Primary Structure of a Linker Subunit of the Tube Worm 3000-kDa Hemoglobin*

(Received for publication, September 11, 1989)

Tomohiko Suzuki†‡, Takashi Takagi§, and Suguru Ohta¶

From the †Department of Biological Research, Faculty of Science, Kochi University, Kochi 780, Japan, the ‡Biological Institute, Faculty of Science, Tohoku University, Sendai 980 Japan, and the ¶Ocean Research Institute, University of Tokyo, Tokyo 164, Japan

The deep-sea tube worm Lamellibrachia contains two giant extracellular hemoglobins, a 3000-kDa hemoglobin and a 440-kDa hemoglobin. The former consists of four heme-containing chains (AI–AIV) and two linker chains (AV and AVI) for the assembly of the heme-containing chains. The 440-kDa hemoglobin consists of only four heme-containing chains (Suzuki, T., Takagi, T., and Ohta, S. (1988) Biochem. J. 255, 541–548). The complete amino acid sequence of a linker subunit (chain AV) has been determined by automated Edman sequencing of the peptides derived by digestion with lysyl endopeptidase and carboxylic acid. The chain is composed of 224 amino acid residues, and the molecular mass for the protein moiety was calculated to be 24,894 Da. An Asx-X-Thr sequence which is possible as a glycosylation site was suggested at positions 108–110. A computer-assisted homology search showed that the sequence shows no notable homology with any other globins and proteins. However, a careful alignment of the linker sequence with a heme-containing chain sequence suggested that there is a slight, but significant homology between the two sequences. The alignment also suggested that the linker resulted from gene duplication of a heme-containing chain with a three exon-two intron structure, and that the first exon of domain 1 and the last exon of domain 2 had been lost during evolution. In our alignment, domain 1 has the heme-binding proximal histidine, but domain 2 does not. This is the first linker subunit to be sequenced completely.

One of the most remarkable differences of annelid 3000–4000-kDa hemoglobins from other invertebrate and vertebrate hemoglobins is characterized by their low heme (iron) content. This reason was derived from the presence of non-heme 31–37-kDa chains that act as ‘linkers’ for the assembly of heme-containing chains in construction of a giant molecular architecture (Vinogradov, 1985; Vinogradov et al., 1986).

Although the deep-sea tube worm Lamellibrachia has been assigned to the phylum of Vestimentifera, it contains two annelid-like giant extracellular hemoglobins, a 3000-kDa hemoglobin and a 440-kDa hemoglobin. Both hemoglobins comprise heme-containing chains with a molecular mass of about 16 kDa, two of which are common to both hemoglobins; furthermore, the 3000-kDa hemoglobin has two additional 32–36-kDa subunits which are proposed to be the linkers (Suzuki et al., 1988, 1989). In this paper, we report the first complete amino acid sequence of a linker subunit from the Lamellibrachia 3000-kDa hemoglobin.

MATERIALS AND METHODS

The tube worm Lamellibrachia sp. (undescribed) was collected from the cold-seep area located off Sagami Bay at a depth of 1160 m, Japan, by a Japanese submersible SHINKAI 2000. The 3000-kDa hemoglobin was separated on a gel filtration column (3 × 100 cm) of Sepharose CL-4B, equilibrated and eluted with 0.1 M phosphate buffer (pH 7.2), and concentrated by ultra centrifugation (50,000 rpm for 1 h, Hitachi 70P). After removal of heme, the sample was reduced with 10 mM dithiothreitol, and the constituent chains were separated on reverse-phase column (Cosmosil 5C8-300, 4.6 × 150 mm, Nacalai Tesque) with a linear gradient of 30–60% acetonitrile in 0.1% trifluoroacetic acid (Suzuki et al., 1990).

The chain AV-1 was carboxymethylated as described previously (Suzuki and Gotoh, 1986). The chain (7 nmol) was digested with lysyl endopeptidase (Wako) in 50 mM Tris-HCl buffer (pH 8.8) at an enzyme-to-substrate ratio of 1:100 at 37 °C for 3 h (Fig. 4 in the Miniprint).1 To obtain overlap peptides, the chain (9 nmol) was digested with endoproteinase Asp-N (Boehringer) in 50 mM phosphate buffer (pH 8) at an enzyme-to-substrate ratio of 1:100 at 37 °C for 13.5 h (Fig. 5 in the Miniprint).

The peptides were purified on reverse-phase column (Cosmosil 5C8-300, 4.6 × 150 mm, Nacalai Tesque) with a linear gradient of 0–80% acetonitrile in 0.1% trifluoroacetic acid over 120 min at a flow rate of 1 ml/min. Some peptides were further purified by rechromatography. Amino acid compositions of the peptides are shown in Table 1 (Miniprint). Amino acid sequences of intact protein and peptide were determined by use of an automated sequencer (Applied BioSystems 477A Protein Sequencer) (Table 2 in the Miniprint).

The peptides were numbered from the N terminus. In this paper, L indicates a lysyl endopeptidase peptide, and A, an endoproteinase Asp-N peptide.

The peptide L6 was applied to an immobilized concanavalin A column (4.6 × 150 mm, Hohein Oil Co., Ltd.) equilibrated with 20 mM Tris-HCl buffer (pH 8) containing 150 mM NaCl and 1 mM CaCl2 and eluted with the same buffer containing 150 mM glucose at a flow rate of 0.5 ml/min. The eluate was monitored for absorbance at 280 nm, and the elution profile was obtained by subtracting its baseline.

RESULTS

Lamellibrachia 3000-kDa hemoglobin has two types of linker subunits, chains AV and AVI at a ratio of about 5:1. The major linker chain AV can be separated further into two isomers AV-1 and AV-2 by reverse-phase chromatography, but their amino acid sequences differ only at one position in the N-terminal 40 residues examined (Suzuki et al., 1988). In the present study, we confirmed that peptide mapping (lysyl

Portions of this paper (including Tables 1 and 2 and Figs. 4 and 5) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.

* This work was supported in part by a grant-in-aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan (to T. S.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed.

‡ To whom correspondence should be addressed.

§ To whom correspondence should be addressed.

¶ To whom correspondence should be addressed.

1 Portions of this paper (including Tables 1 and 2 and Figs. 4 and 5) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are included in the microfilm edition of the Journal that is available from Waverly Press.
endopeptidase digestion) for both chains gave an almost identical result and that all the amino acid compositions of peptides corresponding to each other were indistinguishable (data not shown). Thus the amino acid sequences of chains AV-1 for sequence determination.

endopeptidase digestion) for both chains gave an almost identical result and that all the amino acid compositions of peptides corresponding to each other were indistinguishable (data not shown). Thus the amino acid sequences of chains AV-1 for sequence determination.

The amino acid sequence of a linker subunit (chain AV-1) of the 3000-kDa hemoglobin from the deep-sea tube worm Lamellibrachia was determined by automated Edman degradation of intact protein and peptides derived by cleavage with lysyl endopeptidase and endoproteinase Asp-N. The strategies used to establish the complete sequence are summarized in Fig. 1. The alignment of the peptides is supported by at least 3 overlapped residues. Peptide L5 was eluted in two different positions on a reverse phase column (Fig. 4 in the Miniprint). However no difference was found in amino acid compositions and N-terminal sequences of 34 residues for the two L5 peptides. To confirm the C-terminal sequence, the chain AV was cleaved with CNBr. Consequently, the homoserine-free peptide corresponding to the positions 212-224 was obtained, indicating the accuracy of the C-terminal sequence proposed (data not shown). The residue at position 108 was not detected by Edman degradation of the peptides L5 and A4, but the amino acid composition of peptide A4 clearly suggested that the residue is Asx. Moreover, since the peptide L5 bound to an immobilized concanavalin A column (data not shown), we assumed that it is asparagine attached with carbohydrate group in consideration of the threonine (position 110) present 2 residues after. An Asn-X-Thr sequence is well known as a glycosylation site. As shown in Fig. 1, Lamellibrachia linker chain AV-1 is composed of 224 amino acid residues and has a calculated molecular mass of 24,894 Da for the protein moiety. There is a difference in the molecular masses from sequencing (25 kDa) and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (32 kDa) reported previously (Suzuki et al., 1988), but the carbohydrate groups might contribute to the difference to some extent, although we could not determine them because of a very small amount of the sample available.

**Discussion**

The invertebrate 3000-4000-kDa hemoglobins found in annelids and tube worms have a characteristic linker subunit for the assembly of heme-containing chains (Vinogradov et al., 1986; Suzuki et al., 1988). The molecular mass of linker chains is about the double size of usual heme-containing chain, showing a value of 30-37 kDa on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Vinogradov, 1985). In addition, at least a major part of the linker chains was suggested to exist as a non-heme protein from heme and protein analyses. Although the stoichiometry of the linker chain to heme-containing chains is not well established, it is likely that one or two linker subunits are present per one-twelfth subunit of intact molecule (Vinogradov et al., 1986; Fushitani and Riggs, 1998). One of the biological interests for linker chain is its evolutionary origin, that is whether the linker is derived from a protein quite different from hemoglobin or results from gene duplication of heme-containing chain. This work is aimed at answering this question.

We have determined the amino acid sequence of 224 residues of Lamellibrachia chain AV-1. This is the first linker subunit of annelid-like giant multisubunit hemoglobin to be sequenced completely. The N-terminal amino acid sequence of Lamellibrachia linker chain shows a low, but significant homology (23% identity) with the partial sequence (28 residues) of Lumbricus linker chain (Fushitani and Riggs, 1988) as follows:

Lamellibrachia AAAQPLSVSDAMGARVDQAWVRDQFLTKQF

Lumbricus AS DPYQERRFQYL VKNQNIILDYLKAKL

A computer-assisted homology search (the GENAS, Kyushu University) for the complete amino acid sequence of Lamellibrachia linker chain AV was carried out, but no protein with notable homology was found. However, in consideration of the fact that the sequence homology between the homoglobin chains is relatively low, then we have tried to align carefully the sequence of linker chain with those of related heme-containing chains, Lamellibrachia chain III (Suzuki et al., 1990) and Lumbricus chain c (Jhiang et al., 1988). First, taking notice of the two structurally important cysteine residues (positions 8 and 146 in Fig. 2) characteristic of all heme-containing chains of annelid and Lamellibrachia hemoglobins, we aligned them roughly by eye. Then the alignment was improved by a computer program based on the algorithm of Feng et al. (1985). As shown in Fig. 2, consequently, it was found that there is a slight, but significant similarity between the three chains. This alignment also suggests that the linker chain has a two-domain structure (domain 1, positions 1-116; domain 2, positions 117-224) that the Lam. L-D1, Lam. L-D2, and Lam. L-D3 lack the N-terminal 40 residues of domain 1 and the C-terminal 40 residues of domain 2. The sequences of domains 1 and 2 of
the linker show 14 and 12% homologies, respectively, with that of Lamellibrachia chain AIII. Here if we consider that the globin gene, including that of a heme-containing chain (chain c) of giant hemoglobin from the earthworm Lumbricus terrestris (Jhiang et al., 1988), has a two intron-three exon structure, it would be found that such large deletions found in N- and C-terminal regions correspond just to the deletions of first exon of domain 1 and of last exon of domain 2, respectively. If so, this also supports the idea that the linker chain might result from gene duplication of a heme-containing chain. Consistent with this, Lightbody et al. (1988) found from a monoclonal antibody study that the linker chain of Lumbricus giant hemoglobin is related to its heme-containing chain I.

The hydropathy profile (Kyte and Doolittle, 1982) of Lamellibrachia linker chain AV-1 was compared with that of heme-containing chain AIII (Suzuki et al., 1990). As shown in Fig. 3, the profile of domain 1 of the linker slightly resembled corresponding profile of heme-containing chain, but that of domain 2 was rather different, suggesting that the domain 1 might take a globin folding even though incomplete. Since the heme-binding proximal histidine is conserved in domain 1 in our alignment (Fig. 2), it might be possible that the domain 1 carries a heme group. The functionally important distal histidine is replaced by glutamine in the domains of linker chain (Fig. 2), as in the case of several invertebrate globins. In contrast, domain 2, lacking the proximal histidine, could never carry a heme.

The steric location of linker chains in the gross quaternary structure is unknown. Vinogradov et al. (1986) assumed that the linker chains form a closed circular collar (“bracelet” model) with 12 complexes of 16 heme-containing chains each, or it is also possible that they act for the assembly of heme-containing chains without a continuous bracelet. Anyway, three-dimensional structure of intact giant hemoglobin must be awaited.

The tube worm giant hemoglobins are believed to have a function to transport H₂S (Arp et al., 1987). As a H₂S binding site, we proposed previously that a free cysteine residue is responsible for sulfide binding ability (Suzuki et al., 1990). Since Lamellibrachia chain AV-1 is rich in half-cystines (9 residues, at positions 64, 71, 77, 84, 90, 101, 120, 173, and 216 (Fig. 1)) and our preliminary experiment by using a fluorescent labeling suggested that chain AV has a reactive thiol in the intact Lamellibrachia 3000 kDa molecule (Suzuki et al., 1990), it is likely that the chain AV acts also as a H₂S carrier.

REFERENCES

Arp, A. J., Childress, J. J. & Vetter, R. D. (1987) J. Exp. Biol. 128, 139–158
Feng, D. F., Johnson, M. S. & Doolittle, R. F. (1985) J. Mol. Biol. 21, 112–125
Fushitani, K. & Riggs, A. F. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 9461–9463
Jhiang, S. M., Carey, J. R. & Riggs A. F. (1988) Science 240, 334–336
Kyte, J. & Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
Lightbody, J. J., Quabar, A. N., Mainwaring, M. G., Young, J. S., Waltz, D. A., Vinogradov, S. N. & Gotoh, T. (1988) Comp. Biochem. Physiol. 90B, 301–305
Suzuki, T. & Gotoh, T. (1986) J. Biol. Chem. 261, 9257–9267
Suzuki, T., Takagi, T. & Ohta, S. (1988) Biochem. J. 255, 541–545
Suzuki, T., Takagi, T., Okuda, K., Furukohri, T. & Ohta, S. (1989) Zool. Sci. 6, 915–926
Suzuki, T., Takagi, T. & Ohta, S. (1990) Biochem. J., in press
Vinogradov, S. N. (1985) Comp. Biochem. Physiol. 82B, 1–15
Vinogradov, S. N., Lugo, S. D., Mainwaring, M. C., Kappr, O. H. & Crewe, A. V. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 8024–8028
Table 9. Amino acid composition of certain proteins and peptides of Leucon Americana.

| A | Mole(600/154) | (51/1500) |
|---|---------------|-----------|
| 1 | mass | 116 | 117 |
| 2 | 176 | 116 |
| 3 | 312 | 312 |
| 4 | 357 | 357 |
| 5 | 297 | 297 |
| 6 | 275 | 275 |
| 7 | 277 | 277 |
| 8 | 279 | 279 |
| 9 | 280 | 280 |
| 10 | 281 | 281 |

Table 10. (continued)

| A | 1554 | 1555 |
|---|------|------|
| 1 | Mole(600/154) | (51/1500) |
| 2 | mass | 116 | 117 |
| 3 | 176 | 116 |
| 4 | 312 | 312 |
| 5 | 357 | 357 |
| 6 | 297 | 297 |
| 7 | 275 | 275 |
| 8 | 277 | 277 |
| 9 | 279 | 279 |
| 10 | 280 | 280 |

Table 2. (continued)

| A | 1554 | 1555 |
|---|------|------|
| 1 | Mole(600/154) | (51/1500) |
| 2 | mass | 116 | 117 |
| 3 | 176 | 116 |
| 4 | 312 | 312 |
| 5 | 357 | 357 |
| 6 | 297 | 297 |
| 7 | 275 | 275 |
| 8 | 277 | 277 |
| 9 | 279 | 279 |
| 10 | 280 | 280 |
Table 2 (continued)

| Time (min) | 0 | 20 | 40 | 60 |
|------------|---|----|----|----|
|             |   |    |    |    |
| 200 A       |   |    |    |    |
| 1200 (cont.)|   |    |    |    |
| 1800        |   |    |    |    |
| 2400        |   |    |    |    |
| 3000        |   |    |    |    |
| 3600        |   |    |    |    |
| 4200        |   |    |    |    |
| 4800        |   |    |    |    |
| 5400        |   |    |    |    |
| 6000        |   |    |    |    |
| 6600        |   |    |    |    |
| 7200        |   |    |    |    |
| 7800        |   |    |    |    |
| 8400        |   |    |    |    |
| 9000        |   |    |    |    |
| 9600        |   |    |    |    |
| 10200       |   |    |    |    |
| 10800       |   |    |    |    |
| 11400       |   |    |    |    |
| 12000       |   |    |    |    |

Fig. 3. Reversed-phase chromatography of the endoproteinase Asp-N peptide of Linker subunit 1555. The column (Vydac 3100 3000) was equilibrated with Buffer A, followed by a linear gradient of acetonitrile in Buffer B at a flow rate of 0.5 mL/min.

Fig. 4. Reversed-phase chromatography of the heparin-Sepharose peptide of Linker subunit 1555. The column (Hepac 3100 3000) was equilibrated with Buffer A, followed by a linear gradient of acetonitrile in Buffer B at a flow rate of 0.5 mL/min.

Downloaded from http://www.jbc.org/ by guest on March 23, 2020.
Primary structure of a linker subunit of the tube worm 3000-kDa hemoglobin.
T Suzuki, T Takagi and S Ohta

J. Biol. Chem. 1990, 265:1551-1555.

Access the most updated version of this article at http://www.jbc.org/content/265/3/1551

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/265/3/1551.full.html#ref-list-1