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

Authors: Shishikura, Fumio, and Nakamura, Masako

Source: Zoological Science, 13(6) : 849-856

Published By: Zoological Society of Japan

URL: https://doi.org/10.2108/zsj.13.849

ZOOLOGICAL SCIENCE 13: 849–856 (1996)

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

Fumio Shishikura[1,*] and Masako Nakamura[2]

[1]*Department of Biology, Nihon University School of Medicine, Ohyaguchi-kami machi, Itabashi-ku, Tokyo 173, Japan*
[2]*Faculty of Economics, Chuo University, Higashinakano, Hachioji, Tokyo 192-03, Japan*

ABSTRACT—The monomer subunits of giant extracellular hemoglobins from earthworms *Pontodrilus matsushimensis* and *Pheretima communissima* that belong to the family Megascolecidae, Oligochaeta, were purified by a reversed-phase column, Resource RPC, and sequenced. The complete amino acid sequences of the two monomeric globin chains were determined: 141 amino acid residues with a molecular weight of 16,366 Da for *Pontodrilus matsushimensis* and 140 amino acid residues with a molecular weight of 16,000 for *Pheretima communissima*, respectively. The *Pontodrilus matsushimensis* monomer globin has three cysteine residues, and the two located at positions 2 and 131 are conserved as those observed in all annelids and contribute to form a disulfide-bonded interchain. The third cysteine residue at position 73 is the first evidence for the annelid monomer globin subunits. The physiological functions of the third cysteine residue, however, are still unknown. The monomer sequences of the two species were aligned with those of five known sequences from annelids, including a polychaete, *Tylorrhynchus heterochaetus*, and four oligochaetes, *Pheretima hilgendorfi*, *Pheretima sieboldi*, *Lumbricus terrestris* and *Tubifex tubifex*. Using computer analysis, a 87.9% identity of the amino acid sequences between two monomeric subunits of *Pheretima communissima* and *Pheretima hilgendorfi* hemoglobins showed the highest degree of sequence similarity. A molecular phylogenetic tree of seven species of annelids has constructed, suggesting that the divergence times among the three species of *Pheretima* and between *Pheretima* and *Pontodrilus* were 50 to 100 and about 209 million years ago, respectively.

INTRODUCTION

Annelid hemoglobin is a giant extracellular protein comprised of two main heme-containing subunits, a monomer and a disulfide-bonded trimer, and a linker including two or three chains (Vinogradov, 1985). The four constituent chains of two main heme-containing subunits that are classified into two strains of globins, strain A and strain B, by Gotoh *et al.* (1987) are akin to vertebrate α- and β-globin. To date, five monomeric globins that belong to the strain A have been purified from annelid hemoglobins whose amino acid sequences have been established: chain I of *Tylorrhynchus heterochaetus* by Suzuki *et al.* (1982), chain I of *Lumbricus terrestris* by Shishikura *et al.* (1987), chain I of *Pheretima sieboldi* by Suzuki (1989), chain I of *Tubifex tubifex* by Stern *et al.* (1990) and the monomer chain of *Pheretima hilgendorfi* by Shishikura (1996). Compared with all known sequences of annelid hemoglobins described elsewhere (Kapp *et al.*, 1995), hemoglobin molecules sharing an orthologous relationship, such as a lineage of monomer globin chains in strain A, are good materials for analyzing the molecular evolution of proteins and genetic divergence patterns among annelids (Shishikura, 1996).

The *Pheretima* group of genera, family Megascolecidae, contains over 750 nominal species from southern and eastern mainland of Asia and is a dominant member of the earthworm fauna (Easton, 1984). In Japan, there are as many as 155 species of the genus of *Pheretima* (Kobayashi, 1941) including such abundant species as *Pheretima hilgendorfi* and *Pheretima communissima* in the central and northern regions. A previous study (Shishikura, 1996) showed that a 81.6% identity of the amino acid sequence was found between the

* To whom all correspondence should be addressed.

two monomeric globin chains from the *Pheretima* species, *Pheretima hilgendorfi* and *Pheretima sieboldi*. The value, as expected, was the highest degree of sequence similarity among known annelid sequences (Goodman *et al.*, 1988; Vinogradov *et al.*, 1993; Kapp *et al.*, 1995). Hence, here we aimed to sequence another monomer globin chains of the member of *Pheretima*. Consequently, only 17 amino acid residues out of 140 in total were different between *Pheretima hilgendorfi* and *Pheretima communissima*, which accounts for a 87.9% identity.

On the contrary, *Pontodrilus matsushimensis* also belongs to the family of Megascolecidae (Jamieson, 1971, 1988). The genus *Pontodrilus* is a terrestrial earthworm, normally found burrowing in sand around the high tide mark of sea shores (Iizuka, 1897). These worms have a wide tolerance to the salinity and may be different from common terrestrial earthworms in their osmoregulatory functions (Takeuchi, 1980a, b). Such internal conditions may cause adaptation of protein functions involving amino acid substitutions of primary structure (Perutz, 1984; Clementi *et al.*, 1994). Therefore, we also aimed to establish the amino acid sequence of the monomer subunit of *Pontodrilus matsushimensis* hemoglobin and compared its primary structure with known annelid structures. Here we provide significant evidence concerning the *Pontodrilus* monomer globin sequence: The monomer globin chain has three cysteine residues, and the two located at positions 2 and 131 might be responsible for linking together by a disulfide bond, while the other at position 73 is noteworthy in annelid monomer globin subunits but its biological function is still unknown.

In this study we also describe a convenient method for separating individual chains from the whole hemoglobin. A reversed-phase column, Resource RPC, recently introduced by Pharmacia Biotech., showed very powerful globin chain resolution with high yield, in particular for relatively larger molecules such as monomer and trimer subunits of hemoglobin constituents. The two monomer globins of Megascolecidae were thus obtained and sequenced.

A phylogenetic tree was constructed for seven known monomer sequences of annelid hemoglobins, supporting the traditional relationship among these annelids.

MATERIALS AND METHODS

Materials

Live earthworms *Pheretima communissima* were collected in the field of the Nihon University School of Medicine and *Pontodrilus matsushimensis* collected from high tide of Matsushima Bay were kindly donated by Dr. N. Takeuchi, Miyagi College of Education, Sendai, Japan. Acetonitrile, *p*-amidinophenylmethanesulfonyl fluoride-HCl, *tri-n*-butyl phosphine, and 4-vinylpyridine were purchased from Nacalai Tesque (Kyoto, Japan). Methylethylketone, trifluoroacetic acid (TFA), and 1-L-tosylamide-2-phenylethyl chloromethyl ketone treated trypsin from bovine pancreases (type XIII; 10,000-13,000 BAEE units per mg protein) were obtained from Sigma Chemical Co. (St. Louis, MO, USA). Lysyl endopeptidase (*Achromobactor* protease I) was purchased from Wako Pure Chemicals Co. (Tokyo, Japan). A Bradford's protein assay kit was purchased from Nippon Bio-Rad Laboratories KK (Tokyo, Japan). Separation columns, Resource RPC (3 ml packed with Source 15 RPC gel matrix) and PepRPC HR 5/5 placed in a Fast Protein Liquid Chromatography (FPLC), were purchased from Pharmacia Biotech. (Upsala, Sweden).

Preparation of hemoglobin

Samples of hemoglobin were prepared as previously described (Shlom and Vinogradov, 1973; Shishikura, 1996). The hemoglobin was converted to the cyanomet form by adding 2% $K_3Fe(CN)_6$, 0.5% KCN, and 0.1% $NaHCO_3$ (Moss and Ingram, 1968). Heme was removed from the hemoglobin using methylethylketone as described (Teale, 1959).

Purification of monomer globin subunits

Each of the monomer globin subunits of *Pontodrilus matsushimensis* and *Pheretima communissima* was separated and purified by reversed-phase chromatography on Resource RPC 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).

SDS-polyacrylamide gel electrophoresis (PAGE)

SDS-PAGE was conducted on a mini-slab gel apparatus, Mighty Small II (Hoefer Scientific Instruments, San Francisco, CA, USA) according to the instructions of the manufacturer with slight modification of Laemmli's (Laemmli, 1970) procedure. Molecular weight markers were: 66 kDa (bovine serum albumin), 45 kDa (ovalbumin), 29 kDa (carbonic anhydrase) and 18.4 kDa (β-lactoglobulin).

Protein concentration

Bradford's (1976) protein assay was performed according to the micro-assay procedure described in the Bio-Rad technical bulletin attached to the kit, using γ-globulin as the standard. Protein fractions eluted from reversed-phase columns were lyophilized and then redissolved in 50 mM ammonium bicarbonate buffer, pH 8.9. A 10 μl of redissolved sample was measured by a spectrophotometer Uvidec 340Q (Jasco Co., Tokyo, Japan).

Protein modification

Reduction and *S*-pyridylethylation of globins were performed by methods described elsewhere (Friedman *et al.*, 1970; Shishikura, 1996). Finally, the remaining reagents in the reaction mixture were completely removed from the sample by reversed-phase chromatography on a Resource RPC column.

Enzymatic cleavages

Lysyl endopeptidase digestion was performed essentially as described (Jekel *et al.*, 1983), and was detailed in the previous study (Shishikura, 1996). Large peptides derived from the intact globin by the digestion of lysyl endopeptidase were digested further with trypsin at an enzyme/substrate ratio of 1/200 (by mass) at 37°C for 8 hr in 0.1 M Tris-HCl, pH 8.0, containing 2 M urea.

Peptide separation

All peptides derived from monomer globin as well as peptides from the lysyl endopeptidase digestion were separated using a PepRPC HR 5/5 column in a 0.1% 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 nm and 280 nm. Re-chromatographies of selected peptides were performed as previously described (Shishikura *et al.*, 1987).

Amino acid sequence determination

Sequence analysis was performed using a gas phase protein sequencer, PPSQ-10 (Shimadzu Co., Kyoto, Japan), equipped with a PTH-10 amino acid analyzer. PTH-derivatives from the sequencer

were separated and quantitated.

Computer analysis of aligned sequences

Sequences were aligned by a multiple alignment program, Clustal W (Thompson *et al.*, 1994), which can be found on the home page of the Baylor College of Medicine (curator; Dr. Smith RF, Human Genome Center, Baylor College of Medicine). Pairwise distances and a rooted tree with branch lengths among seven monomer sequences were analyzed using the computer program Protdist under the Dayhoff PAM matrix option, Neighbor under the unweighted pair group method with arithmetic (UPGMA) means and Neighbor-Joining option of the PHYLIP package (Felsenstein, 1993). Other computations were performed using the DNASIS programs (NBRF-Protein Sequence Database, version 23, 1995) of Hitachi Software Engineering Co., Ltd. (Yokohama, Japan).

RESULTS

Chain separation

Figure 1A shows a typical chain separation pattern of the whole hemoglobin of *Pontodrilus matsushimensis* on a reversed-phase column (Resource RPC), whose SDS-PAGE patterns are shown in Fig. 2A, B. Similarly, Fig. 1B and Fig. 3A, B show a typical separation pattern of whole hemoglobin of *Pheretima communissima* and SDS-PAGE patterns of fractions eluted from the column, respectively. In both species, the subunit constituents are: monomer (Fig. 2A4, B4 for *Pontodrilus*, Fig. 3A5, B5 for *Pheretima*), trimer (Fig. 2A5, B5 for *Pontodrilus*, Fig. 3A7, B7 for *Pheretima*) and several protein bands including linker chains (Fig. 2A1, 2, 3, B1, 2, 3 for *Pontodrilus*, Fig. 3A1, 2, 3, 4, B1, 2, 3, 4 for *Pheretima*). Both FPLC chromatograms and SDS-PAGE analyses also show two monomer components from both *Pontodrilus matsushimensis* and *Pheretima communissima*. The components in the peaks indicated by arrows are identified to be monomer subunits by SDS-PAGE analyses under reduced and non-reducued conditions. One of the minor components of *Pheretima communissima* monomer globins is shown in Figs. 3A6 and 3B6 but others are not shown here. The recoveries of the first eluted monomer globins (indicated by bars in Fig. 1) were as much as 520 μg (20.8%) for *Pontodrilus matsushimensis* and 480 μg (19.2%) for *Pheretima communissima*, when 2.5 mg of each hemoglobin sample was applied.

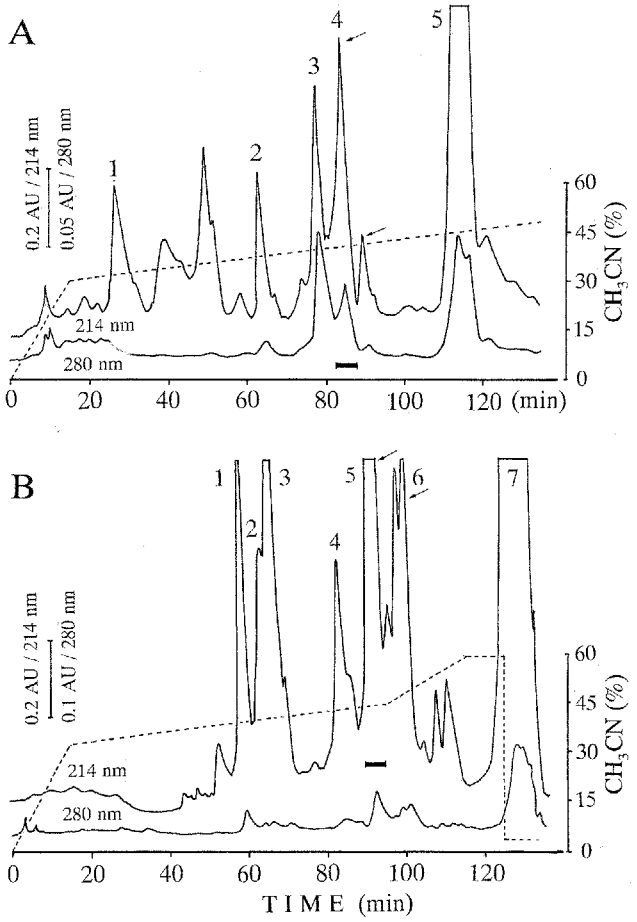

Fig. 1. Elution profiles of the reversed-phase FPLC separation of whole hemoglobins of the two species, *Pontodrilus matsushimensis* (A) and *Pheretima communissima* (B). The column (Resource RPC 3 ml packed with Source 15 RPC) was eluted with 0.1% TFA buffered gradient to 60% acetonitrile in 0.08% TFA at a flow rate of 0.5 ml/min. Arrows indicate monomer subunits and bars are those used for amino acid sequence studies.

Peptide mapping of the two monomeric globins

The major monomer globins from *Pontodrilus matsushimensis* and *Pheretima communissima* were reduced and *S*-pyridylethylated and were then digested, separately, by lysyl endopeptidase. Figures 4A and 4B show their FPLC patterns on a PepRPC HR 5/5 column. These lysyl endopeptidase digested (K-) peptides of *Pontodrilus* (Fig. 4A) and *Pheretima* (Fig. 4B) were designated by considering the order of elution, position and homology in the sequences of the two species; for example, *Pontodrilus matsushimensis* (Pm) K-6 peptide and *Pheretima communissima* (Pc) K-6 peptide indicate the last eluted peaks sharing sequence homology and the peptide positions of the two species. Definitive differences between *Pontodrilus matsushimensis* and *Pheretima communissima* on the chromatograms were observed: One peak with a retention time of 60.9 min (PmK-5 peptide) in *Pontodrilus* disappeared in *Pheretima*. In contrast, two different peaks with retention times of 18.8 min (PcK-5a peptide) and 36.2 min (PcK-5b peptide) emerged in *Pheretima*. The elution order of K-2 and K-3 peptides in *Pontodrilus* appeared in reverse order in *Pheretima*.

Sequencing and alignment

The amino acid sequences of monomer globin chains of *Pontodrilus matsushimensis* and *Pheretima communissima* were determined as follows; their reduced and *S*-pyridylethylated heme deprived proteins were digested with lysyl endopeptidase and separated by reversed-phase chromatography whose elution pattern was depicted above (Fig. 4). The lysyl endopeptidase digested peptides were sequenced by a Shimadzu amino acid sequencer PPSQ-10. All the K-peptides except for PmK-6 and PcK-6 were

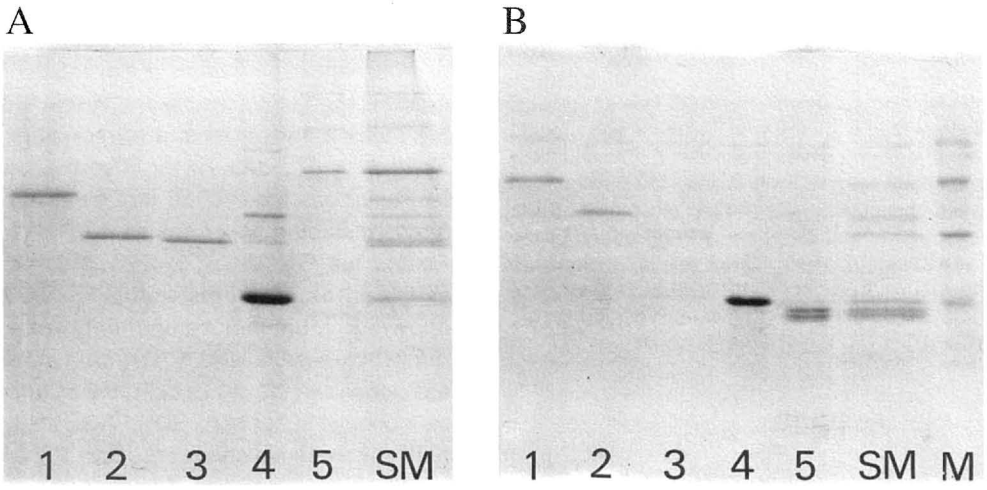

Fig. 2. SDS-PAGE of fractions recovered from FPLC of *Pontodrilus matsushimensis* hemoglobin (Fig. 1A). Non reducing gradient gel (A) and reducing gradient gel (B) between 10 and 20% (top to bottom) were used. The numbers of lanes in both panels A and B correspond to those of peaks in Fig. 1A. SM, starting material; M, molecular mass markers, 66, 45, 29, and 18.4 kDa.

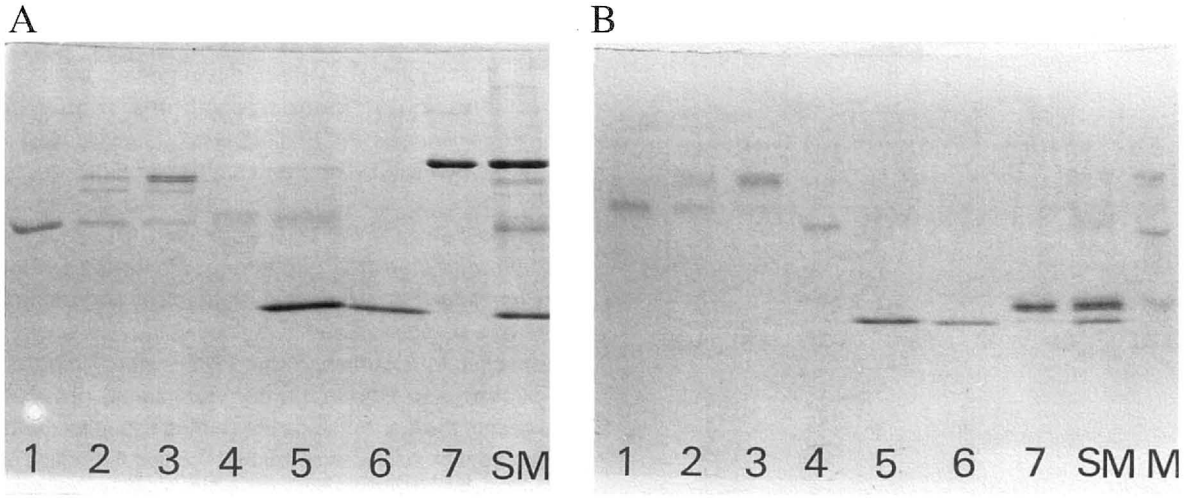

Fig. 3. SDS-PAGE of fractions recovered from FPLC of *Pheretima communissima* hemoglobin (Fig. 1B). Non reducing gradient gel (A) and reducing gradient gel (B) between 10 and 20% (top to bottom) were used as described. The numbers of lanes in both panels A and B correspond to those of peaks in Fig. 1B. SM, starting material; M, molecular mass markers, 66, 45, 29, and 18.4 kDa.

sequenced completely. The rests were digested further by trypsin and separated by reversed-phase chromatography on a Pep RPC column (data not shown), followed by sequencing. Sequence similarities to the monomer globin of *Pheretima hilgendorfi* previously determined (Shishikura, 1996) as well as to those of annelid hemoglobins were used to determine the position of sequenced peptides. Only one small peptide derived from the monomer globin of *Pheretima communissima* could not be found in the elution of lysyl endopeptidase digestion peptides, however, its amino acid sequences, Val^{10}-Lys^{11}, were determined by the analysis of the first 20 amino-terminal amino acid sequences. The peptide fragment containing the third cysteine residue in *Pontodrilus* monomer subunit was analyzed in triplicate. The results are shown in Fig. 5 with the five known sequenced monomers of annelids, including *Tylorrhynchus heterochaetus* (Suzuki *et al.*, 1982), *Lumbricus terrestris* (Shishikura *et al.*, 1987), *Pheretima sieboldi* (Suzuki, 1989), *Tubifex tubifex* (Stern *et al.*, 1990) and *Pheretima hilgendorfi* (Shishikura, 1996). The *Pontodrilus matsushimensis* monomer subunit constitutes of 141 amino acid residues, containing three cysteine residues, and has a calculated molecular mass of 16,366 Da. The *Pheretima communissima* monomer subunit has 140 amino acid residues, containing two cysteine residues, and has a calculated molecular mass of 16,000 Da. There are 82 positions of identity among these four monomeric globins of the family Megascolecidae. There are also 26 invariant positions in this grouping of seven globins of annelids. When

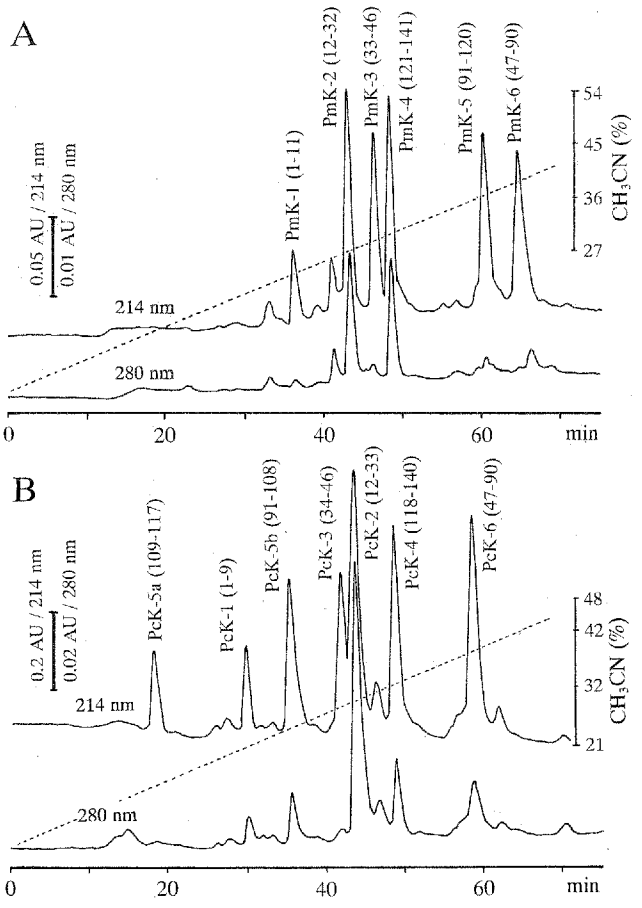

Fig. 4. Elution profiles of the reversed-phase FPLC separation of the lysylendopeptidase digested peptides of the monomer subunits of *Pontodrilus matsushimensis* hemoglobin (A) and *Pheretima communissima* hemoglobin (B). The column (PepRPC HR 5/5) was eluted with 0.1% TFA buffered gradient to 60% acetonitrile in 0.08% TFA at a flow rate of 0.3 ml/min. The numbers in parentheses are those of amino acid residues in the sequences of monomer subunits of *Pontodrilus matsushimensis* and *Pheretima communissima* hemoglobins.

these globins are compared with the *Pheretima communissima*, there are 123 identities (87.9%) in *Pheretima hilgendorfi*, 121 identities (85.8%) in *Pheretima sieboldi*, 103 identities in *Pontodrilus* (66.5%), 82 identities (56.2%) in *Lumbricus*, 78 identities (51.3%) in *Tubifex* and 70 identities (45.8%) in *Tylorrhynchus*. A matrix of amino acid identities among the rest was also developed (Table 1). Figure 6 shows a phylogenetic tree of seven species of annelids based on the amino acid sequence analyses of their monomeric globins.

DISCUSSIONS

Many attempts have been carried out to separate and purify individual chains of annelid hemoglobins: gel filtration (Shlom and Vinogradov, 1973; Suzuki *et al.*, 1983b), DEAE-ion exchange column chromatography (Suzuki *et al.*, 1983a), preparative gel electrophoresis (Ochiai and Enoki, 1979), chromatofocusing (Gotoh, 1982; Shishikura *et al.*, 1986), reversed-phase column chromatography (Fushitani *et al.*, 1988; Suzuki *et al.*, 1990a; Ownby *et al.*, 1993) and others (Bonaventura and Bonaventura, 1981). Our separation method described here had a similar efficiency as previously obtained in separation of *Lumbricus* hemoglobin by Ownby *et al.* (1993) and *Lamellibrachia* hemoglobin by Suzuki *et al.* (1990b), who used reversed-phase columns, Synchropak RP-P C_{18} and Cosmosil $5C_{18}$-300, respectively. However, several advantages still exist: Resource RPC has a large quantity of sample capacity with high chemical stability, gives excellent resolution in separation at high flow rates with low back pressure (*versus* flow rates), and can be run in pH ranging from 1-12. In this study, amounts as large as 520 μg (20.8%) and 480 μg (19.2%) for monomer subunits and 1.16 mg (46.4%) and 1.56 mg (62.4%) for trimer subunits of *Pontodrilus matsushimensis* and *Pheretima communissima*, respectively, were obtained from 2.5 mg of each whole hemoglobin sample.

According to the SDS-PAGE analyses under reduced and unreduced conditions, the *Pontodrilus* monomer subunit which

```
P.mat   DCDVLARFRVKHQWQEVFL-GSNRMEFSHDLWKEFFHDHSDLVALFKRVHGENINSPQFQAHGIRVLAGLDGLCALDEEDT  80
P.com   DCDVLERFKVKHQWQAVFS-EAHRTDFSLHFWKEFLHDHPDLVGLFKRVNGDNIYSPEFQAHGIRVLAGLDSVIGVLDEED  80
P.hil   ECDVLERFKVKHQWQTVFS-EAHRTEFSLHFWKEFLHDHPDLVELFTRVNGANIYSPEFQAHGIRVLAGLDSVIGVLDEIP  80
P.sie   DCNTLKRFKVKHQWQQVFSGEHHRTEFSLHFWKEFLHDHPDLVSLFKRVQGENIYSPEFQAHGIRVLAGLDSVIGVLDEDD  81
L.ter   ECLVTEGLKVKLQWASAFGHAHQRVAFGLELWKGILREHPEIKAPFSRVRGDNIYSPQFGAHSQRVLSGLDITISMLDTPD  81
T.tub   ECDALQRFKVKHQWAEAFGTSHHRLDFGLKLWNSIFRDAPEIRGLFKRVDGDNAYSAEFEAHAERVLGGLDMTISLLDDQA  81
T.het   TDCGILQRILVLQQWAQVYSVGESRTDFAIDVFNNFFRTNPD-RSLFNRVNGDNVYSPEFKAHMVRVFAGFDILISVLDDKP  81
          *       *  **          *  *                 * ** * *  *  * **  **  * *

P.mat   TLNQLLVHLKGQHTERGTKPEYFDIFGVHLLRILDDHLGKTYFARQEWQDCYAV-IAAGIKP-  141
P.com   TFNVQLAHLKAQHTERGTKPEYFDLFGKHLADHLGDKLG-THFDFGAFRDCYAV-IAAGIKP-  140
P.hil   TLTVQLAHLKAQHTERGTKPEYFDLFGKHLASHLGDELG-THFDYAAFRDCYDF-IASGIKP-  140
P.sie   TFTVQLAHLKAQHTERGTKPEYFDLFGTQLFDILGDKLG-THFDQAAWRDCYAV-IAAGIKP-  140
L.ter   MLAAQLAHLKVQHVERNLKPEFFDIFLKHLLHVLGDRLG-THFDFGAWHDCVDQ-IIDGIKDI  141
T.tub   AFDAQLAHLKSQHAERNIKADYYGVFVNELLAVLPDYLG-TKLDFKAWSECLGV-ITGAIHD-  140
T.het   VLDQALAHYAAFHLQFGTIP--FKAFGQTMFQTIAEHIH--GADIGAWRACYAEQIVTGITA-  139
              * *   *          *                    *   *   *
```

Fig. 5. Alignment of the amino acid sequences of monomer globins of *Pontodrilus matsushimensis* and *Pheretima communissima*, with five of other members of oligochaetes, including *Pheretima hilgendorfi* (*P. hil*) (Shishikura, 1996), *Pheretima sieboldi* (*P. sie*) (Suzuki, 1989), *Lumbricus terrestris* (*L. ter*) (Shishikura *et al.*, 1987), *Tubifex tubifex* (*T. tub*) (Stern *et al.*, 1990) and *Tylorrhynchus heterochaetus* (*T. het*) (Suzuki *et al.*, 1982). The alignment was achieved by a multiple alignment program, Clustal W (Thompson *et al.*, 1994). The 26 residues conserved in the seven globins are shown by asterisks.

Table 1. Pair wise distances between the complete amino acid sequences of monomer subunits of *Pontodrilus*, three species of *Pheretima*, *Lumbricus*, *Tubifex* and *Tylorrhynchus* hemoglobins

	P. mat	*P. com*	*P. hil*	*P. sie*	*L. ter*	*T. tub*	*T. het*
P. mat	–	66.5/103/155	63.2/ 96/152	66.9/ 99/148	45.1/74/164	44.6/70/157	37.8/62/164
P. com	0.4458	–	87.9/123/140	85.8/121/141	56.2/82/146	51.3/78/152	45.8/70/153
P. hil	0.5616	0.1533	–	81.6/115/141	55.1/81/147	46.7/71/152	43.1/66/153
P. sie	0.4458	0.1631	0.2451	–	51.7/77/149	48.7/73/150	46.1/71/154
L. ter	0.9709	0.7707	0.8240	0.8442	–	53.3/82/154	36.5/61/167
T. tub	1.0221	0.8106	0.9568	0.8715	0.8037	–	39.8/64/161
T. het	1.1952	1.0127	1.1286	1.0698	1.3905	1.2534	–

Values in the upper right half of the table are the matching percentages (on the left), numbers of identical residues (on the center) and total windows of amino acid residues (on the right), which are calculated from complete sequence data by an amino acid maximum homology program placed on DNASIS. Values in the lower left half of the table are distances computed from amino acid sequences corresponding to 145 sites of monomer globins (see Fig. 5) by the program Protdist under the Dayhoff PAM matrix option (Felsenstein, 1993). Abbreviations of species correspond to those used in Fig. 5.

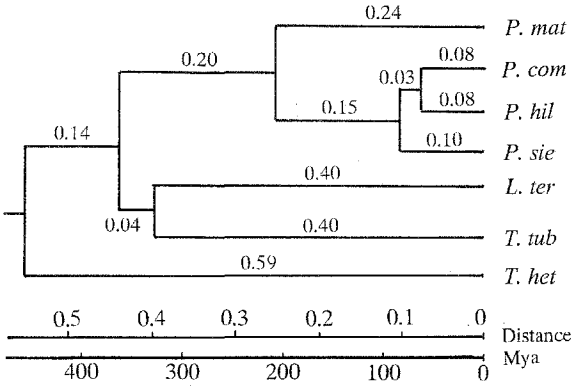

Fig. 6. A rooted tree produced from distance matrices shown in the lower left half of Table 1 by the program Neighbor under the UPGMA option of the PHYLIP package (Felsenstein, 1993). Branch lengths are proportional to protein distances at the bottom of the tree. Also shown on the tree are individual branch lengths from the neighbor-joining analysis (Felsenstein, 1993) that produced the same branching structure as the UPGMA analysis. Mya, million years ago. Abbreviations of species correspond to those used in Fig. 5.

migrated slower than any of the constituent polypeptide chains of the trimer subunit (Fig. 2B, lane 5) was unique feature because corresponding subunits of the other annelid hemoglobins, in general, migrated faster than the three chains of trimer subunit (Vinogradov, 1985). Furthermore, an apparent molecular weight of the *Pontodrilus* monomer subunit was estimated to be about 18,400 Da on SDS-PAGE. After sequenced, however, the molecular mass of the monomer subunit was determined to be 16,366 Da. This value was in good agreement with those reported on monomer chains of Megascolecidae (Suzuki, 1989; Shishikura, 1996), including that of *Pheretima communissima*. The value obtained by SDS-PAGE analysis is, therefore, misunderstanding and probably due to dissimilarity in molecular shape and apparent net charge of SDS-treated *Pontodrilus* monomer subunit. Figures 1, 2 and 3 also indicated that heterogeneities of monomer globin subunits existed in both *Pontodrilus matsushimensis* and *Pheretima hilgendorfi* hemoglobin as demonstrated previously in *Lumbricus terrestris* hemoglobin (Shishikura *et al.*, 1987; Ownby *et al.*, 1993) as well as *Neanthes diversicolor* hemoglobin (Suzuki *et al.*, 1994). In this experiment, only the major fractions of both hemoglobin monomers were purified and sequenced.

One of the most striking features of sequenced monomers is that the *Pontodrilus* monomer globin has been shown to have three cysteine residues at positions of 2, 73 and 131. In annelids, the globin chains belonging to the member of strain A have a conserved characteristic containing two cysteine residues located in the amino- and carboxyl-termini. The two cysteine residues in both species apparently form a disulfide-bonded interchain as already reported in the known sequences of monomer subunits of annelid hemoglobins. Therefore, the third cysteine residue at position 73 may be free and is a conspicuous characteristic of monomer globin of *Pontodrilus matsushimensis*. This is the first observation of the presence of the third cysteine residue in the monomer subunits of annelid hemoglobins. Recently, Suzuki *et al.* (1990b) demonstrated the presence of three cysteine residues in the constituent polypeptide chain (AIII) of deep-sea tube-worm *Lamellibrachia* (phylum Vestimentifera) hemoglobin. They showed that the third cysteine residue at position 74 in *Lamellibrachia* hemoglobin is responsible for sulphide-binding ability because the hemoglobin may have been acquired high-molecular adaptation to transport enough sulphide to internal sulphide-oxidizing bacterial symbionts (Childress *et al.*, 1987). It is, however, difficult to directly adjust their discussion to explain the function of the third cysteine residue in *Pontodrilus* hemoglobin because *Pontodrilus matsushimensis* is only found in sea shore environments where the animals live and spend their time 50 cm underground in winter (Dr. N. Takeuchi, personal communication). In this respect of their adaptations to hypoxic conditions, it might be said that they have developed particular mechanisms for maintaining an adequate O_2 supply to tissues. Although there is a lack of data on globin structures, in particular on aquatic or marine oligochaetes with tolerance to a wide range of salinity, it is noteworthy that the *Pontodrilus* hemoglobin has a distinct similarity in the amino acid sequences compared with the deep-sea tube worm

Lamellibrachia. Hence, it would be interesting to investigate the biochemical meanings of the third cysteine residue in the *Pontodrilus* hemoglobin for the molecular evolution of hemoglobins as well as physiology of invertebrate hemoglobins.

Previous works (Gotoh *et al.*, 1987; Shishikura, 1996) suggested that homologous subunits sharing an orthologous relationship in the hemoglobin molecules are useful clues for analyzing the genetic relationship of earthworms. Table 1 lists the percentage differences and protein distances between complete amino acid sequences of *Pheretima* and *Pontodrilus* compared with the known sequences of oligochaetes and polychaetes. The identical scores (81.6%-87.9%) among three species of *Pheretima* showed very high degrees of similarity and might provide evidence biochemical support for the previous morphological study of Easton (1984) who grouped several species of *Pheretima*, including above three species into *Amynthas hilgendorfi* species-complex. Recently, Boore and Brown (1995) sequenced the complete mitochondrial DNA of earthworms (from *Lumbricus terrestris*). It is also hoped that it may be used for analyzing genetic relationships among these species as well as annelids.

The UPGMA analysis (Felsenstein, 1993) of distance matrices shown in the lower left half of Table 1 produced a rooted tree with the assumption of an evolutionary clock (Fig. 6). The same tree topology was obtained by a neighbor-joining analysis (Felsenstein, 1993). This rooted tree was coincident to the traditional phylogeny of annelids. Furthermore, the evolutionary rates of monomer subunits of annelid hemoglobins were also calculated by assuming that the divergence between the Oligochaeta and the Polychaeta occurred about 450 million years ago (Fushitani *et al.*, 1988). This gave an estimated divergence time of *Pheretima communissima* and *Pheretima hilgendorfi* of 62.5 million years ago, that of *Pheretima sieboldi* and the other two species of *Pheretima* 83.3 million years ago, and that of *Pontodrilus* and three species of *Pheretima* 208.8 million years ago. Although very close species of *Pheretima* were analyzed, it was clearly shown that evolutionary divergence patterns can be studied using the amino acid sequence data from the monomer subunits of annelid hemoglobins.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. N. Takeuchi of the Biology Department, Miyagi College of Education, Sendai, Japan, for his cooperation in collecting the *Pontodrilus matsushimensis* and Dr. M. Esumi and the staff of the First Department of Pathology, Nihon University School of Medicine, for their generosity in allowing the use of DNASIS and helping set up the equipment for analyzing amino acid compositions and amino acid sequences. We also express our gratitude to the staff of the Department of Biology, Nihon University School of Medicine for their cooperation.

REFERENCES

Bonaventura J, Bonaventura C (1981) Preparation of high molecular weight invertebrate hemoglobins. Meth Enz 76: 43–54

Boore JL, Brown WM (1995) Complete sequence of the mitochondrial DNA of the annelid worm *Lumbricus terrestris*. Genetics 141: 305–319

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248–254

Childress JJ, Felbeck H, Somero GN (1987) Symbiosis in the deep sea. Science 203: 107–112

Clementi ME, Condò SG, Castagnola M, Giardina B (1994) Hemoglobin function under extreme life conditions. Eur J Biochem 223: 309–317

Easton EG (1984) Earthworms (Oligochaeta) from islands of the south-western Pacific, and a note on two species from Papua New Guinea. New Zealand J Zool 11: 111–128

Felsenstein J (1993) PHYLIP: Phylogeny Inference Package, Version 3.5c. Department of Genetics SK-50, University of Washington, Seattle, WA 98195, U.S.A.

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: 3868–3871

Fushitani K, Matsuura MSA, Riggs AF (1988) The amino acids 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

Goodman M, Pedwaydon J, Czelusniak J, Suzuki T, Gotoh T, Moens L, Shishikura F, Walz D, Vinogradov S (1988) An evolutionary tree for invertebrate globin sequences. J Mol Evol 27: 236–249

Gotoh T (1982) Studies on erythrocruorin of *Tylorrhynchus heterochaetus*. V. Dissociation of 54,000 molecular weight subunit. J Sci Ccollege General Education, University of Tokushima XV: 1–10

Gotoh T, Shishikura F, Snow JW, Ereifej KI, Vinogradov SN, Walz DA (1987) Two globin strains in the giant annelid extracellular haemoglobins. Biochem J 241: 441–445

Iizuka A (1897) On a new species of littoral Oligochaeta (*Pontodrilus matsushimensis*). Ann Zool Jap 2: 21–27

Jamieson BGM (1971) A review of the megascolecoid earthworm genera (Oligochaeta) of Australia. Part III-The subfamily Megascolecinae. Mem Queensland Mus 16: 69–102

Jamieson BGM (1988) On the phylogeny and higher classifcation of the Oligochaeta. Cladistics 4: 367–410

Jekel PA, Weijer WJ, Beintema JJ (1983) Use of endoproteinase Lys-C from *Lysobacter enzymogenes* in protein sequence analysis. Anal Biochem 134: 347–354

Kapp OH, Moens L, Vanfleteren J, Trotman CNA, Suzuki T, Vinogradov SN (1995) Alignment of 700 globin sequences: Extent of amino acid substitution and its correlation with variation in volume. Protein Sci 4: 2179–2190

Kobayashi S (1941) An outlook of distribution of terrestrial oligochaetes in the western regions of Japan (Nishinihon ni okeru rikusei hinnmourui no bunnpugaikyou; in Japanese). Doubutsugaku Zasshi 53: 371–384

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685

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

Ochiai T, Enoki Y (1979) Earthworm (*Pheretima communissima* and *Pheretima hilgendorfi*) hemoglobins: Giant assemblies and subunits. Biochim Biophys Acta 579: 442–451

Ownby DW, Zhu H, Schneider K, Beavis RC, Chait BT, Riggs AF (1993) The extracellular hemoglobin of the earthworm, *Lumbricus terrestris*. Determination of subunit stoichiometry. J Biol Chem 268: 13539–13547

Perutz MF (1984) Species adaptation in a protein molecule. Adv

Protein Chem 36: 213–244
Shishikura F (1996) Amino acid sequence of the monomer subunit of the extracellular hemoglobin of the earthworm, *Pheretima hilgendorfi*. Zool Sci 13: 551–558
Shishikura F, Mainwaring MG, Yurewicz EC, Lightbody JJ, Walz DA, Vinogradov SN (1986) A disulfide-bonded trimer of myoglobin-like chains is the principal subunit of the extracellular hemoglobin of *Lumbricus terrestris*. Biochim Biophys Acta 869: 314–321
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
Shlom JM, Vinogradov SN (1973) A study of the subunit structure of the extracellular hemoglobin of *Lumbricus terrestris*. J Biol Chem 248: 7904–7912
Stern MS, Vinogradov SN, Sharma PK, Ereifej K, Walz DA (1990) Amino acid sequence of the monomer subunit of the giant extracellular hemoglobin of the aquatic oligochaete, *Tubifex tubifex*. Eur J Biochem 194: 67–73
Suzuki T (1989) Amino acid sequence of the monomer subunit of the giant multisubunit hemolgobin from the earthworm *Pheretima sieboldi*. Eur J Biochem 185: 127–134
Suzuki T, Takagi T, Gotoh T (1982) Amino acid sequence of the smallest polypeptide chain containing heme of extracellular hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. Biochim Biophys Acta 708: 253–258
Suzuki T, Koboyashi M, Furukohri T, Gotoh T (1983a) Autoxidation of extracellular oxyhaemoglobin from the polychaete *Perinereis brevicirris* (Grube). Comp Biochem Physiol 75B: 17–21
Suzuki T, Takagi T, Furukohri T, Gotoh T (1983b) Separation of constituent polypeptide chains containing heme from extracellular hemoglobin of the polychaete *Perinereis brevicirris* (Grube). Comp Biochem Physiol 75B: 567–570
Suzuki T, Takagi T, Gotoh T (1990a) Primary structure of two linker chains of the extracellular hemoglobin from the polychaete *Tylorrhynchus heterochaetus*. J Biol Chem 265: 12168–12177
Suzuki T, Takagi T, Ohta S (1990b) Primary structure of a constituent polypeptide chain (AIII) of the giant haemoglobin from the deep-sea tube worm *Lamellibrachia*. A possible H_2S-binding site. Biochem J 266: 221–225
Suzuki T, Ohta T, Yuasa HJ, Takagi T (1994) The giant extracellular hemoglobin from the polychaete *Neanthes diversicolor*. The cDNA-derived amino acid sequence of linker chain L2 and the exon/intron boundary conserved in linker genes. Biochim Biophys Acta 1217: 291–296
Takeuchi N (1980a) Neuroendocrine control of hydration in megascolecid earthworms. Comp Biochem Physiol 67A: 341–345
Takeuchi N (1980b) Control of coelomic fluid concentration and brain neurosecretion in the littoral earthworm *Pontodrilus matsushimensis* Iizuka. Comp Biochem Physiol 67A: 357–359
Teale FWJ (1959) Cleavage of the haem-protein link by acid methylethylketone. Biochim Biophys Acta 35: 543
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 (1985) The structure of invertebrate extracellular hemoglobins (erythrocruorins and chlorocruorins). Comp Biochem Physiol 82B: 1–15
Vinogradov SN, Walz DA, Pohajdak B, Moens K, Kapp OH, Suzuki T, Trotman CNA (1993) Adventitious variability? The amino acid sequences of nonvertebrate globins. Comp Biochem Physiol 106B: 1–26

(Received July 19, 1996 / Accepted August 29, 1996)