Vol. 182, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1992 Pages 195-200

MOLECULAR CLONING OF THE cDNA FOR THE MAJOR HEMOGLOBIN COMPONENT FROM *PARAMECIUM CAUDATUM* +

Kiyoshi Yamauchi[1]*, Masanori Mukai[1], Takehiko Ochiai[2] and Itaru Usuki[3]

[1]Department of Biology, Faculty of Science, Shizuoka University, Oya 836, Shizuoka 422, JAPAN

[2]Department of Biology, Fukushima Medical College, Hikarigaoka 1, Fukushima 960-12, JAPAN

[3]Department of Biology, College of General Education, Niigata University, Igarashi 2-no-cho 8050, Niigata, JAPAN

Received November 7, 1991

SUMMARY. Nucleotide sequence of the cDNA for the major hemoglobin component of *Paramecium caudatum* was determined. An oligonucleotide was synthesized on the basis of the amino acid sequence, and the *Paramecium* cDNA library constructed in phage λgt11 was screened with it. Three positive clones, of which insert sizes were 0.4, 0.6, and 0.9 kbp, were obtained. Sequence analysis made clear that the 0.4-kbp cDNA retains a full length of the nucleotides encoding 116 amino acid residues, and that in the coding region it contains four TAA codons which are known to encode glutamine.

Hemoglobin (Hb) distributes widely in a variety of organisms, including vertebrates, invertebrates, bacteria and plants. The ciliated protozoa, *Paramecium*, is known to possess a heterogeneous monomeric Hbs showing 11-13 kilo dalton (kDa) in molecular mass (1,2). The Hbs of *P. caudatum* consists of a major component and several minors (3,4). The concentration of the major Hb component reaches to 75-90% of the whole Hb (4), which corresponds to 0.4-0.9% of the total protein (5). By direct protein sequencing it has recently been shown that the major Hb component of *P. caudatum* consists of 116 amino acid residues, and that the sequence is very unique when it is compared with those of vertebrate and invertebrate Hbs (6). In this paper, we report for the first time the cDNA sequence encoding the major Hb component from *P. caudatum*, to make a contribution to increase our knowledge on the globin gene structure of many invertebrates.

+ The sequence presented has been submitted to the EMBL/GenBank database under the accession number M57542.

* Author to whom correspondence should be addressed.

0006-291X/92 $1.50

MATERIALS AND METHODS

P. caudatum (syngen 3, stock YC) was cultured, then the Hb was extracted and partially purified as described in the previous papers (4,7). The Hb sample obtained (400 μg/lane) was further purified by native polyacrylamide gel electrophoresis (PAGE) as described by Laemmli (8) except that SDS was not added, to prepare specific polyclonal antibody against it. The eluted Hb (70 μg) was mixed with an equal amount of the complete Freund's adjuvant (Calbiochem-Behring, CA), and it was injected subcutaneously into two rabbits. Two or three additional immunizing doses of the immunogen were administered four and six weeks later. Blood was collected weekly after first additional immunizing. Western blotting was carried out by the method of Burnette (9). Immunoreactive bands were detected using the goat anti-rabbit IgG (Fc) alkaline phosphatase conjugate, according to the protocol supplied by the manufacture (Promega Corp. MD)

Oligonucleotide probe was synthesized on the basis of the amino acid sequence of the Hb from *P. caudatum* (Fig. 1), using the Applied Biosystems Model 380A DNA synthesizer (Foster City, CA). Twenty pmol of oligonucleotide was labeled at the 5' end with T4 polynucleotide kinase (Takara Shuzo, Japan) in a 70 μl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 5 mM dithiothreitol, 0.1 mM spermidine, 0.1 mM EDTA and 200 μCi (7.4 MBq) [γ-^{32}P]ATP. The reaction was carried out at 37°C for 1 h, and the labeled oligonucleotide was separated from free [γ-^{32}P]ATP by Nick-Columns (Pharmacia LKB Biotechnology, Sweden). *Paramecium* RNA was prepared by the guanidine isothiocyanate method (10). Poly(A)$^+$ RNA was isolated by adsorption to and elution from oligo(dT) cellulose (11). A cDNA library was prepared according to the method of Young & Davis (12) with slight modifications. The cDNA was ligated into the *Eco*RI restriction site of λgt11, packaged *in vitro*, and recombinant phages plated on *E. coli* Y1090.

The library was screened using the polyclonal antibody or the 32-mer synthetic oligonucleotide as a probe. In the antibody screening, the recombinants on the nylon filters (MSI, Micron Separations Inc, MA) were incubated with the anti-*Paramecium* Hb antiserum (1:200) and followed by incubation with the biotinylated anti-rabbit IgG, subsequently with the streptavidin-biotinylated alkaline phosphatase preformed complex according to the protocol supplied by the manufacture (Amersham International plc, UK). In the oligonucleotide screening, the recombinants on the nylon filters were hybridized with the ^{32}P end-labeled 32-mer oligonucleotide probe (2 x 10^7 cpm/ml) in the hybridization buffer containing 10 x Denhardt's solution, 10 x SSC and 0.1% SDS at 30°C overnight. The filters were washed two times with 5 x SSC, then with 4 x SSC containing 0.1% SDS at room temperature, and they were autoradiographed on Kodak XAR-5 film for 8-16 h at -80°C. The insert cDNA in recombinant phage DNA was subcloned into the *Eco*RI site of pUC118 or pUC119 plasmid (13). To confirm whether the isolated cDNAs encodes Hb, the *Pvu*II-*Sal*I fragment, a 3' region of cDNA, was ligated to *Sma*I and *Sal*I sites in the expression vector, pUEX1, which can produce the β-galactosidase-fusion protein (14). After verifying the construction of the recombinant pUEX1 by the Southern blotting, we

	(32)										(42)
Amino acid sequence:	Phe	Phe	Asn	Gly	Ile	Asp	Met	Pro	Asn	Gln	Thr
Deduced codon:	TTC	TTC	AAT	GGA	ATT	GAT	ATG	CCA	AAT	TAA	ACA
						*			*		
Codon frequency:	21	21	78	52	18	57	12	19	78	36	97
	30	30	97	89	48	72	12	28	97	52	206

Figure 1. Synthetic oligonucleotide probe for isolation of the *Paramecium* hemoglobin cDNA. The oligonucleotide probe was synthesized at a region corresponding to the position of 32-42 in the amino acid sequence reported for the major Hb component of *Paramecium* (6). Codon selections were made on the basis of the highest frequency in the codon usage reported for the *Paramecium* surface antigens (20). Mismatches between the probe sequence and the actual cDNA sequence are marked (*).

cultured the transformant *E. coli* HB101 possessing the intact pUEX1 (for a control) or the recombinant pUEX1 (for the experiment) at 32°C. At a late logarithmic phase, the cells were subsequently cultured for 2 h at 42°C to express the β-galactosidase-fusion protein under the control of λP_L promoter and cI857. The collected cells (about 6×10^7 cells) were immediately disrupted by an addition of the SDS-sample buffer, and they were boiled for 3 min. The resultants were resolved by SDS-PAGE, and immunoreactivity of proteins transferred to the PVDF filter was checked up with the anti-*Paramecium* Hb antibody. Sequencing of the cDNA was performed on a single stranded plasmid by the dideoxynucleotide chain termination method (15).

RESULTS AND DISCUSSION

Figure 2A shows that the antiserum recognizes the Hb of *Paramecium*, but it does not reacted with those of other species, such as *Xenopus*, mouse and human (data not shown). We therefore screened *Paramecium* expression library in λgt11 (7×10^5 recombinants) by using the specific antibody. However, this trial was failed to obtain any positive clones through two examinations. On the other hand, the same library gave three positive clones by screening of 3×10^5 recombinants using the ^{32}P-labeled synthetic oligonucleotide probe (Fig. 1). To estimate insert sizes, each of DNAs in the clones was digested with *Eco*RI. The digest was size-fractionated on a 1% agarose gel, and the resultant was transferred to nylon filter, then the objective cDNA was probed with the ^{32}P-labeled oligonucleotide. Size of the insert DNA from the three clones were 0.9, 0.6 and 0.4 kbp, which were termed Clone HbI, Clone HbII and Clone HbIII, respectively.

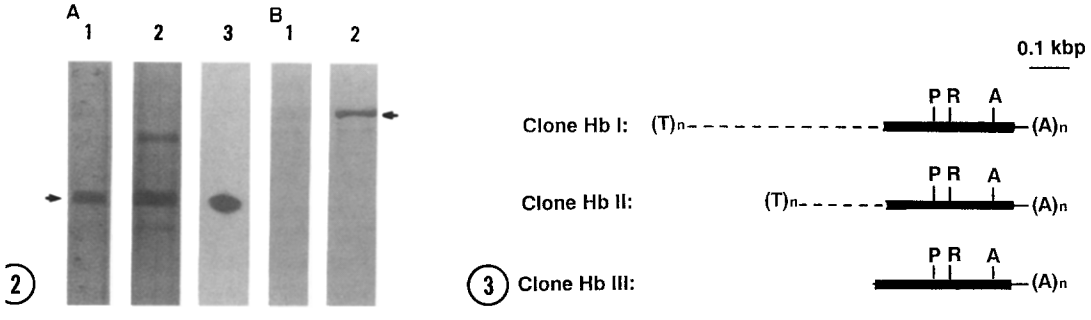

Figure 2. Immunological detection of the *Paramecium* hemoglobin (A) and of the epitope of hemoglobin fused with β-galactosidase (B). (A) Partially purified major Hb component (10μg/lane) from *Paramecium caudatum* was electrophoresed in 15% native polyacrylamide gel. Lane 1: No staining. Lane 2: Staining with Coomassie brilliant blue. Lane 3: Arrow shows a band indicating the presence of major component of hemoglobin. (B) Bacterial extracts containing a *Paramecium* Hb-β-galactosidase-fusion protein were subjected to SDS-PAGE. Lane 1: Control showing no immunoreactivity. Lane 2: Arrow shows a distinct band indicating a production of β-galactosidase-fusion protein which reveals an immunoreactive epitope to the anti-*Paramecium* Hb antibody.

Figure 3. Restriction map of the isolated cDNAs for the *Paramecium* hemoglobin. Solid boxes and smooth lines indicate the coding and noncoding regions, respectively. Broken lines indicate unknown cDNA which might be ligated artificially with Hb cDNA in a step of the construction of cDNA library. A, P, R, $(T)_n$ and $(A)_n$ represent *Aat*II, *Pvu*II and *Rsa*I restriction sites, and poly(A) and poly(T) tracts, respectively.

As shown in Fig. 3, the 3' portion of these cDNA have a common restriction map each other. Furthermore, the insert of Clone HbI, in particular, was longer in size than those of other known Hb mRNAs. These result suggest that some of the isolated clones have an artificial construct. To verify whether the isolated cDNAs encode hemoglobin, the 3' portion of it, downstream from a internal *PvuII*I site, was expressed as β-galactosidase-fusion protein in *E. coli.* The cell extract transformed with pUEX1 containing a 3' region of the cDNA, produced a fusion protein immunoreactive to the anti-*Paramecium* Hb antibody (Fig. 2B, lane 2), but the cell extract transformed with intact pUEX1 not reacted it (Fig. 2B, lane 1), showing that at least 3' portion of the cDNA encodes the *Paramecium* Hb. Although we also tried to express the 5' portion of the cDNAs, upstream from the *Pvu*II site, in *E. coli* with a similar method, no immunoreactive band was obtained.

The nucleotide sequencing of the cDNAs showed that both of Clone HbI and Clone HbII lacked a certain length of 5' coding region, and that the head of the globin cDNAs may be artificially ligated with unknown cDNA at the position of nucleotide 23, probably in a process of the preparation of cDNA library (Fig. 3). Whereas, Clone HbIII was found to retain a full length of nucleotide sequence from the translational initiation site to the termination site as shown in Fig. 4. All of the insert cDNAs preserved a stretch of 20-30 adenylate residues corresponding to 3' polyadenylation terminal of the mRNA.

```
1                                                          GAAAATG
                                                             *

8       TCTCTCTTCGAACAATTAGGAGGATAAGCCGCCGTCCAAGCCGTCACAGCTCAATTCTAT
        S  L  F  E  Q  L  G  G  Q  A  A  V  Q  A  V  T  A  Q  F  Y     20

68      GCTAACATCTAAGCTGATGCTACAGTAGCCACATTCTTCAATGGAATTGACATGCCAAAC
        A  N  I  Q  A  D  A  T  V  A  T  F  F  N  G  I  D  M  P  N     40

128     TAAACAAACAAGACAGCTGCCTTCTTATGCGCTGCCCTCGGAGGACCAAATGCCTGGACC
        Q  T  N  K  T  A  A  F  L  C  A  A  L  G  G  P  N  A  W  T     60

188     GGAAGAAATTTGAAGGAAGTACATGCCAACATGGGAGTCAGTAACGCCCAATTCACAACT
        G  R  N  L  K  E  V  H  A  N  M  G  V  S  N  A  Q  F  T  T     80

248     GTCATTGGACATTTGAGAAGTGCATTGACTGGCGCTGGTGTTGCTGCTGCTTTGGTTGAA
        V  I  G  H  L  R  S  A  L  T  G  A  G  V  A  A  A  L  V  E    100

308     TAAACGGTCGCTGTTGCTGAAACAGTCAGAGGAGACGTCGTCACAGTCTGATTTATATTC
        Q  T  V  A  V  A  E  T  V  R  G  D  V  V  T  V  *             116

368     ACACAATAAACAAATATATC(A)n
```

Figure 4. Nucleotide and deduced amino acid sequence of the major hemoglobin component cDNA from *Paramecium caudatum.* Nucleotide number is marked in the left column, and amino acid number is given in the right column. The region referred for synthesis of the oligonucleotide probe used is underlined. The initiation codon and termination codon for translation were marked (*). The AATAAA polyadenylation signal (16) is broken-lined.

The sequence of AATAAA, which was proposed to act as a signal sequence for the addition of poly(A) tract at the 3' end of mRNA (16), was located at 16 nucleotides upstream from the poly(A) tract. The nucleotides flanking the predicted initiation codon ATG at the position of nucleotide 5-7 differed extremely from the Kozak's consensus sequence CCA/GCCATG (17), suggesting that *Paramecium* has an unique sequence flanking the translational initiation site (18). The nucleotide sequence contained four TAA codons in this reading frame at the positions of 32-34, 78-80, 128-130 and 308-310, indicating failure to form an open reading frame in ordinary cases. To encode for amino acids, Glx, however, *Paramecium* has been reported to use two codons (TAA and TAG) out of three, which act universally as the termination signals in most of eukaryotes (19,20). Hereupon, *Paramecium* can realize the open reading frame extending the nucleotides from 5 to 358, which is deducible a protein composing of 116 amino acids and a molecular mass of 11,948 dalton. If TAA encodes Gln, our deduced amino acid sequence was completely the same with the primary structure which was reported to the Hb on the basis of direct Edman degradation method (6). In contrast to this, translation of the cDNA in *E. coli* should be interrupted by TAA at an early position corresponding to 9 in order of the amino acid sequence. Therefore, it may cause a difficulty of detection of the immunoreactive signal in the screening of λgt11 expression library.

In the open reading frame resolved, G + C content was 48%. This value falls into a range of those obtained from other ciliate genes (21). The codon usage exhibits a strong bias, since only 59% of 63 codons are used. For example, 9 codons out of 11 encoding Gly were GGA, and 8 codons out of 12 encoding Thr were ACA. Such biases of the codon usage was well agreed with those for other *Paramecium* genes summarized by Martindale (21).

Our most unexpected finding was that the *Paramecium* Hb does not share any sequence homology with other globins at the nucleotide level as well as the amino acid level. This fact suggests that the Hb gene in protozoa had been separated from an ancestral globin gene at very early period of the evolution of eukaryotes.

The acquisition of cDNA clones encoding the protozoa globin is expected to provide a powerful tool to study the mechanism involved in regulation of the synthesis of the cognate Hbs, the structure of these gene family and the search of ancestral type of Hbs.

ACKNOWLEDGMENTS

We wish to acknowledge the assistance of Mr. Itaru Hirai to perform this experiment. This work was supported in part by a Grant-in aid from the Ministry of Education, Science and culture (No. 02640553), Japan, to I.U.

REFERENCES

1. Irie, T. and Usuki, I. (1980) Comp. Biochem. Physiol. **67B**, 549-554.
2. Steers, E. J., Barnett, A. and Clifford, E. L. (1981) Comp. Biochem. Physiol. **70B**, 185-191.
3. Usuki, I. and Irie, T. (1983) Comp. Biochem. Physiol. **75B**, 421-424.
4. Usuki, I. Hino, A. and Ochiai, T. (1989) Comp. Biochem. Physiol. **93B**, 555-559.
5. Usuki, I. and Hino, A. (1987) Cell. Mol. Biol. **33**, 601-609.
6. Iwaasa, H., Takagi, T. and Shikama, K. (1989) J. Mol. Biol. **208**, 355-358.
7. Usuki, I. and Irie, T. (1983) Comp. Biochem. Physiol. **75B**, 415-420.
8. Laemmli, U. K. (1970) Nature **227**, 680-685.
9. Burnette, W. N. (1981) Anal. Biochem. **112**, 195-203.
10. Chirgwin, J. M., Przybyla, A. F., MacDonald, R. J. and Rutter, W. F. (1979) Biochemistry **18**, 5294-5299.
11. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. U.S.A. **69**, 1408-1412.
12. Young, R. A. and Davis, R. W. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 1194-1198.
13. Vieira, J. and Messing, J. (1987) Methods Enzymol. **153**, 1-11.
14. Bressan, G, M. and Stanley, K. K. (1987) Nucleic Acids Res. **15**, 10056.
15. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) in Molecular Cloning: A Laboratory Manual. 2nd edt, Cold Spring Harvor Laboratory Press, Cold Spring Harvor.
16. Birnstiel, M. L., Busslimger, M. and Strub, K. (1985) Cell **41**, 349-359.
17. Kozak, M. (1984) Nucleic Acids Res. **12**, 857-872.
18. Yamauchi, K. (1991) Nucleic Acids Res. **19**, 2715-2720.
19. Caron, F. and Meyer, E. (1985) Nature **314**, 185-188.
20. Preer, J. R., Jr., Preer, L. B., Rudman, M. B. and Barnett, A. J. (1985) Nature **314**, 188-190.
21. Martindale, D. W. (1989) J. Protozool. **36**, 29-34.