Primary structure of a dimeric haemoglobin from the deep-sea cold-seep clam *Calymptogena soyoae*

Tomohiko SUZUKI,* Takashi TAKAGI† and Suguru OHTA‡

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

The heterodont clam *Calymptogena soyoae*, living in the cold-seep area of the upper bathyal depth of Sagami Bay, Japan, has two homodimeric haemoglobins (Hb I and Hb II) in erythrocytes. The complete amino acid sequence of 136 residues of *C. soyoae* Hb II was determined. The sequence showed low homology with any other globins (at most 20% identity) and lacked the *N*-terminal extension of seven to nine amino acid residues characteristic of all the molluscan haemoglobins sequenced hitherto. Although the subunit assembly of molluscan haemoglobin is known to be 'back-to-front' relative to vertebrate haemoglobin, *C. soyoae* Hb II is unlikely to undergo such a subunit assembly because it lacks homology in the sequence involving subunit interaction. These structural features suggest that *C. soyoae* haemoglobin may have accomplished a unique molecular evolution. The distal (E7) histidine residue of *C. soyoae* Hb II is unusually replaced by glutamine. However, the oxyhaemoglobin is stable enough to act as an O₂ carrier, since the autoxidation rate at near physiological temperature (3 °C) is about 3 times lower than that of human haemoglobin at 37 °C. H.p.l.c. patterns of peptides (Figs. 5–7), amino acid compositions of intact protein and peptides (Table 1) and amino acid sequences of intact protein and peptides (Tables 2 and 3) have been deposited as Supplementary Publication SUP 50150 (11 pages) at the British Library Document Supply Centre, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies may be obtained on the terms indicated in Biochem. J. (1989) 257, 5.

INTRODUCTION

Deep-sea investigations revealed that unique biological communities are present around the areas of hydrothermal vents or cold seeps at a depth of 1100–6000 m (Corliss *et al.*, 1979; Kennicutt *et al.*, 1985; Ohta & Laubier, 1987), in which the most conspicuous animals are the giant clam *Calymptogena* and the giant tube worms *Riftia* and *Lamellibrachia*. These invertebrate animals are sustained by mutual symbiosis with sulphide-oxidizing bacteria (Childress *et al.*, 1987), and their blood, containing abundant haemoglobin(s), is believed to facilitate O₂ transport to the site of carbon fixation (Arp & Childress, 1981; Terwilliger *et al.*, 1983).

The subunit structure and preliminary O₂-binding studies of the hydrothermal-vent clam *Calymptogena magnifica* haemoglobin have been reported previously (Terwilliger *et al.*, 1983). Here we report the subunit structure, primary structure and stability at near physiological conditions of haemoglobin from the cold-seep clam *Calymptogena soyoae*.

MATERIALS AND METHODS

C. soyoae clams (11–13 cm long) were collected from the cold-seep area located 34°59.90'N, 139°13.60'E at a depth of 1160 m in Sagami Bay, southeast of Hatsushima, Japan, by the Japanese submersible *Shinkai 2000* during November of 1987 (dives 314 and 315) (Okutani & Egawa, 1985; Ohta *et al.*, 1987). The temperature of the water bathing the animals was about 2.8 °C (3.1 °C

within the sediment 10 cm below the sea floor), and the O₂ concentration was 56–84 μM. As soon as the animals were brought to the surface, retained within a tightly closed heat-insulation box, blood was removed from the circulating system by using a syringe and stored at –80 °C or –40 °C until use.

The haemolysate was centrifuged at 32600 *g* for 15 min at 2 °C, and the resultant haemoglobin solution was applied to a gel-filtration column (1 cm × 30 cm) of Superose 12 (Pharmacia, Uppsala, Sweden). The column was equilibrated with 50 mM-sodium phosphate buffer, pH 7.2, containing 150 mM-NaCl and eluted with the same buffer at a flow rate of 0.5 ml/min at 15–17 °C. Haemoglobin solution in 10 mM-Tris/HCl buffer (pH 8.5) was applied to a column (7.5 mm × 75 mm) of TSK-Gel DEAE-5PW (Tosoh, Tokyo, Japan) and eluted with a linear gradient of 0–0.2 M-NaCl in 10 mM-Tris/HCl buffer, pH 8.5, over 60 min at a flow rate of 1 ml/min at 15–17 °C. Two haemoglobin components and haemolysate were each applied to a reverse-phase column (4.6 mm × 150 mm) of Cosmosil 5C₁₈-300 (Nakarai, Kyoto, Japan) equilibrated with 40% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid and eluted with a linear gradient of 40–80% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid over 30 min at a flow rate of 1 ml/min.

C. soyoae Hb II was selected for sequence determination. Methods of removal of haem, carboxymethylation of cysteine and manual Edman degradation were the same as described previously (Suzuki, 1986;

Abbreviation used: Hb, haemoglobin.

* To whom all correspondence should be addressed.

Suzuki & Gotoh, 1986). The whole protein (65 nmol) was digested with lysyl endopeptidase (Wako, Tokyo, Japan) at an enzyme/substrate ratio of 1:200 (w/w) in 10 mM-Tris/HCl buffer, pH 8.9, at 37 °C for 6 h, and the resultant peptides were purified by h.p.l.c. on a reverse-phase column (Fig. 5 in Supplement SUP 50150). Peptide

L1 was digested further with trypsin in 0.1 M-NH₄HCO₃ at 37 °C for 2 h (Fig. 6 in Supplement SUP 50150). To obtain overlap peptides, the whole protein (50 nmol) was also digested with pepsin at an enzyme/substrate ratio of 1:100 (w/w) in 5% (v/v) formic acid at 37 °C for 40 min (Fig. 7 in Supplement SUP 50150). Amino acid com-

Fig. 1. (a) Gel filtration of *C. soyae* haemoglobin by h.p.l.c., (b) separation of *C. soyae* haemoglobins (Hb I and Hb II) by h.p.l.c. and (c) reverse-phase h.p.l.c. of *C. soyae* Hb I, of *C. soyae* Hb II and of *C. soyae* haemolysate

(a) Gel filtration of *C. soyae* haemoglobin (200 μ M-haem) by h.p.l.c. The column of Superose 12 was equilibrated with 50 mM-phosphate buffer, pH 7.5, containing 150 mM-NaCl and eluted with the same buffer at a flow rate of 0.5 ml/min. *C. soyae* haemoglobin was eluted at a retention time (R.T.) of 30.27 ± 0.10 min. The markers used were human haemoglobin (M_r 64000, R.T. 29.00 min), sperm-whale myoglobin (M_r 17500, R.T. 31.25 min) and horse cytochrome *c* (M_r 12500, R.T. 31.65 min). (b) Separation of *C. soyae* haemoglobins (Hb I and Hb II) by h.p.l.c. The sample from the gel-filtration column of Superose 12 was dialysed against 10 mM-Tris/HCl buffer, pH 8.5, and then applied to a TSK-Gel DEAE-5PW column. The column was equilibrated with 10 mM-Tris/HCl buffer, pH 8.5, and eluted with a linear gradient of NaCl (----) at a flow rate of 1 ml/min. Two major peaks, Hb I and Hb II, retained haem. (c) Reverse-phase h.p.l.c. of (i) *C. soyae* Hb I, of (ii) *C. soyae* Hb II and of (iii) *C. soyae* haemolysate. In each case a column of Cosmosil 5C₁₈-300 was equilibrated with 40% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid and eluted with a linear gradient of acetonitrile (----) at a flow rate of 1 ml/min. The sample was dissolved in 70% (v/v) formic acid.

positions of peptides are shown in Table 1 in Supplement SUP 50150. The *N*-terminal 43 residues were determined directly by use of an automated sequencer (Applied BioSystems 477A Protein Sequencer) (Table 2 in Supplement SUP 50150), and the peptides obtained by digestion with enzymes were sequenced by the manual Edman method (Table 3 in Supplement SUP 50150).

The autoxidation measurements were made under air-saturated conditions (Suzuki, 1986) at 3, 10, 15, 20, 25 and 26 °C. The autoxidation rate is given by the equation :
$$-d[\text{Oxyhaemoglobin}]/dt = k_{\text{obs.}}[\text{oxyhaemoglobin}]$$

where $k_{\text{obs.}}$ represents the observed first-order rate constant at a given temperature.

Fig. 2. Summary of data establishing the amino acid sequence of *C. soyae* Hb II
Automated sequencer or manual Edman degradation (□) was employed for sequence determination. Key: L, a lysyl endopeptidase peptide; T, a tryptic peptide; P, a peptic peptide.

Fig. 3. Alignment of the amino acid sequencer of *C. soyoae* Hb II with those of human and *A. trapezia* β -chains

The alignment is based on the assumption that the helical segments present in most other globins are also present in *C. soyoae* Hb II. Several deletions are used to obtain significant correspondence between the sequences. The boxed residues indicate the 15 invariable residues in the three globins. Helices and β -turns predicted by the method of Chou & Fasman (1978) are shown by --- and — respectively. The distal (E7) amino acid residues are indicated by the arrow.

RESULTS AND DISCUSSION

On a gel-filtration column of Superose 12 *C. soyoae* haemoglobin was eluted as a single peak corresponding to an apparent M_r of 30000, a value for a dimeric structure (Fig. 1a). Purified dimeric haemoglobin was separated on a column of TSK-Gel DEAE-5PW into two components, Hb I and Hb II, as shown in Fig. 1(b). The M_r of Hb I and Hb II was confirmed to be 30000 on a gel-filtration column of Superose 12. By reverse-phase h.p.l.c. each component was found to consist of a single polypeptide chain (Fig. 1c). From these results we concluded that *C. soyoae* haemoglobin is composed of two homodimers in approximately equal proportions.

Among the invertebrate haemoglobins, the molluscan forms show a remarkable diversity in subunit structure; so far, homodimeric, heterodimeric, tetrameric and didomain-polymeric haemoglobins have been found in the circulating erythrocytes of several arcid clams (Terwilliger & Terwilliger, 1985). The diversity seems to be wide even in congeneric clams, because *C. soyoae* haemoglobin is dimeric, as shown in Fig. 1, and that of *C. magnifica* is tetrameric (Terwilliger *et al.*, 1983).

The amino acid sequence of *C. soyoae* Hb II was determined mainly by the manual Edman method. The strategies used to establish the complete sequence are shown in Fig. 2, and the detailed results are given in Figs. 5–7 and Tables 1–3 in Supplement SUP 50150. *C. soyoae* Hb II is composed of 136 amino acid residues, and the M_r was calculated to be 15846 including a haem group.

The amino acid sequence of *C. soyoae* Hb II is aligned with those of the β -chains of human and the arcid clam *Anadara trapezia* haemoglobins (Gilbert & Thompson, 1985) in Fig. 3. Of the 136 amino acid residues in

C. soyoae Hb II, 28 (20%) and 29 (21%) residues are identical with those in the corresponding positions in human and *A. trapezia* β -chains. In the three polypeptide chains 15 residues including two essential haem contact residues, CD1-Phe and F8-His, appear to be invariant. *C. soyoae* Hb II also shows low homology with other globins (at most 20%). On the other hand, the *N*-terminal 70 residues of *C. soyoae* Hb I (T. Suzuki, T. Takagi & S. Ohta, unpublished work) showed 39% homology with those of *C. soyoae* Hb II.

The polypeptide chain of *C. soyoae* Hb II is the shortest in the globins so far sequenced, mainly owing to the absence of the D helix (positions 60–66 in Fig. 3) and the deletions of nine residues in the G and H helices (positions 129–132 and 137–141) in our alignment. It is well known that α -chains of vertebrate haemoglobins and several invertebrate globins lack the D helix.

The hydropathy profile (Kyte & Doolittle, 1982) of *C. soyoae* Hb II resembled that of human α -chain. Furthermore, according to the method of Chou & Fasman (1978), five turns were predicted at the segments A–B, B–C, D–E, G–H and H–C (Fig. 3), most of which located at the sites where the hydrophobicity is at a local minimum. A, C, E, F and G helices were also predicted. These considerations, taken together, indicate that our alignment of the segments in *C. soyoae* Hb II is reliable and that the globin folding of *C. soyoae* Hb II is similar to that of a vertebrate myoglobin.

An *N*-terminal extension of seven to nine amino acid residues, forming an additional pre-A helix (Royer *et al.*, 1985), is characteristic of all the molluscan haemoglobins sequenced hitherto, namely α -, β - and homodimeric chains of *A. trapezia*, α - and homodimeric chains of *Anadara broughtonii*, a homodimeric chain of *Scapharca*

Fig. 4. Plot of $\log k_{\text{obs}}$ versus $1/T$ for the autoxidation of *C. soyoe* oxyhaemoglobin in 0.1 M-phosphate buffer, pH 7.2, containing 0.1 M-NaCl

The activation energy (E_a) was calculated from the slope of the straight line. ●, *C. soyoe* Hb II; ○, *C. soyoe* haemolysate. The haemoglobin concentration was 31 μM .

and an unusual didomain chain of *Barbatia* (Gilbert & Thompson, 1985; Furuta & Kajita, 1983; Petruzzelli *et al.*, 1985; Riggs *et al.*, 1986). However, *C. soyoe* Hb II lacked this characteristic extension completely (Fig. 3).

The subunit assembly of molluscan dimeric and tetrameric haemoglobins is known to be 'back-to-front' relative to vertebrate haemoglobins (Royer *et al.*, 1985), and therefore E and F helices involving subunit interaction are highly conserved among the seven molluscan haemoglobins so far sequenced (Furuta & Kajita, 1983; Gilbert & Thompson, 1985; Petruzzelli *et al.*, 1985; Riggs *et al.*, 1986); 19 (66%) out of 29 residues in E and F helices are identical, whereas the homology of the remaining segments is at most 30%. However, when the residues in E and F helices of *C. soyoe* Hb II were compared with those of *A. trapezia* β -chain (Fig. 3), only four residues (14%) were identical. Therefore *C. soyoe* Hb II is unlikely to undergo such a 'back-to-front' subunit assembly as seen in other molluscan haemoglobins. Consequently, these two structural features, namely showing low homology with other molluscan haemoglobins even in the E and F helices and lacking the characteristic N-terminal extension, led us to speculate that *C. soyoe* haemoglobin may have accomplished a unique molecular evolution. Alternatively, *C. soyoe* haemoglobin might evolve from a different gene strain compared with that of usual molluscan haemoglobins.

Another interesting feature of *C. soyoe* Hb II is the replacement of the distal (E7) histidine residue by glutamine (Fig. 3). Most globins have histidine at the distal position, which is capable of forming a hydrogen bond to the bound O_2 and stabilizing it (Phillips & Schoenborn, 1981). The replacement by glutamine is found only, but at the highest frequency, in several shark and elephant myoglobins (Fisher *et al.*, 1981; Suzuki, 1987; Romero-

Herrera *et al.*, 1981) and in opossum, hagfish, *Urechis* and *Vitreoscilla* haemoglobins (Stenzel *et al.*, 1979; Liljeqvist *et al.*, 1979; Garey & Riggs, 1986; Wakabayashi *et al.*, 1986). Studies with mutant human haemoglobin with E7-Gln produced by protein engineering show that E7-Gln can form a hydrogen bond (Nagai *et al.*, 1987), like E7-His. It is, however, also well known that the replacement of E7-His often causes a rapid autoxidation that converts oxyhaemoglobin into a physiologically inactive met form. Therefore we examined the stability of *C. soyoe* oxyhaemoglobin with E7-Gln in 0.1 M-sodium phosphate buffer, pH 7.2, at various temperatures.

Fig. 4 shows a plot of $\log k_{\text{obs}}$ versus $1/T$ for the autoxidation of *C. soyoe* oxyhaemoglobin. The autoxidation rate at near-physiological temperature (3 °C) and pH 7.2 was determined to be 0.0014 h^{-1} , which was about 3 times lower than that (0.0037 h^{-1}) of human haemoglobin (β -chain) at 37 °C and pH 7.2 (Mansouri & Winterhalter, 1973). Furthermore the autoxidation reaction of *C. soyoe* haemoglobin appears to be protected by a higher activation energy (E_a 153 kJ/mol) when compared with that (110 kJ/mol) of mammalian myoglobin (Gotoh & Shikama, 1974). Hence it may be concluded that *C. soyoe* haemoglobin is stable enough to act as an O_2 carrier at physiological conditions. However, it should be noted that *C. soyoe* haemoglobin is, surprisingly, autoxidized 1300 times faster than human haemoglobin under the same conditions (37 °C, pH 7.2), if we calculate the rate on the basis of the E_a value of *C. soyoe* haemoglobin. Detailed analyses for both the autoxidation reaction and the three-dimensional structure must be awaited to understand this phenomenon.

Terwilliger *et al.* (1983) showed that *C. magnifica* haemoglobin has a relatively high O_2 affinity, and now we have shown that *C. soyoe* haemoglobin is resistant to autoxidation at least under physiological conditions. These two characteristics of the deep-sea clam *C. soyoe* haemoglobin must play an important role in facilitating the effective uptake of O_2 and the transport of O_2 to be used by the chemoautotrophic sulphide-oxidizing bacteria (Arp & Childress, 1981; Terwilliger *et al.*, 1983; Childress *et al.*, 1987).

REFERENCES

- Arp, A. J. & Childress, J. J. (1981) *Science* **213**, 342–344
- Childress, J. J., Felbeck, H. & Somero, G. N. (1987) *Sci. Amer.* **256**, 106–112
- Chou, P. H. & Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47**, 251–276
- Corliss, J. B., Dymond, J., Gordon, L. I., Edmond, J. M., von Herzen, R. P., Ballard, R. D., Green, K., Williams, D., Bainbridge, A., Crane, K. & van Andel, T. H. (1979) *Science* **203**, 1073–1083
- Fisher, W. K., Koureas, D. D. & Thompson, E. O. P. (1981) *Aust. J. Biol. Sci.* **34**, 5–10
- Furuta, H. & Kajita, A. (1983) *Biochemistry* **22**, 917–922
- Garey, J. R. & Riggs, A. F. (1986) *J. Biol. Chem.* **261**, 16446–16450
- Gilbert, A. T. & Thompson, E. O. P. (1985) *Aust. J. Biol. Sci.* **38**, 221–236
- Gotoh, T. & Shikama, K. (1974) *Arch. Biochem. Biophys.* **163**, 476–481
- Kennicutt, M. C., II, Brooks, J. M., Bidigare, R. R., Fay, R. R., Wade, T. L. & McDonald, T. J. (1985) *Nature (London)* **317**, 351–353
- Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132

- Liljeqvist, G., Braunitzer, G. & Paleus, S. (1979) Hoppe-Seyler's Z. Physiol. Chem. **360**, 125–135
- Mansouri, A. & Winterhalter, K. H. (1973) Biochemistry **12**, 4946–4949
- Nagai, K., Luisi, B., Shih, D., Miyazaki, G., Imai, K., Poyart, C., Young, A. D., Kwiatkowski, L., Nobel, R. W., Lin, S. H. & Yu, N. T. (1987) Nature (London) **329**, 858–860
- Ohta, S. & Laubier, L. (1987) Earth Planet. Sci. Lett. **83**, 329–342
- Ohta, S., Sakai, H., Taira, A., Ohwada, K., Ishii, T., Maeda, M., Fujioka, K., Saino, K., Kogure, K., Gamo, T., Shirayama, Y., Furuta, T., Ishizuka, T., Endow, K., Sumi, T., Hotta, H., Hashimoto, J., Handa, N., Masuzawa, T. & Horikoshi, M. (1987) JAMSTECTR Deepsea Res. (in Japanese with English summary and legends) **3**, 51–60
- Okutani, T. & Egawa, K. (1985) Venus (Jpn. J. Malacol.) **44**, 285–289
- Petruzzelli, R., Goffredo, B. M., Barra, D., Bossa, F., Boffi, A., Verzili, D., Ascoli, F. & Chiancone, E. (1985) FEBS Lett. **184**, 328–332
- Phillips, S. E. V. & Schoenborn, B. P. (1981) Nature (London) **292**, 81–82
- Riggs, A. F., Riggs, C. K., Lin, R. J. & Domdey, H. (1986) in Invertebrate Oxygen Carriers (Linzen, B., ed.), pp. 473–476, Springer-Verlag, Berlin
- Romero-Herrera, A. E., Goodman, M., Dene, H., Bartnicki, D. E. & Mizukami, H. (1981) J. Mol. Evol. **17**, 140–147
- Royer, W. E., Love, W. E. & Fenderson, F. F. (1985) Nature (London) **316**, 277–280
- Stenzel, P., Brimhall, B., Jones, R. T., Black, J. A., McLachlan, A. & Gibson, D. (1979) J. Biol. Chem. **254**, 2071–2076
- Suzuki, T. (1986) J. Biol. Chem. **261**, 3692–3699
- Suzuki, T. (1987) Biochim. Biophys. Acta **914**, 170–176
- Suzuki, T. & Gotoh, T. (1986) J. Biol. Chem. **261**, 9257–9267
- Terwilliger, R. C. & Terwilliger, N. B. (1985) Comp. Biochem. Physiol. B **81**, 255–261
- Terwilliger, R. C., Terwilliger, N. B. & Arp, A. (1983) Science **219**, 981–983
- Wakabayashi, S., Matsubara, H. & Webster, D. A. (1986) Nature (London) **322**, 481–483

Received 1 August 1988/5 December 1988; accepted 13 December 1988