanica* belongs to Haemadipsidae and *M. decora* belongs to Hirudinidae. The suborder of the two species is Hirudiniformes. Siddall and Burreson

(1998) have reported a molecular phylogeny of leeches, including the Haemadipsidae and Hirudinidae, based on the mitochondrial cytochrome *c* oxidase subunit I, and confirmed a close relationship to each other in the traditional phylogeny of leeches (Mann, 1962). This is also supported by this study.

Gotoh and his collaborators (Gotoh et al., 1987) proposed two strains of globins, A and B, for the classification of multiple globins of annelids. The two-strain hypothesis has been revised: strain A was subdivided into A1 and A2 and strain B into B1 and B2 (Gotoh et al., 1991; Suzuki et al., 1993). Recently, Negrisolo et al. (2001) proposed a new model for globin evolution of annelids, vestimentiferans, and pogonophorans. They also pointed out that the subdivision into four homologous groups of globins (A1, A2, B1 and B2) could be an oversimplification of the real situation. This proposal is supported by the molecular relationships of the four groups appeared on Fig. 4.

Fig. 4 shows the phylogenetic relationship of annelid globins based on primary structures, including a land leech (*H. zylanica*) and a fresh water leech (*M. decora*), a terrestrial earthworm (*L. terrestris*; a representative species of Oligochaeta), and a marine polychaete (*T. heterochaetus*; a representative species of Polychaeta). In a preceding paper (Shishikura et al., 1997), we classified the four kinds of globins of *H. zeylanica* into two strains, A and B, based on *N*-terminal sequences, and this classification was confirmed by this study. Fig. 4 and Table 1 clearly support four subdivisions in the cases of leech globins. The tree also partly supports Negrisolo's recommendation that to classify some *Lumbricus* globins and *Tylorrhynchus* globins into four subdivisions is an oversimplification. More information on the primary structures is needed to establish the real molecular relationships among leech globins in Hirudinea because the two kinds of leeches reported here seem to have diverged in very recent times (Siddall and Burreson, 1998; Apakupakul et al., 1999).

In conclusion, this study determined the primary structures of the four kinds of globins from the land leech, *H. zeylanica*, allowing the construction of a molecular phylogenetic tree of globins among representative species belonging to the three orders of Annelida (Hirudinea, Oligochaeta and Polychaeta). Judging from the molecular relationships of annelid globins, the leech globins first separated from the lineage of Annelida.

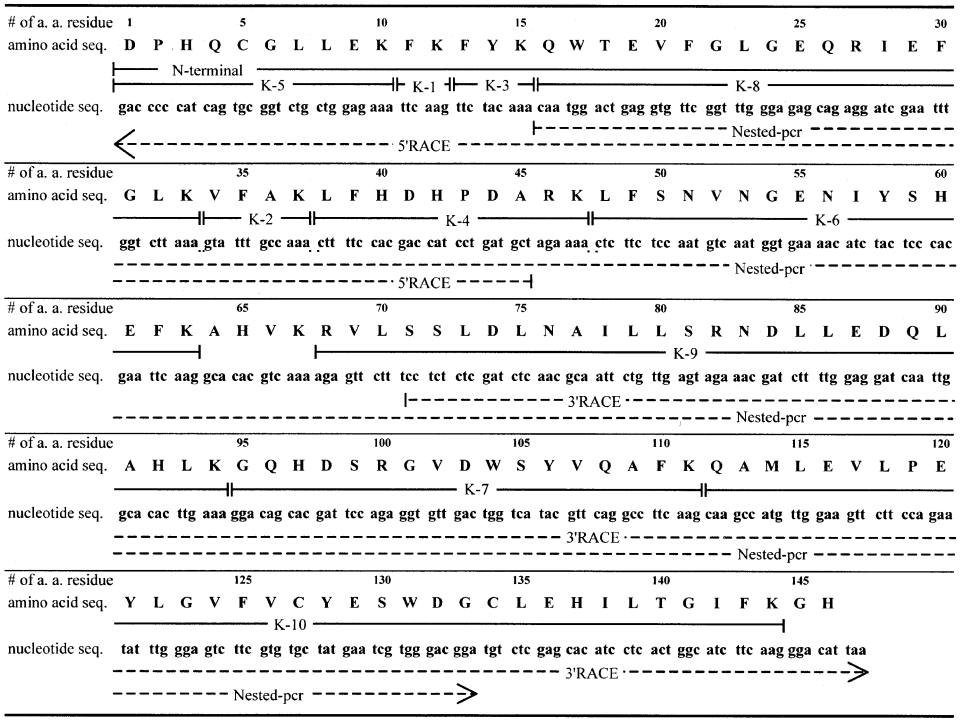
Acknowledgments

The author thanks Dr S. Tani and Mr Y. Shishikura for collecting and providing *H. zeylanica*. This work was supported in part by a Grant from the Ministry of Education, Culture, Sports, Science and Technology to promote multi-disciplinary research projects.

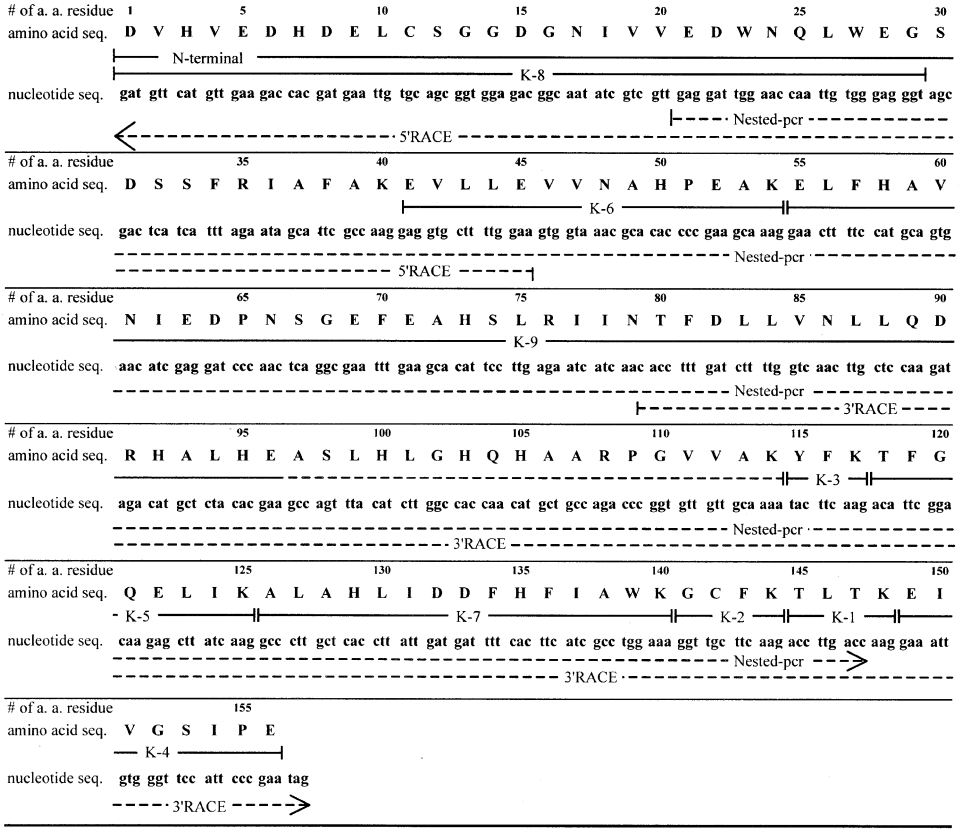
Appendix A:

Appendix B:

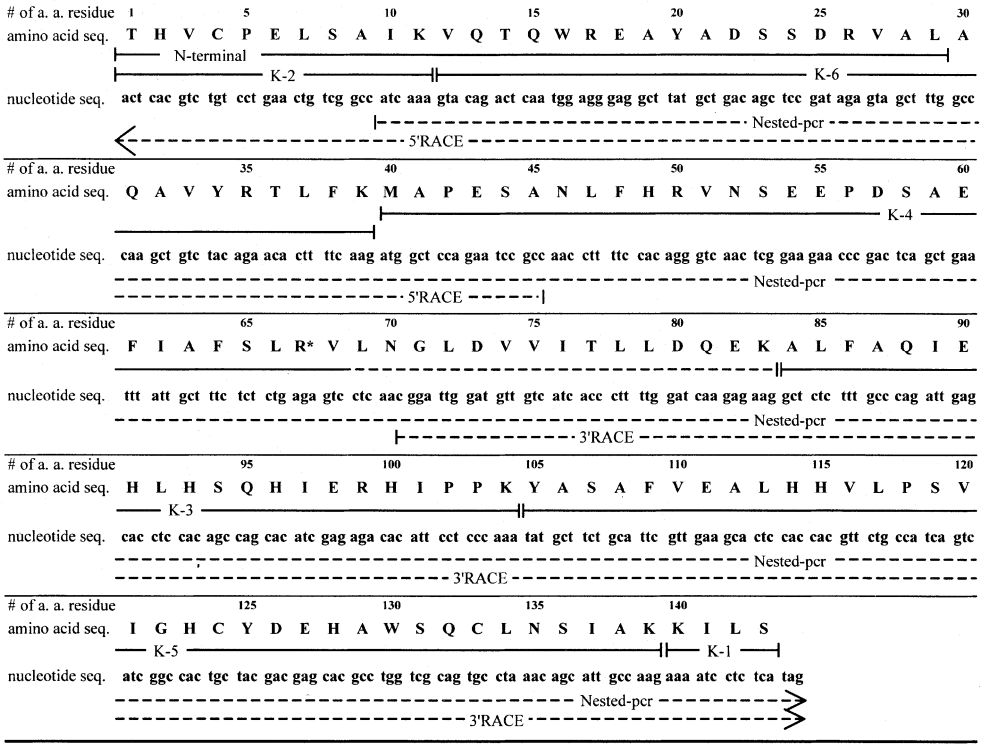
Appendix 1-A. Primary structure and nucleotide sequence of M-1 globin of *H. zeylanica*



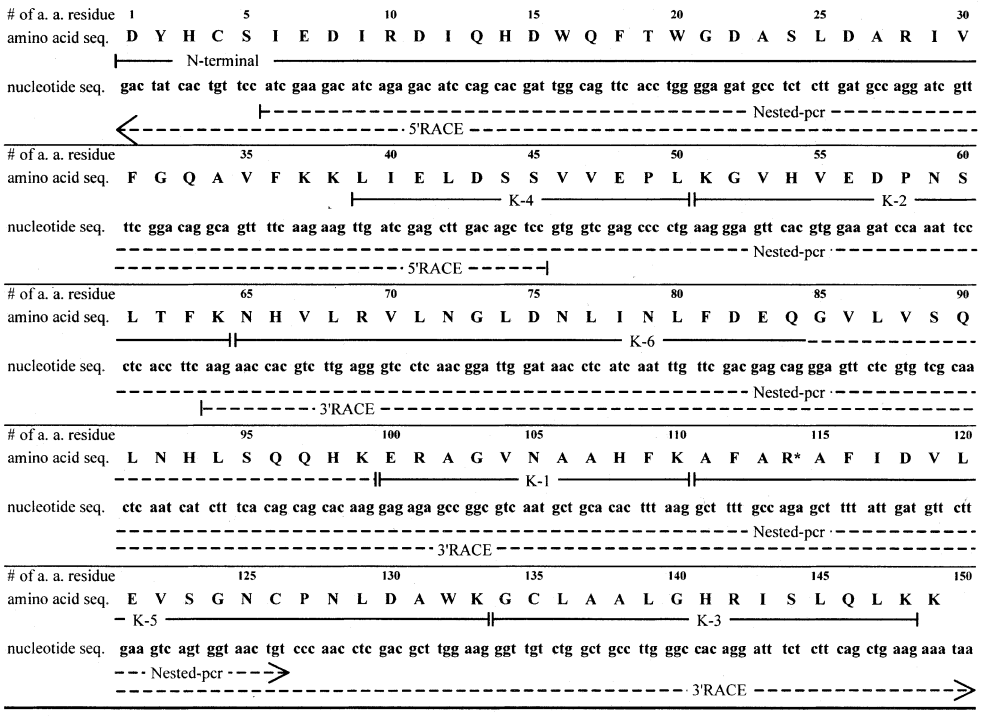
Appendix 1-B. Primary structure and nucleotide sequence of M-2 globin of *H. zeylanica*



Appendix 1-C. Primary structure and nucleotide sequence of D-1 globin of *H. zeylanica*



Appendix 1-D. Primary structure and nucleotide sequence of D-2 globin of *H. zeylanica*



*, microheterogenity: D1: 67 (K), D2:114 (K)

Appendix 2-A. Oligonucleotide primers used in the first PCR of globin cDNAs

HZM1	# of Amino acid residue	1			2			3			4			5			6			7				
N-terminal	Amino acid sequence	**D**			**P**			**H**			**Q**			**C**			**G**			**L**				
	Nucleotide sequence	**g**	**a**	**y**	**c**	**c**	**n**	**c**	**a**	**y**	**c**	**a**	**r**	**t**	**g**	**y**	**g**	**g**	**n**	**y**	**t**			
HZM2	# of Amino acid residue	1			2			3			4			5			6			7				
N-terminal	Amino acid sequence	**D**			**V**			**H**			**V**			**E**			**D**			**H**				
	Nucleotide sequence	**g**	**a**	**y**	**g**	**t**	**n**	**c**	**a**	**y**	**g**	**t**	**n**	**g**	**a**	**r**	**g**	**a**	**y**	**c**	**a**			
HZD1	# of Amino acid residue	1			2			3			4			5			6			7				
N-terminal	Amino acid sequence	**T**			**H**			**V**			**C**			**P**			**E**			**L**				
	Nucleotide sequence	**a**	**c**	**n**	**c**	**a**	**c**	**y**	**t**	**n**	**t**	**g**	**y**	**c**	**c**	**n**	**g**	**a**	**r**	**y**	**t**			
HZD2	# of Amino acid residue	1			2			3			4			5			6			7				
N-terminal	Amino acid sequence	**D**			**Y**			**H**			**C**			**S**			**I**			**E**				
	Nucleotide sequence	**g**	**a**	**y**	**t**	**a**	**y**	**c**	**a**	**c**	**y**	**g**	**y**	**w**	**s**	**n**	**a**	**t**	**h**	**g**	**a**			

Appendix 2-B. Oligonucleotide primers used in nested PCR of globin fragments

HZM1	# of Amino acid residue	16			17			18			19			20			21			22				
	Amino acid sequence	**Q**			**W**			**T**			**E**			**V**			**F**			**G**				
Nested-pcr	Nucleotide sequence	**c**	**a**	**r**	**t**	**g**	**g**	**a**	**c**	**n**	**g**	**a**	**r**	**g**	**t**	**n**	**t**	**t**	**y**	**g**	**g**			
HZM2	# of Amino acid residue	21			22			23			24			25			26			27			28	
	Amino acid sequence	**E**			**D**			**W**			**N**			**Q**			**L**			**W**			**E**	
Nested-pcr	Nucleotide sequence	**g**	**a**	**r**	**g**	**a**	**y**	**t**	**g**	**g**	**a**	**a**	**y**	**c**	**a**	**r**	**y**	**t**	**n**	**t**	**g**	**g**	**g**	
HZD1	# of Amino acid residue	10			11			12			13			14			15			16			17	
	Amino acid sequence	**I**			**K**			**V**			**Q**			**T**			**Q**			**W**			**R**	
Nested-pcr	Nucleotide sequence	**a**	**t**	**h**	**a**	**a**	**r**	**g**	**t**	**n**	**c**	**a**	**r**	**a**	**c**	**n**	**c**	**a**	**r**	**t**	**g**	**g**	**m**	**g**
HZD2	# of Amino acid residue	6			7			8			9			10			11			12				
	Amino acid sequence	**I**			**E**			**D**			**I**			**R**			**D**			I				
Nested-pcr	Nucleotide sequence	**a**	**t**	**h**	**g**	**a**	**r**	**g**	**a**	**y**	**a**	**t**	**h**	**m**	**g**	**n**	**g**	**a**	**y**	**a**	**t**			

Appendix 2-C. Oligonucleotide primers used in 3' RACE of globin fragments

HZM1	# of Amino acid residue	71			72			73			74			75			76			77				
	Amino acid sequence	**S**			**S**			**L**			**D**			**L**			**N**			**A**				
3'RACE	Nucleotide sequence		**c**	**c**	**t**	**c**	**t**	**c**	**t**	**c**	**g**	**a**	**t**	**c**	**t**	**c**	**a**	**a**	**c**	**g**	**c**			
HZM2	# of Amino acid residue	78			79			80			81			82			83			84				
	Amino acid sequence	**N**			**T**			**F**			**D**			**L**			**L**			**V**				
3'RACE	Nucleotide sequence		**c**	**a**	**c**	**c**	**t**	**t**	**t**	**g**	**a**	**t**	**c**	**t**	**t**	**t**	**t**	**t**	**g**	**g**	**t**	**c**		
HZD1	# of Amino acid residue	70			71			72			73			74			75			76			77	
	Amino acid sequence	**N**			**G**			**L**			**D**			**V**			**V**			**I**			**T**	
3'RACE	Nucleotide sequence		**c**	**g**	**g**	**a**	**t**	**t**	**g**	**g**	**a**	**t**	**g**	**t**	**t**	**g**	**t**	**c**	**a**	**t**	**c**	**a**	**c**	**c**
HZD2	# of Amino acid residue	64			65			66			67			68			69			70				
	Amino acid sequence	**K**			**N**			**H**			**V**			**L**			**R**			**V**				
3'RACE	Nucleotide sequence	**a**	**a**	**g**	**a**	**a**	**c**	**c**	**a**	**c**	**g**	**t**	**c**	**t**	**t**	**g**	**a**	**g**	**g**	**g**	**t**	**c**		

Appendix 2-D. Oligonucleotide primers used in 5' cRACE of globin fragments

Fragment	Primer	# of Amino acid residue	Amino acid sequence	Nucleotide sequence
HZM1	P (reverse)	95 94 93 92 91 90	G K L H A	c c t t t c a a g t g t g c c
	1st PCR-F	71 72 73 74 75 76 77	S S L D L N A	c c t c t c t c g a t c t c a a c g c
	1st PCR-R	66 65 64 63 62 61 60 59	V H A K F E H S	a c g t g t g c c t t g a a t t c g t g g g
	2nd PCR-R	45 44 43 42 41 40	A D P H D H	a g c a t c a g g a t g g t c g t g
HZM2	P (reverse)	98 97 96 95 94	S A E H L	t g g c t t c g t g t a g a g
	1st PCR-R	45 44 43 42 41 40 39	E L L V E K A	t t c c a a a a g c a c c t c c t t g
	1st PCR-F	79 80 81 82 83 84 85	N T F D L L V	c a c c t t t g a t c t t t t g g t c
HZD1	P (reverse)	92 91 90 89 88	L H E I Q	g a g g t g c t c a a t c t g
	1st PCR-F	70 71 72 73 74 75 76 77	N G L D V V I T	c g g a t t g g a t g t t g t c a t c a c c
	1st PCR-R	58 57 56 55 54 53 52 51	S D P E E S N V	g a g t c g g g t t c t t c c g a g t t g a c
	2nd PCR-R	45 44 43 42 41 40 39	A S E P A M K	g c g g a t t c t g g a g c c a t c
HZD2	P (reverse)	87 86 85 84 83	L V G Q E	a g a a c t c c c t g c t c
	1st PCR-F	64 65 66 67 68 69 70	K N H V L R V	a a g a a c c a c g t c t t g a g g g t c
	1st PCR-R	56 55 54 53 52 51 50 49	E V H V G G L P	t t c c a c g t g a a c t c c c t t c a g g
	2nd PCR-R	45 44 43 42 41 40 39	S S D L E I L	g g a g c t g t c a a g c t c g a t c

References

Apakupakul, K., Siddall, M.E., Burreson, E.M., 1999. Higher level relationships of leeches (Annelida: Clitellata: Euhirudinea) based on morphology and gene sequences. Mol. Phy. Evol. 12, 350–359.

Braunitzer, G., Gehring-Muller, R., Hilschmann, N., Hilse, K., Hobom, G., Rudloff, V., et al., 1961. Die Konstitution des normalen adulten Humanhamoglobins. Hoppe-Seyler's Z. Physiol. Chem. 325, 283–286.

Chomczynski, P., Sacchi, N., 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. Anal. Biochem. 162, 156–159.

Felsenstein, J., 1993. PHYLIP: Phylogen inference Package, Version 3.5c. Department of Genetics SK-50. University of Washington, Seattle, WA.

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-sepecific oligonucleotide primer. Proc. Natl. Acad. Sci. USA 85, 8998–9002.

Fushitani, K., Matsuura, M.S.A., Riggs, A.F., 1988. The amino acid sequences of chains a, b and c that form the trimer subunit of the extracellular hemoglobin from *Lumbricus terrestris*. J. Biol. Chem. 263, 6502–6517.

Gotoh, T., Shishikura, F., Snow, J.W., Ereifej, K.I., Vinogradov, S.N., Walz, D.A., 1987. Two globin strains in the giant annelid extracellular haemoglobins. Biochem. J. 241, 441–445.

Gotoh, T., Suzuki, T., Takagi, T., 1991. In: Vinogradov, S.N., Kapp, O.H. (Eds.), Nomenclature of the Major Constituent Chains Common to Annelid Extracellular Hemoglobins. Springer-Verlag, Berlin, pp. 279–283.

Imamura, T., Baldwin, T.O., Riggs, A., 1972. The amino acid sequence of the monomeric hemoglobin component from the bloodworm, *Glycera dibranchiata*. J. Biol. Chem. 247, 2785–2797.

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

Kapp, O.H., Qabar, A.N., Bonner, M.C., Stern, M.S., Walz, D.A., Schmuck, M., et al., 1990. Quaternary structure of the giant extracellular hemoglobin of the leech *Macrobdella decora*. J. Mol. Biol. 213, 141–158.

Lent, C., 1986. New medicinal and scientific uses of the leech. Nature 323, 494.

Mann, K.H., 1962. Leeches (Hirudinea). Their Structure, Physiology, Ecology and Embryology. Pergamon Press, New York, pp. 201.

Maruyama, I.N., Rakow, T.L., Maruyama, H.I., 1995. cRACE: a simple method for identification of the 5′ end of mRNAs. Nucleic Acids Res. 23, 3796–3797.

Nagai, K., Luisi, B., Shih, D., Miyazaki, G., Imai, K., Poyart, C., et al., 1987. Distal residues in the oxygen binding site of haemoglobin studied by protein engineering. Nature 329, 858–860.

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

Negrisolo, E., Pallavicini, A., Barbato, R., Dewilde, S., Ghiretti-Magaldi, A., Moens, L., et al., 2001. The evolution of extracellular hemoglobins of annelids, vestimentiferans, and pogonophorans. J. Biol. Chem. 276, 26 391–26 397.

Perutz, M.F., 1970. Stereochemistry of cooperative effects in haemoglobin. Nature 228, 726–739.

Phillips, S.E.V., Schoenborn, B.P., 1981. Neutron diffraction reveals oxygen–histidine hydrogen bond in oxymyoglobin. Nature 292, 81–82.

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

Shishikura, F., Ochiai, T., Yamanaka, I., 1997. Leech extracellular hemoglobin: two globin strains that are akin to vertebrate hemoglobin α and β chains. Zool. Sci. 14, 923–930.

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.

Siddall, M.E., Burreson, E.M., 1998. Phylogeny of leeches (Hirudinea) based on mitochondrial cytochrome *c* oxidase subunit I. Mol. Phylogenet. Evol. 9, 156–162.

Suzuki, T., Gotoh, T., 1986. The complete amino acid sequence of giant multisubunit hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. J. Biol. Chem. 261, 9257–9267.

Suzuki, T., Takagi, T., Ohta, S., 1989a. Primary structure of a dimeric haemoglobin form the deep-sea cold-seep clam *Calyptogena soyoae*. Biochem. J. 260, 177–182.

Suzuki, T., Takagi, T., Ohta, S., 1989b. Amino acid sequence of the dimeric hemoglobin (Hb I) from the deep-sea cold-seep clam *Calyptogena soyoae* and the phylogenetic relationship with other molluscan globins. Biochim. Biophys. Acta 993, 254–259.

Suzuki, T., Takagi, T., Ohta, S., 1993. N-terminal amino acid sequences of 440 kDa hemoglobins of the deep-sea tube worms, Lamellibrachia sp. 1, Lamellibrachia sp. 2 and slender vestimentifera gen. Sp. 1 evolutionary relationship with annelid hemoglobins. Zool. Sci. 10, 141–146.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequences weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Walsmann, P., Markwardt, F., 1985. On the isolation of the thrombin inhibitor Hirudin. Thromb. Res. 40, 563–569.

Protein and Peptide Letters, Vol. 10, No. 4, pp. 422-425, 2003
Bentham Science Publishers Ltd.
0929-8665/03 $ 41.00 + .00

CRYSTALLIZATION AND PRELIMINARY X-RAY DIFFRACTION STUDY OF HEMOGLOBIN D FROM THE ALDABRA GIANT TORTOISE, *Geochelone gigantea*

Takao Kuwada[1], Tomokazu Hasegawa[2], Isamu Satoh[1], Koichi Ishikawa[3] and Fumio Shishikura[1, 4, *]

[1]Institute of Quantum Science, Nihon University, Funabashi 274-8501, Japan
[2]X-ray Research Laboratory, Rigaku Corporation, Akishima, 196-8666, Japan
[3]Department of Pharmacology, Nihon University School of Medicine, Itabashi, Tokyo 173-8610, Japan
[4]Department of Biology, Nihon University School of Medicine, Itabashi, Tokyo 173-8610, Japan
Corresponding author: Fumio Shishikura; Phone: 81-3-3972-8111; Fax. 81-3-3972-0027; E-mail: fshishi@med.nihon-u.ac.jp

ABSTRACT: Hemoglobin D (Hb D) from the Aldabra giant tortoise, *Geochelone gigantea*, was crystallized by the hanging drop vapor diffusion technique with a precipitant solution containing 10% polyethylene glycol 3350 and 50 mM HEPES-Na, pH 7.5. The Hb D crystals of *G. gigantea*, which diffract to at least a 2.0 Å resolution, belong to the monoclinic space group *C2* with unit cell dimensions of a = 112.1 Å, b = 62.4 Å, c = 54.0 Å, and β = 110.3°. One $\alpha\beta$ dimer molecule of Hb D existed in an asymmetric unit, with a calculated value of Vm of 2.77 $Å^3Da^{-1}$.

Keywords: hemoglobin D, crystallization, X-ray diffraction, Reptilia, the Aldabra giant tortoise, *Geochelone gigantea.*

INTRODUCTION

Amniota (reptiles, birds, and mammals) have two or more hemoglobin components that are expressed according to the demands of different physiological conditions [1, 2]. In reptiles, two hemoglobin components have been detected in the erythrocytes [3]: hemoglobin A (Hb A) is common to vertebrates [4, 5], and hemoglobin D (Hb D) was first identified in birds as a minor component of the embryonic and adult definitive erythrocytes [6, 7]. To our knowledge, the presence of Hb D has only been confirmed in many but not all birds, in sphenodon, in after lizards and snakes, and in turtles.

The Aldabra giant tortoise, *Geochelone gigantea*, is one of the two extant giant tortoises [8, 9]. Our previous study on *G. gigantea* hemoglobin [10] revealed that the primary structures and genomic

nucleotide sequences of the constituent globin molecules strongly resembled those of Hb D in birds [11]. These findings provided biochemical evidence of a close relationship between the two Phyla, Aves and Reptilia.

In the course of this study, we encountered a similar phenomenon to that found in birds Hb by Cobb *et al.* [12] and Morrow *et al.* [13], who reported that the solubility of Hb D but not Hb A decreased greatly and that Hb D formed crystalline aggregates upon deoxygenation. In fact, the solubility of *Geochelone* Hb D decreased greatly and crystalline aggregates were formed if there was no coexistence of the cognate Hb A. Riggs [14] also reported that the Hbs of most birds and of some reptiles and amphibians are characterized by a further deoxygenation-dependent self-association of tetramers. These tetramer-tetramer complexes have a greatly lowered oxygen affinity and, in turn, partial oxygenation results in dissociation of tetramers of higher oxygen affinity. Thus, it would be of useful to analyze the tertiary structure of *Geochelone* Hb D for better understanding the deoxygenation-dependent self-association of tetramers, the cooperativity of the oxygen binding mechanism, and, in particular, the tetramer-tertramer contact first suggested for chicken Hb by Huisman *et al.* [15], since reptiles and birds have two different components, Hb A and Hb D, which share common β-globin chains.

We aim to establish the crystal structure of *G. gigantea* Hb D and analyze the association of the dimer-dimer and tetramer-tetramer molecules in comparison with the known crystal structures of the Hb D of chicken [16] as well as the mammalian embryonic hemoglobin [17]. Here, we describe the crystallization and preliminary X-ray diffraction analysis of the *G. gigantea* Hb D. This is the first crystallization report for the reptilian Hb D.

EXPERIMENTAL RESULTS

Preparation of the hemoglobin D

Preparation of native Hb D has been reported previously [10, 18]. The purified Hb D for crystallization was desalted and concentrated in 50 mM Tris-HCl (pH 7.5) with Ultrafree-4 centrifugal ultrafiltration devices (Nihon Millipore Ltd., Tokyo, Japan).

Crystallization

Crystallization of *G. gigantea* Hb D was carried out by the hanging drop vapor diffusion technique, using VDX plates (Hampton, CA, USA) as follows: Hb D solution (2.0 μl of 3.5-4.0 mg/ml) in 50 mM Tris-HCl (pH 7.5) was placed on a siliconized cover glass and mixed with an equal amount of the reservoir solution containing 10% (w/v) polyethylene glycol (PEG, mean molecular weight of 3350 Da, Hampton, CA, USA) as a precipitant in 50 mM HEPES-Na, pH 7.5. The crystallization drops on the siliconized cover glasses were sealed with vacuum grease and the vapor diffusion reaction was performed at 20°C against 0.75 ml of reservoir solution.

The purified *Geochelone* Hb D formed very small and very thin crystalline aggregates called microcrystals during concentration as well as during incubation of the purified Hb D samples at 4°C. These microcrystals were too small for analysis at our facilities and also inhibited the growth of large

crystals by the hanging drop vapor diffusion technique. However, we found that the removal of microcrystals by filtration (pore size: 0.1 μm in diameter) proved to be quite useful for the growth of large crystals of the *Geochelone* Hb D. In these experiments, the crystals were first crushed and used as microseeds. The micro-seeding trials gave small but nicely formed single crystals in a day. Finally, a cycle of macro-seeding with these single crystals yielded crystals, which were large enough for proper analysis in five days. Figure 1 shows a single crystal of *Geochelone* Hb D whose size was approximately 0.2 x 0.3 x 0.2 mm.

X-ray diffraction analysis

X-ray diffraction data were collected from flash-cooled crystals at 100 °K using a Rigaku R-AXIS IV^{++} image plate detector with CuKa radiation from a Rigaku ultraX-18 rotating anode generator (Rigaku Co., Tokyo, Japan). A crystal was immersed in anti-freezing solution containing 20% glycerol for 10-15 s, was picked-up with a loop, and was then rapidly transferred to a cold stream. The conditions of data collection were camera length = 150 mm, exposure time = 1 min, and X-ray source = 50 kV, 100 mA (fine focused). The diffraction data were processed using the software CrystalClear (version 1.3, Rigaku Co., Tokyo, Japan). The crystal system was found to be monoclinic, with a space group of *C2*, and unit cell dimensions of a = 112.1 Å, b = 62.4 Å, c = 54.0 Å, and β = 110.3°. The resulting data set was 99% complete at a 53.65 - 2.02 Å resolution with an overall R_{merge} of 4.0% (Table 1). The value of Vm was 2.77 $Å^3Da^{-1}$, suggesting that there is one αβ dimer in a symmetric unit. The tertiary structural determination using these crystals is currently in progress by molecular replacement technique on the model of chicken Hb D [16].

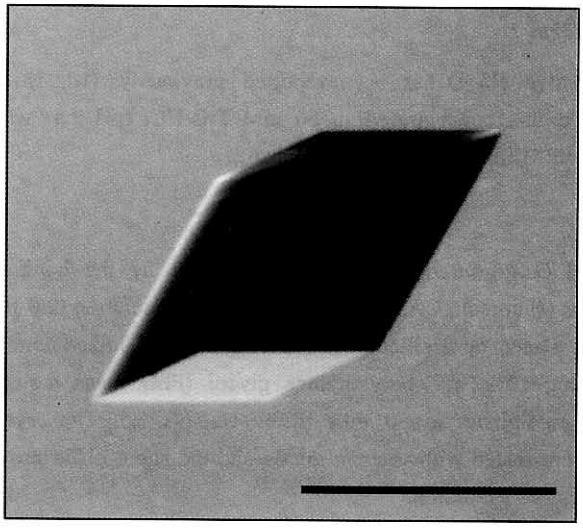


Figure 1. A micrograph of the *G. gigantea* Hb D crystal (a bar : 0.2 mm).

Table 1. Data collection and processing statistics. Note. Values in parentheses are for the highest resolution shell (2.09-2.02 Å).

Space group	C2
Unit cell dimensions	
a (Å)	112.1
b (Å)	62.4
c (Å)	54.0
ß (°)	110.3
Resolution (Å)	53.65 - 2.02
No. of recorded observations	63949
No. of unique reflections	22787
R_{merge} ($\Sigma_h\Sigma_i\|I_{hi} - \bar{I}_h\|/\Sigma_h\Sigma_i\bar{I}_h$)	0.040 (0.135)
Completeness (%)	98.8 (100.0)
Multiplicity	2.78 (2.71)
I /σI	11.8 (3.3)

ACKNOWLEDGEMENTS

We gratefully thank Dr. A. Yamano and Mr. K. Miyazaki of Rigaku Co., for their help in data collection. We express our sincere thanks to Prof. N. Inokuchi, Prof. K. Watanabe, Dr. K. Kusama, Dr. A. Suda, and the entire staff of Inokuchi and Watanabe laboratories, College of Pharmacy, Nihon University, for the use of their laboratory facilities. This work was supported in part by a Grant from the Ministry of Education, Culture, Sports, Science and Technology to promote multi-disciplinary research projects.

REFERENCES

[1] Di Prisco, G., Condo, S.G., Tamburrini, M. and Giardina, B. (**1991**) *Trends Biochem. Sci., 16*, 471-474.
[2] Maina, J.M. (**2000**) *J. Exp. Biol., 203*, 3045-3064.
[3] Rucknagel, K.P. and Braunitzer, G. (**1988**) *Biochem. J., 316*, 959-965.
[4] Fushitani, K., Higashiyama, K., Moriyama, E.N., Imai, K. and Hosokawa, K. (**1996**) *Mol. Biol. Evol., 13*, 1039-1043.
[5] Gorr, T.A., Mable, B.K. and Kleinschmidt, T. (**1998**) *J. Mol. Evol., 47*, 471-485.
[6] Brown, J.L. and Ingram, V.M. (**1974**) *J. Biol. Chem., 249*, 3960-3972.
[7] Moss, B.A. and Hamilton, E.A. (**1974**) *Biochim. Biophys. Acta, 371*, 379-391.
[8] Grubb, P. (**1971**) *Phil. Trans. Roy. Soc. Lond. B, 260*, 327-372.
[9] Van Denburgh, J. (**1914**) *Proc. Calif. Acad. Sci., 2*, 203-374.
[10] Shishikura, F. (**2002**) *Zool. Sci., 19*, 197-206.
[11] Chapman, B.S., Hood, L.E. and Tobin, A.J. (**1982**) *J. Biol. Chem., 257*, 651-658.
[12] Cobb, J.A., Manning, D., Kolatkar, P.R., Cox, D.J. and Riggs, A.F. (**1992**) *J. Biol. Chem., 267*, 1183-1189.
[13] Morrow, J.S., Wittebort, R.J. and Gurd F.R.N. (**1974**) *Biochem. Biophys. Res. Commun., 60*, 1058-1065.
[14] Riggs, A.F. (**1998**) *J. Exp. Biol., 201*, 1073-1084.
[15] Huisman, T.H.J., Van Veen, J.M.S., Dozy, A.M. and Nechtman, C.M. (**1964**) *Biochim. Biophys. Acta, 88*, 352-366.
[16] Knapp, J.E., Oliveira, M.A., Xie, Q., Ernst, S.R., Riggs, A.F. and Hackert, M.L. (**1999**) *J. Biol. Chem., 274*, 6411-6420.
[17] Surtherland-Smith, A.J., Baker, H.M., Hofmann, O.M., Brittain, T. and Baker, E.N. (**1998**) *J. Mol. Biol., 280*, 475-484.
[18] Shishikura, F. and Takami, K. (**2001**) *Zool. Sci., 18*, 515-526.

Received on February 11, 2003, accepted on May 10, 2003.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 315 (2004) 815–822
BBRC
www.elsevier.com/locate/ybbrc

A novel microperoxidase activity: methyl viologen-linked nitrite reducing activity of microperoxidase☆

Kohei Suruga,[a] Kiriko Murakami,[a] Yohei Taniyama,[a] Toshiyuki Hama,[a] Hirotaka Chida,[a] Tadashi Satoh,[b] Seiji Yamada,[c] Wataru Hakamata,[d] Ryu Kawachi,[a] Yasuhiro Isogai,[e] Toshiyuki Nishio,[a] and Tadatake Oku[a,*]

[a] *Department of Biological Chemistry, College of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa, Kanagawa 252-8510, Japan*
[b] *Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization (KEK), Tsukuba, Ibaraki 305-0801, Japan*
[c] *RIKEN, Harima Institute/SPring-8, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan*
[d] *Division of Organic Chemistry, National Institute of Health Sciences, Setagaya Tokyo 158-8501, Japan*
[e] *RIKEN, Hirosawa 2-1, Wako, Saitama 351-0198, Japan*

Received 13 January 2004

Abstract

To investigate the nitrite reducing activity of microperoxidases (mps) in the presence of methyl viologen and dithionite, the fragments C14-K22 (mp9), V11-L32 (mp22), and G1-M65 (mp65) containing heme were prepared by enzymatic hydrolysis of commercially equine heart cytochrome *c* (Cyt *c*), in which His is axially coordinated to heme iron, and acts as its fifth ligand. The nitrite reducing activity of mps was measured under anaerobic condition, and the nitrite reducing activity of mps increased with the cutting of the peptide chain. The activity of the shortest nonapeptide mp9 was approximately 120-fold that of Cyt *c* (104 amino acid residues) and 3.2-fold that of nitrite reductase (EC 1.7.7.1) from *Escherichia coli*. In the nitrite reduction by mp, nitrite was completely reduced to ammonia. We presumed that ferrous mps reduced NO_2^- to NO by donating one electron, the NO was completely reduced to NH_4^+ under anaerobic condition via ferrous–NO complexes as a reaction intermediate using visible spectra and ESR spectra, and this overall reaction was a 6-electron and 8-proton reduction. Sepharose-immobilized mp9 had a nitrite reducing activity similar to that of mp9 in solution, and the resin retained the activity after five uses and even 1-year storage. The mp will be able to use as a substitute for nitrite reductase.

Keywords: Microperoxidase; Cytochrome *c*; Nitrite reducing activity; Hemoprotein; Peptide chain

In a natural nitrogen cycle, nitrite (NO_2^-) is reduced by two types of nitrite reductase. First, in dissimilatory reduction, also called denitrification, NO_2^- is used as a respiratory terminal substrate, and two types of nitrite reductase, copper protein [1], and Cyt cd_1 [2], are involved. Second, NAD(P)H–nitrite reductase (EC 1.7.99.3), ferredoxin–nitrite reductase (EC 1.7.7.1), and Cyt *c* nitrite reductase (EC 1.7.2.2), also known as assimilatory nitrite reductases, catalyze the reduction of NO_2^- to ammonia (NH_4^+) [3–5]. The NH_4^+ reduced by these assimilatory nitrite reductases is mainly used in the synthesis of amino acids.

A heme-containing peptide microperoxidase (mp) is prepared by the proteolytic digestion of equine heart Cyt *c*. The two cysteines (Cys14 and Cys17) of mp are covalently attached to the iron (III)-protoporphyrin IX, and histidine (His) 18 is coordinated to heme iron. Thus far, there are some reports about the amino acid composition and carbon monoxide (CO) reactivity of mps. For example, Santucci et al. [6] reported that the heme-containing undecapeptide (Val11-Glu21), also called mp11, is prepared by the pepsin-catalyzed hydrolysis of equine heart Cyt *c*, and mp11 reacts with free His or His-composed α-helix to investigate the reaction

☆ *Abbreviations:* mp, microperoxidase; Cyt *c*, cytochrome *c*; NO_2^-, nitrite; NO, nitric oxide; CD, circular dichroism; ESR, electron spin resonance; Mb, myoglobin; Hb, hemoglobin.

* Corresponding author. Fax: +81-466-84-3950.
E-mail address: oku@brs.nihon-u.ac.jp (T. Oku).

0006-291X/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2004.01.133

of mp11 heme iron and His. As a result, the His-composed α-helix bound to the sixth position of the heme iron of mp11, and the structural stability of the mp11 complex and His-composed α-helix was higher than that of mp11 and free His. There are also some reports that the heme of mp11 reacts with CO, and the visible absorption maxima at 564 (α), 533 (β), and 413 (soret) nm have been determined [7–9]. Ricoux et al. [10] reported that the octapeptide mp8 (Cys14-Glu21) obtained by peptic and tryptic digestion of equine heart Cyt *c* reacts with the nitrosoalkane (RNO) produced by nitroalkane (RNO_2) reduction, and the new absorption maxima at 415 nm (soret peak) were obtained by the formation of the complex between mp8 and RNO. Moreover, the heme nonapeptide mp9 (Cys14-Lys22) prepared by the trypsin-catalyzed hydrolysis of equine heart Cyt *c* shows a peroxidase-like activity, and its mp9 activity is approximately 1/50 that of horseradish peroxidase [11]. However, the NO_2^- reducing activity of mp and the relationship between the length of the peptide chain of mp and the NO_2^- reducing activity have not yet been reported. Nitrite reductase does not come into the market, because of low content of the enzyme in plants and microbes, and of requirement of a lot of time for the purification. We have tried to prepared a substitute for nitrite reductase by cutting of the peptide chain of commercial equine heart Cyt *c* and determine NO_2^- in an aqueous solution.

In our laboratory, the structural/functional relationships of *c*-type Cyt have been investigated, and the crystal structure of Cyt c_6 from the red alga *Porphyra yezoensis* has been determined at 1.57 Å resolution (PDB code: 1gdv) [12]. We reported that the conformational stability of M58C mutant in which the sixth ligand Met58 of *P. yezoensis* Cyt c_6 was replaced by cysteine to be approximately 2.2-fold that of nonmutant Cyt c_6 [13]. Moreover, the NO_2^- reducing activity of equine heart Cyt *c* that reacts at 100 °C for 30 min is ca. fivefold that of native Cyt *c*, and we presumed that the NO_2^- reducing activity of Cyt *c* increases with heat treatment, because of the following intramolecular changes that occur in Cyt *c* with heat denaturation: (1) unfolding of the peptide chain, (2) exposure of heme to the solvent, (3) dissociation of the sixth ligand (Met80) from heme–iron, and (4) autoxidation [14].

In this work, mp9 (C14-K22), mp22 (V11-L32), and mp65 (G1-M65) were prepared from equine heart Cyt *c* (104 residues) to investigate NO_2^- reducing activity as a novel activity, and the relationship between the activity and the length of the peptide chain of mp. The physicochemical properties of these mps were investigated using UV/visible spectra, CD spectra, and redox titration. Moreover, mp9 which showed the highest NO_2^- reducing activity was immobilized in CNBr–Sepharose and acrylamide, and its NO_2^- reducing activity was measured.

Materials and methods

Preparation of microperoxidase 9, 22, and 65. Mp9 and 22 were prepared according to Plattner et al. [11] and Cheek et al. [15]. For the preparation of mp9 and mp22, equine heart Cyt *c* (Wako Pure Chemical Industries) was digested by incubation of Cyt *c* in 0.1 M Tris–HCl buffer (pH 8.0) containing 0.5 M urea, with trypsin for mp9 or chymotrypsin for mp22 (protein substrate:protease ratio 1:50 by mass) at 37 °C for 24 h. Mp65 was prepared by treatment of intact equine heart Cyt *c* with cyanogen bromide at 20 °C for 4 h [16]. The reaction mixture was loaded on a Toyopearl HW-40F gel filtration column (Tosoh; 1.0 × 85 cm). The purity of mp9 and 22 was checked using a BioCAD 700E perfusion chromatograph equipped with a Poros R2/20 column (Applied Biosystems: 4.6 × 100 mm), and the purity check of mp65 was performed by tricine SDS–PAGE.

Measurement of nitrite reducing activity. The NO_2^- reducing activity of mps was determined according to Vega [17]. The vial containing 0.675 ml of 100 mM sodium phosphate buffer (pH 7.0), 0.4 ml of 10 mM sodium nitrite, 0.5 ml of 3 mM methyl viologen, and 0.125 ml of 20 μM mp solution was sealed with a butyl rubber cap. After pre-incubating the mixture at 37 °C for 5 min, the reaction was initiated by adding 0.3 ml of 100 mM sodium dithionite dissolved in 50 mM sodium bicarbonate. The reaction was conducted under anaerobic condition at 37 °C. The reaction mixture was placed in a test tube, left to stand for several minutes, and then vigorously shaken until complete decolorization.

NO_2^- detection was performed according to Ramirez et al. [18]. For NO_2^- detection, 1.95 ml water, 1 ml of 1% sulfanilamide, 1 ml of 0.02% *N*-1-naphthylethylenediamine, and 1 ml pure water were added to 50 μl of the reaction mixture. After allowing the reaction mixture to stand at room temperature for 20 min, its optical density at 540 nm was determined. One unit of activity was defined as the amount of enzyme needed to reduce 1 μmol of NO_2^- per minute. The k_{cat} value was calculated from Lineweaver–Burk plots.

NH_4^+ detection was based on the methods described by Scheiner et al. [19]. For the NH_4^+ detection 1.9 ml water, 1 ml of 50 mg% sodium nitroprusside in 10% phenol, 1 ml of 0.06% sodium hypochlorite solution dissolved in 0.1 M disodium hydrogen phosphate, 0.25 M sodium hydroxide, and 1 ml pure water were added to 100 μl of the reaction mixture. After allowing the reaction mixture to stand at room temperature for 60 min, its optical density at 630 nm was determined.

The inhibition of NO_2^- reducing activity of mps was determined according to Vega [17].

Spectroscopic analysis. The visible spectra of mps were monitored with a Hitachi U3310 spectrophotometer using quartz cuvettes of 1.0-cm path length at 25 °C. Mp concentration was determined using a pyridine hemochromogen method [20]. Circular dichroism (CD) spectra were recorded at 25 °C in 20 mM sodium phosphate buffer (pH 7.0), with a JASCO J-700 spectropolarimeter using 0.2-cm path length rectangular quartz cuvettes [21]. Electron spin resonance (ESR) spectrum of mp9 was recorded at cryogenic temperatures with a JOEL ESR spectrometer, JES-FA200.

Redox titration study. The redox titrations of mps were performed under anaerobic condition, with a continuous stream of argon, in 100 mM sodium phosphate buffer, pH 7.0, at 25 °C [22]. The potentials were measured with a Horiba F-13 pH meter equipped with an ORP electrode, and the optical spectra were monitored throughout the titration on a Hitachi U3310 spectrophotometer. The redox mediators were used to stabilize the solution redox potential as described by Yamada et al. [14]. The redox data were analyzed with a theoretical curve based on the Nernst equation ($n = 1$): $E = E^0 + (RT/nF)$ ln([ferric]/[ferrous]) [23].

Preparation of Sepharose-immobilized and acrylamide-immobilized microperoxidase. Sepharose-immobilized mp9 and acrylamide-immobilized mp9 were prepared according to the Shin and Oshino [24]. The

binding rate of mp9 to CNBr–Sepharose 4B and the rate of inclusion of mp9 in acrylamide were 99.98% and 99.78%, respectively.

Results and discussion

Physicochemical properties of microperoxidase

The UV/visible spectra of mp65 (G1-M65), mp22 (V11-L32), and mp9 (C14-K22) are shown in Fig. 1. For mp65, its ferric forms showed visible absorption maxima at 535 ($\alpha + \beta$) and 404 (soret peak) nm, and its ferrous forms at 549 (α), 520 (β), and 415 (soret peak) nm (Fig. 1B). These absorption peaks of mp65 indicate that mp65 has a 6-coordination low spin [25]. The absorption peak at 695 nm that was observed in Met80 sulfur charge transfer to heme iron disappeared in the ferric forms of mp65. There are three His residues (His18, His26, and His32) in equine heart Cyt *c*. His18 is axially coordinated to heme iron and acts as its fifth ligand [26]. Santucci et al. [27] reported that the fragment G1-G56 was obtained by the thermolysin-catalyzed hydrolysis of equine heart Cyt *c*, and this fragment has two His residues (His18 (fifth ligand) and His26 or His33 (sixth ligand)) axially bound to the heme iron. From this study, it was considered that two His residues (His18 and His26 or His33) are coordinated to heme iron in mp65. For mp22, its ferric forms showed visible absorption maxima at 398 nm (soret peak), and its ferrous forms at 549 (α), 520 (β), and 416 (soret peak) nm (Fig. 1C). This indicates that mp22 has a 6-coordination low spin with His18 and His26 as axial ligands. For mp9, its ferric forms showed visible absorption maxima at 619 (high spin marker band) and 395 (soret peak) nm, and its ferrous forms at 549 (α) and 412 (soret peak) nm; moreover, the absorption peak at 520 (β) disappeared in the ferrous forms of mp9 (Fig. 1D). The absorption peak at 619 nm of the ferric forms of mp9 was consistent with those of deoxy-Mb [28] and mp11 [29], which have a 5-coordination high spin with His as an axial ligand. The absorption peak at 520 (β) nm disappeared in the ferrous forms of the M80A Cyt *c* mutant [25], similar to the case of mp9. From these, it was assumed that mp9 has a 5-coordination high spin with His18 as the fifth ligand.

The physicochemical properties of mp65, mp22, and mp9 are summarized in Table 1. The isoelectric points (pI) of mp65, mp22, and mp9 were 9.52 (theoretical value: 10.06), 6.02 (9.72), and 4.28 (6.99), respectively,

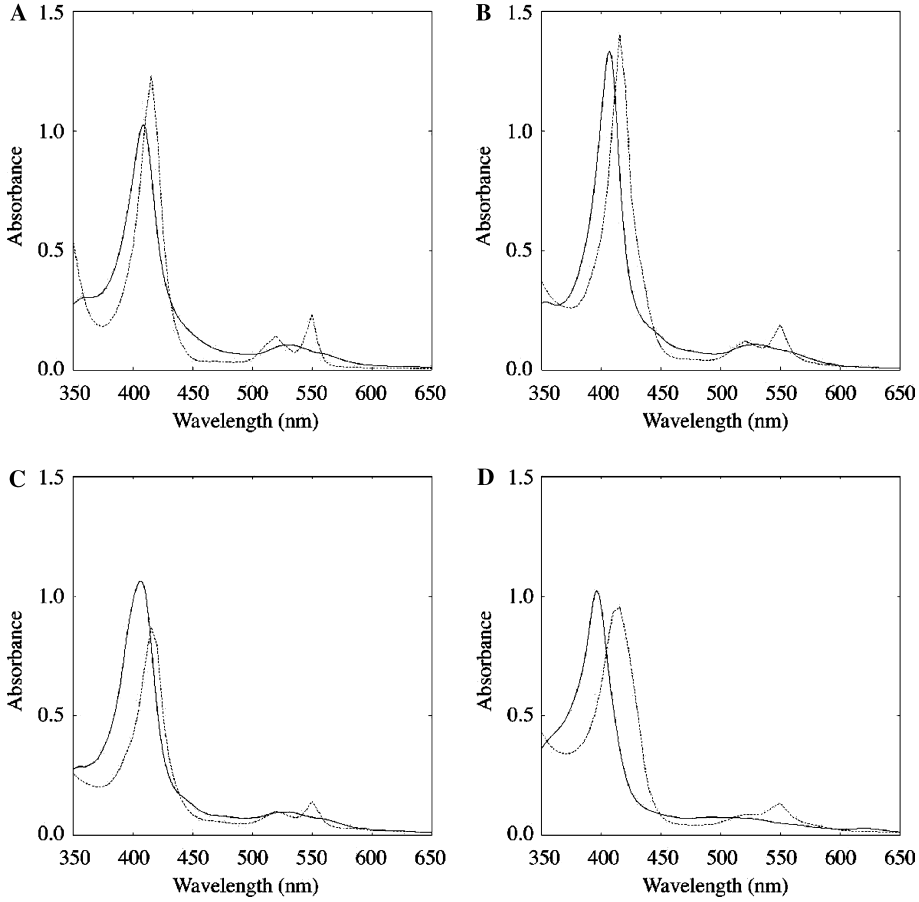


Fig. 1. UV/visible spectra of the ferric (solid line) and ferrous (broken line) forms of Cyt *c* and mps. Absorption spectra of 10 μM Cyt *c* and mps were measured in 10 mM sodium phosphate buffer (pH 7.0) at 25 °C. (A) Cyt *c*; (B) mp65; (C) mp22; and (D) mp9.

Table 1
Physicochemical properties of mps prepared from equine heart Cyt *c*

	mp9	mp22	mp65	Cyt *c*
p*I*	4.28	6.02	9.52	10.50
Absorption maxima (nm)				
Ferric (Fe^{III})	395	398	404	409
	619		535	535
Ferrous (Fe^{II})	412	416	415	415
		520	520	520
	549	549	549	550
E^0 (mV, pH 7.0, 25 °C)	−132	−67	−62	260
MW	1630	3065	8900	12,500

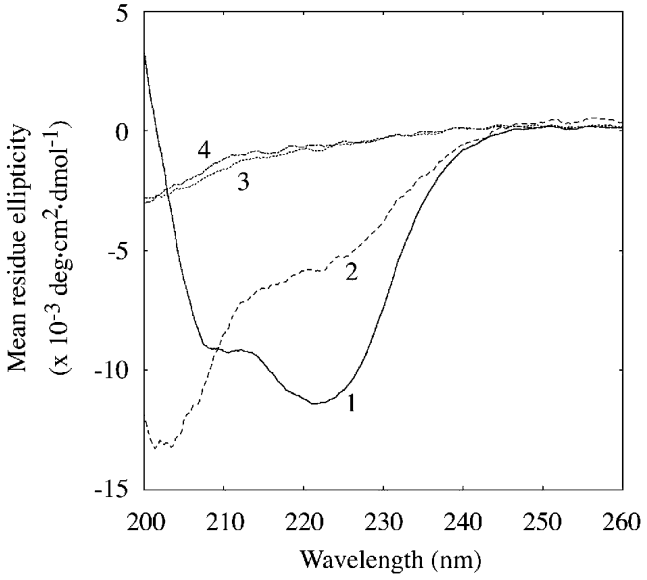


Fig. 2. Circular dichroism spectra of Cyt *c* and mps. The spectra of Cyt *c* and mps in 10 mM sodium phosphate buffer (pH 7.0) at 25 °C were recorded. Lane 1: Cyt *c*; 2: mp65; 3: mp22; and 4: mp9.

and the pI of mp9 was the same as that of mp11 [11]. The redox potentials of mp65, mp22, and mp9 were −62, −67, and −132 mV, and were 322, 327, and 392 mV lower than that of Cyt *c* (260 mV), respectively. The redox potential of mp9 was similar to that of mp11 (−134 mV) [30].

The secondary structures of mp65, mp22, and mp9 were examined, and the far-UV CD spectra of these mps were measured (Fig. 2). It was found that mp9 (C14-K22) has only random coil conformations, because it has no α-helix segment of equine heart Cyt *c* (Figs. 2–4). Mp22 (V11-L32) has the helix II (K22-K27) segment of equine heart Cyt *c*, but its CD spectra agreed closely with that of mp9, which has no α-helix (Figs. 2–3). From these results, it was assumed that mp22 has a random coil conformation. The CD spectra of mp9 and 22 were almost the same as that of mp11, which has a random coil conformation [31]. Mp65 (G1-M65) has helix I (G1-A15), helix II (K22-K27) and helix III (D50-G56) segments, and part of the helix IV (E61-E69) segment. From the CD spectra, the absorbance at 222 nm of mp65 was ca. 1/2 that of equine heart Cyt *c* (Fig. 2, lane 1), but the absorbance at 200 nm which corresponds to the random coil configuration increased (Fig. 2, lane 2). This result was similar to that of the fragment G1-G56 that was prepared by the thermolysin-catalyzed hydrolysis of equine heart Cyt *c* [27], and it was considered that the α-helices of mp65 are unfolded. From the results of redox potential and CD spectra, we assumed that the heme of mps would be exposed to solvent with cutting of the peptide chain.

Nitrite reducing activity of microperoxidase and its reaction mechanism

The NO_2^- reducing activity of mp65, mp22, and mp9 was measured in the presence of methyl viologen and dithionite under anaerobic condition, and the NO_2^- reducing activity of the mps increased with the cutting of the peptide chain (Fig. 3). The reason for this is given in

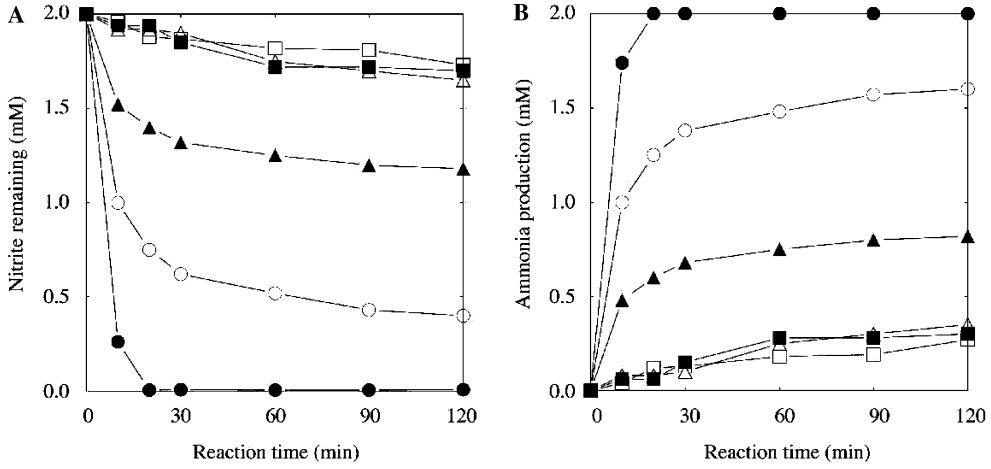


Fig. 3. Nitrite reduction and ammonia production of Cyt *c* and mps. The nitrite reducing activities of 1.25 μM Cyt *c* and mps were measured at 37 °C under anaerobic condition in the presence of dithionite and methyl viologen. (□) ferrous sulfate; (■) hemin; (△) Cyt *c*; (▲) mp65; (○) mp22; and (●) mp9. (A) nitrite reduction; (B) ammonia production.

the middle of this section. Both the NO_2^- reducing activities of ferrous sulfate which composes mp, and hemin which has no peptide chains were approximately 1/150 that of mp9 (Fig. 3). Generally known, the hemin was insoluble in all solvents except aqueous strong base. The k_{cat} (s^{-1}) values of equine heart myoglobin (Mb) and the fragment (V17-K46, 26 residues) prepared by the lysyl-endopeptidase-catalyzed hydrolysis of equine heart Mb were 0.004 and 0.018, respectively. These values were similar to that of Cyt *c* (k_{cat} (s^{-1}) = 0.015), but were approximately 1/460 and 1/100 that of mp9. In *c* type Cyt, the protoheme is covalently bonded via thioether to the two cysteines (Cys14 and 17) of the peptide chain [32], but *b*-type hemoproteins such as Mb and hemoglobin (Hb) have no thioether linkage [33]. From these, it was proved that the covalent bond between heme and Cys may be necessary for enhancing of the NO_2^- reducing activity of mp. A kinetic study of mp for determining its NO_2^- reducing activity was conducted, and the results are shown in Table 2. The k_{cat} (s^{-1}) values of mp65, 22, and 9 were 0.120, 0.800, and 1.834, respectively. The NO_2^- reducing activity of the shortest nonapeptide mp9 was approximately 1/27 and 1/6 that of the spinach nitrite reductase (k_{cat} (s^{-1}) = 49.650) [34] and *P. yezoensis* nitrite reductase (k_{cat} (s^{-1}) = 10.000) [35], respectively, but this activity was ca. 3.2-fold that of the nitrite reductase from *Escherichia coli* (k_{cat} (s^{-1}) = 0.570) [36] and 120-fold that of Cyt *c* (104 amino acid residues).

Table 2
Kinetic study on nitrite reducing activity of Cyt *c* and mps

	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM/s^{-1})
Cyt *c*	2.453	0.015	0.006
mp65	2.140	0.120	0.056
mp22	1.830	0.800	0.440
mp9	1.420	1.834	1.292

From these results, it was assumed that the lower molecular in prepared mps, which were water-soluble heme peptide, showed the higher NO_2^- reducing activity, because the collision probability of mps to NO_2^- would be increased with the cutting of the peptide chain. As shown in Figs. 1B and C (see preceding section), mp65 and mp22 have a 6-coordination. On the other hand, mp9 has a 5-coordination and the sixth position of mp9 was vacant (Fig. 1D). Consequently, the shortest

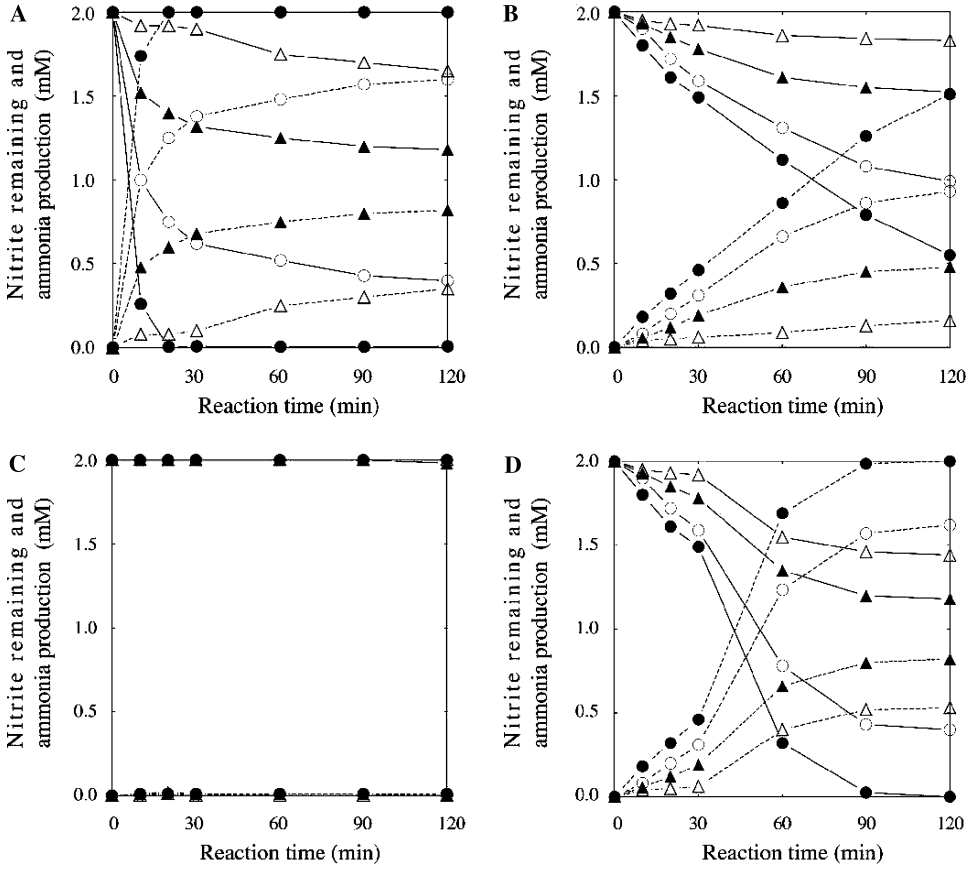


Fig. 4. (A) Nitrite reduction and ammonia production of Cyt *c* and mps. (B) Effect of methyl viologen on nitrite reducing activity of Cyt *c* and mps. (C) Effect of dithionite on nitrite reducing activity of Cyt *c* and mps. (D) Effect of methyl viologen on nitrite reducing activity of Cyt *c* and mps. Methyl viologen was added to reaction vial at 30 min. (△) Cyt *c*; (▲) mp65; (○) mp22; and (●) mp9. Solid line: nitrite reduction; Dashed line: ammonia production.

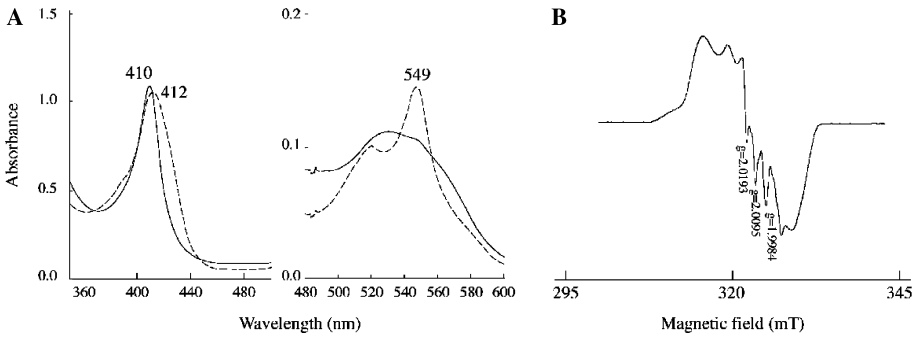


Fig. 5. (A) Visible spectra of ferrous mp9 reacted with nitrite. Ferrous mp9 (dashed line) was prepared by the addition of portion of dithionite under anaerobic conditions at pH 7.0. Then a portion of 1 M sodium nitrite was added (solid line). (B) Electron spin resonance spectrum of mp9 during the nitrite reducing activity. The spectrum was monitored under liquid nitrogen.

nonapeptide mp9 had the highest NO_2^- reducing activity in prepared mps.

In the NO_2^- reduction by mp, NO_2^- was reduced to NH_4^+ in the presence of methyl viologen and dithionite under anaerobic condition, and the conversion rate was approximately 100% (Fig. 4A). This reaction was the same as that of the assimilatory nitrite reductases from spinach [34], *P. yezoensis* [35], and *E. coli* [36]. When the methyl viologen as electron carrier was absent in the reaction mixture, the NO_2^- reducing activity of mp was approximately 1/45 that of mp in the presence of methyl viologen (Fig. 4B). When the sodium dithionite as reductant was absent in the reaction mixture, the mp did not show the NO_2^- reducing activity at all (Fig. 4C). After the addition of methyl viologen to the reaction mixture at 30 min, the NO_2^- reduction and NH_4^+ production were rapidly proceeded (Fig. 4D).

The visible spectra of reaction products of mp9 with NO_2^- were measured at pH 7.0 (Fig. 5A). The ferric mp9 did not react with NO_2^- at pH 7.0 (data not shown), but the spectral changes of the ferrous mp9 reacted with NO_2^- were observed. The soret peak shifted to 410 nm (ferrous mp9–NO_2^- complex) from 412 nm (ferrous mp9) and the α-peak (549 nm) disappeared, and these spectral changes were also recognized in the case of mp11 reacted with CO [7]. These spectral changes were probably due to the formation of NO_2^-–Fe–His coordination. The ESR spectrum of mp9 during the NO_2^- reducing activity was measured under cryogenic temperature (Fig. 5B). The ESR spectrum showed a hyperfine structure near and at $g = 2$, and this indicated the ferrous–NO complex [37]. From this, it was assumed that the ferrous–NO complexes were a reaction intermediate in the NO_2^- reducing activity of mps.

From these results, we presumed that ferrous mps reduced NO_2^- to NO by giving one electron, and the NO was completely reduced to NH_4^+ under anaerobic condition via ferrous–NO complexes as a reaction intermediate. This overall reaction was a 6-electron and 8-proton reduction as well as that of the nitrite reductase [34–36].

The pH dependence of the NO_2^- reducing activity of mp9 was studied. The optimum pH for NO_2^- reducing activity of mp9 was 7.0 with sodium phosphate buffer (Fig. 6). This value was close to those of the nitrite reductase from spinach (pH 7.5), and *P. yezoensis* (pH 7.5), [35].

The effect of various inhibitors on the NO_2^- reducing activity of mp9 was investigated (Table 3). Potassium cyanide (KCN) was found to be effective inhibitor, and at a KCN concentration of 2 mM, the mp9 was more than 99% inhibition. When the CO was added to the reaction mixture, CO inhibited the NO_2^- reducing activity of mp9 completely. These results were similar degree to those of nitrite reductase from spinach [37,38]. A metal chelating reagent, *o*-phenanthroline, and EDTA showed no inhibition of NO_2^- reducing activity of mp9.

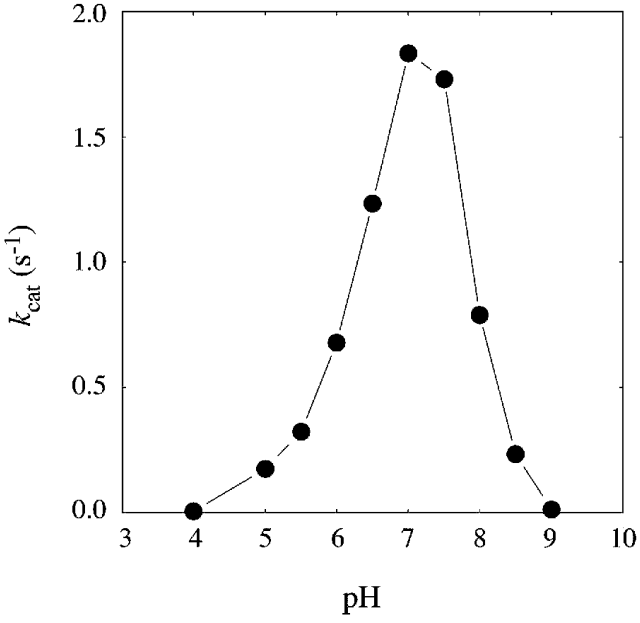


Fig. 6. The pH profile of the nitrite reduction of mp9. Each pH was maintained with sodium citrate (pH 4–5), sodium phosphate (pH 6–7), Tris–chloride (pH 8), and glycine–NaOH (pH 9).

Table 3
Effect of some inhibitors on nitrite reducing activity of mp9

Inhibitors	Concentration (mM)	Inhibition (%)
Potassium cyanide	0.02	91.18
	0.2	92.40
	2.0	99.13
Carbon monoxide	Saturated	100.0
o-Phenanthroline	0.02	0.00
	0.2	0.00
	2.0	0.00
EDTA	0.02	0.00
	0.2	0.00
	2.0	1.69

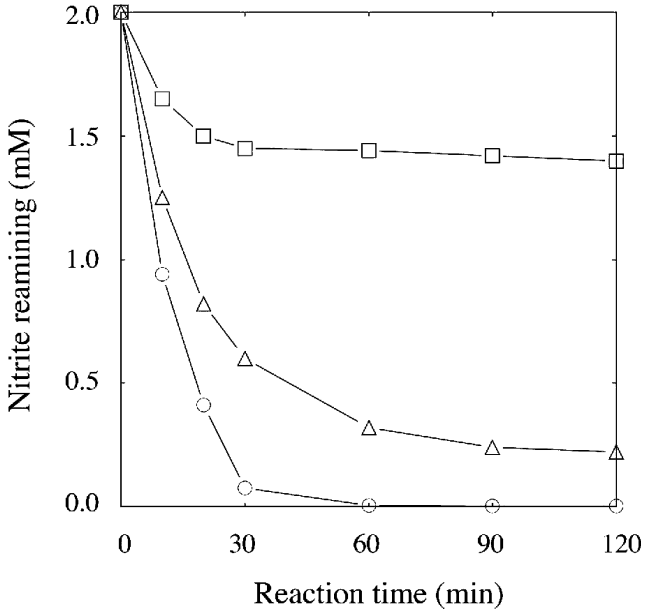


Fig. 7. Nitrite reducing activity of immobilized mp9. (○) free mp9; (△) Sepharose-immobilized; and (□) acrylamide-immobilized.

These results were close to those of nitrite reductase from spinach and green alga *Chlorella fusca* [37,39].

Immobilized microperoxidase and its nitrite reducing activity, stability, and reuse

As a rule, proteins and enzymes are unstable at high temperatures and their recovery after a reaction is difficult. We attempted to improve the reusability and stability of mp9, which showed the highest NO_2^- reducing activity, and prepared the mp9 immobilized in CNBr–Sepharose or acrylamide. The NO_2^- reducing activity of CNBr–Sepharose or acrylamide-immobilized mp9 was then measured. The NO_2^- reducing activity of CNBr–Sepharose-immobilized mp9 was fourfold that of acrylamide-immobilized mp9, and CNBr–Sepharose-immobilized mp9 had nearly the same activity as soluble mp9 (Fig. 7). Heme in CNBr–Sepharose-immobilized mp9 was exposed to a solvent, but, in the case of the acrylamide-immobilized mp9, mp9 is surrounded by a matrix. It is considered that the matrix prevents its contact with NO_2^- or an electron donor [14]. Thus, it was presumed that the NO_2^- reducing activity of CNBr–Sepharose-immobilized mp9 is higher than that of acrylamide-immobilized mp9. The residual activities of CNBr–Sepharose and acrylamide-immobilized mp9 were 99.0% and 99.1%, respectively, even after storage at 4 °C for 1 year. The activity of CNBr–Sepharose-immobilized mp9 after the fifth use was 99.4%, and thus mp9 is reusable. Moreover, the immobilized mp9 also showed the NO reducing activity in addition to NO_2^- reducing activity, and the activity was nearly the same as that of soluble mp9 (data not shown).

Conclusion

In this work, we reported for the first time that the mps prepared by enzymatic hydrolysis of commercial equine heart Cyt *c* showed the NO_2^- reducing activity as a novel activity. The NO_2^- reducing activity of mps increased with the cutting of the peptide chain, and mp9 after use had the same NO_2^- reducing activity as that of freshly mp9 by immobilization. This mp may thus be a new molecule for simple applications such as the determination of NO_2^- in an aqueous solution and will be able to use as a substitute for nitrite reductase.

Acknowledgments

We are indebted to Misses Kanako Kashiwai and Mayumi Takaku, and Messrs. Koichi Kamiuse and Osamu Tozawa, Nihon University, for the preparation of microperoxidases, and Dr. Toshiki Masumizu, JEOL Ltd., for measurement of the ESR spectrum. This work was supported in part by a Grant-in-Aid for the 21 Century's Center of Excellence (COE) Program, for Scientific Research (C) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and for the Research for the Future Program (JSPS-RFTF97R16001) from the Japan Society for the Promotion of Science.

References

[1] T. Matsubara, Studies on denitrication. XII. Gas production from amines and nitrite, J. Biochem. (Tokyo) 67 (1970) 229–235.
[2] P.R. Alefounder, S.J. Ferguson, The location of dissimilatory nitrite reductase and the control of dissimilatory nitrate reductase by oxygen in *Paracoccus denitrificans*, Biochem. J. 192 (1980) 231–240.
[3] R.A. Lazzarini, D.E. Atkinson, A triphospyridine nucleotide-specific nitrite reductase from *Escherichia coli*, J. Biol. Chem. 256 (1961) 9711–9717.
[4] S. Ida, Y. Morita, Purification and general properties of spinach leaf nitrite reductase, Plant Cell Physiol. 14 (1973) 661–671.

[5] O. Eisle, A. Messerschmidt, P. Stach, G.P. Bourenkov, H.D. Bartunik, R. Huber, P.M.H. Kroneck, Structure of cytochrome *c* nitrite reductase, Nature 400 (1999) 476–480.
[6] R. Santucci, A. Picciau, G. Antonini, L. Campanella, A complex of microperoxidase with a synthetic peptide: structure and functional characterization, Biochim. Biophys. Acta 1250 (1995) 183–188.
[7] M. Laberge, A.J. Vreugdenhil, I.S. Butler, Microperoxidase-11: molecular dynamics and Q-band excited resonance Raman of the oxidized, reduced and carbonyl forms, J. Biomol. Struct. Dyn. 15 (1998) 1039–1050.
[8] V.S. Sharma, H.M. Renney, J.F. Geibel, T.G. Traylor, A new for the determination of ligand dissociation rate constant of carboxyhemoglobin, Biochem. Biophys. Res. Commun. 66 (1975) 1301–1306.
[9] V.S. Sharma, M.R. Schmidt, H.M. Ranney, Dissociation of CO from carboxyhemoglobin, J. Biol. Chem. 251 (1976) 4267–4272.
[10] R. Ricoux, J.L. Boucher, D. Mansuy, J.P. Mehy, Formation of iron (II)-nitrosoalkane complex: a new activity of microperoxidase 8, Biochem. Biophys. Res. Commun. 278 (2000) 217–223.
[11] H. Plattner, E. Wachter, P. Grobner, A heme-nonapeptide tracer for electron microscopy, Histochemistry 53 (1977) 223–242.
[12] S. Yamada, S.-Y. Park, H. Shimizu, Y. Koshizuka, K. Kadokura, T. Satoh, K. Suruga, M. Ogawa, Y. Isogai, T. Nishio, Y. Shiro, T. Oku, Structure of cytochrome c_6 from the red alga *Porphyra yezoensis* at 1.57 Å resolution, Acta Cryst. D 56 (2000) 1577–1578.
[13] T. Satoh, A. Itoga, Y. Isogai, M. Kurihara, S. Yamada, M. Natori, N. Suzuki, K. Suruga, R. Kawachi, M. Arahira, T. Nishio, C. Fukazawa, T. Oku, Increasing the conformational stability by replacement of heme axial ligand in *c*-type cytochrome, FEBS Lett. 531 (2002) 543–547.
[14] S. Yamada, K. Suruga, M. Ogawa, T. Hama, T. Satoh, R. Kawachi, T. Nishio, T. Oku, Appearance of nitrite reducing activity of cytochrome *c* upon heat denaturation, Biosci. Biotechnol. Biochem. 66 (10) (2002) 2044–2051.
[15] J. Cheek, D.W. Low, H.B. Gray, J.H. Dawson, Histidine-tailed microperoxidase-10: a pH-dependent ligand switch, Biochem. Biophys. Res. Commun. 253 (1998) 195–198.
[16] M. Hoshino, Y. Kawata, Y. Goto, Interaction of GroEL with conformational states of horse heart cytochrome *c*, J. Mol. Biol. 262 (1996) 575–587.
[17] J.M. Vega, J. Cardenas, M. Losada, Ferredoxin–nitrite reductase, Methods Enzymol. 69 (1980) 255–270.
[18] J. Ramirez, F.F.D. Campo, A. Paneque, M. Losada, Ferredoxin–nitrite reductase from spinach, Biochim. Biophys. Acta 118 (1966) 58–71.
[19] D. Scheiner, Determination of ammonia and Kjeldahl nitrogen by indophenol method, Water Res. 10 (1976) 31–36.
[20] D.L. Drabkin, Spectrophotometric studies, J. Biol. Chem. 146 (1942) 605–617.
[21] C.N. Pace, B.A. Shirly, J.A. Thompson, Protein Structure: A Practical Approach, IRL, Oxford, 1989, pp. 311–330.
[22] S. Yamada, T. Nakamura, Y. Tanaka, Y. Isogai, T. Nishio, T. Oku, Characterization and amino acid sequences of cytochrome c_6 from two strains of the green alga *Chlorella vulgaris*, Biosci. Biotechnol. Biochem. 64 (2000) 628–632.
[23] P.L. Duton, Redox potentiometry: determination of midpoint potentials of oxidation-reduction components of biological electron-transfer systems, Methods Enzymol. 54 (1978) 411–435.
[24] M. Shin, R. Oshino, Ferredoxin–Sepharose 4B as a tool for the purification of ferredoxin-$NADP^+$ reductase, J. Biochem. (Tokyo) 83 (1978) 357–361.
[25] J.M. Wallace, I. Cardenes, Functional role of heme ligation in cytochrome *c*, J. Biol. Chem. 267 (1992) 3852–3861.
[26] G.W. Bushnell, G.V. Louie, G.D. Brayer, High-resolution three-dimensional structure of horse heart cytochrome *c*, J. Mol. Biol. 20 (1990) 585–595.
[27] R. Santucci, L. Fiorucci, F. Sinibaldi, F. Polizio, A. Desideri, F. Ascoli, The heme-containing N-fragment (residues 1–56) of cytochrome *c* is a bis-histidine functional system, Arch. Biochem. Biophys. 379 (2000) 331–336.
[28] T. Iizuka, T. Yonetani, Spin changes in hemoproteins, Adv. Biophys. 1 (1970) 157–182.
[29] H.M. Marques, C.B. Perry, Hemepeptide models for hemoproteins: the behavior of *N*-acetylmicroperoxidase-11 in aqueous solution, J. Inorg. Biochem. 75 (1999) 281–291.
[30] G. Battistuzzi, M. Borsari, J.A. Cowan, A. Ranieri, M. Sola, Control of cytochrome *c* redox potential: axial ligation and protein environment effects, J. Am. Chem. Soc. 124 (2002) 5315–5324.
[31] W. Huang, Z. Zheling, X. Han, J. Tang, Z. Peng, S. Dong, E. Wang, Electrochemistry and spectroscopy study on the interaction of microperoxidase-11 with lipid membrane, Biophys. Chem. 94 (2001) 165–173.
[32] E. Margoliash, E. Smith, G. Kreil, H. Tuppy, The complete amino-acid sequence, Nature 192 (1961) 1125–1127.
[33] P.A. Cornelius, A.W. Steele, D.A. Chernoff, R.M. Hochstrasser, Different dissociation pathways and observation of an excited deoxy state in picosecond photolysis of oxy-and carboxymyoglobin, Proc. Natl. Acad. Sci. USA 78 (1981) 7526–7529.
[34] C.-H. Ho, G. Tamura, Purification and properties of nitrite reductase from spinach leaves, Agric. Biol. Chem. 37 (1973) 37–44.
[35] M.G. Guerreo, J.M. Vega, M. Losada, The assimilatory nitrite-reducing system and its regulation, Ann. Rev. Plant Physiol. 32 (1981) 169–204.
[36] J.D. Kemp, D.E. Atkinson, Nitrite reductase of *Escherichia coli* specific for reduced nicotinamide adenine dinucleotide, J. Bacteriol. 92 (1966) 628–634.
[37] M. Hirasawasoga, S. Horie, G. Tamura, Further characterization of ferredoxin nitrite reductase and the relationship between the enzyme and methyl viologen-dependent nitrite reductase, Agric. Biol. Chem. 46 (1982) 1319–1328.
[38] J.M. Vega, H. Kamin, Spinach nitrite reductase, J. Biol. Chem. 252 (1977) 896–909.
[39] W.G. Zumft, Ferredoxin: nitrite oxidoreductase from Chlorella purification and properties, Biochim. Biophys. Acta 1040 (1990) 237–244.