



The cDNA-Derived Amino Acid Sequence of Chain II of the Heterodimeric Hemoglobin from the Blood Clam *Barbatia virescens*

Authors: Suzuki, Tomohiko, Arita, Tomikazu, and Kawasaki, Yoshitada

Source: Zoological Science, 12(4) : 453-455

Published By: Zoological Society of Japan

URL: <https://doi.org/10.2108/zsj.12.453>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

The cDNA-Derived Amino Acid Sequence of Chain II of the Heterodimeric Hemoglobin from the Blood Clam *Barbatia virescens*

TOMOHIKO SUZUKI, TOMIKAZU ARITA and YOSHITADA KAWASAKI

Department of Biology, Faculty of Science, Kochi University, Kochi 780 Japan

ABSTRACT—The blood clam *Barbatia virescens* expresses a unique heterodimeric hemoglobin consisting of chains I and II in erythrocytes. This is in sharp contrast to the tetrameric ($\alpha_2\beta_2$) and polymeric two-domain hemoglobins of the congeneric species *Barbatia reeveana* and *Barbatia lima*. The 3' and 5' parts of the cDNA of *B. virescens* chain II have been amplified separately by polymerase chain reaction (PCR), and the complete nucleotide sequence of 690 bp was determined. The open reading frame is 477 nucleotides in length and encodes a protein with 158 amino acid residues, of which 120 amino acids were identified directly by the protein sequencing of the peptides obtained from digestions with trypsin, *S. aureus* V8 protease and pepsin. The mature protein begins with the blocked Ser, and thus the N-terminal Met is cleaved away. The molecular mass for the protein was calculated to be 17605 Da. The cDNA-derived amino acid sequence of *B. virescens* heterodimeric chain II shows the highest homology (42%) with that of *B. virescens* chain I, but shows lower homology (32–35%) with those of tetrameric α and β chains of *B. lima*. This indicates that *B. virescens* chains I and II do not correspond to *B. lima* α and β chains, namely the heterodimeric hemoglobin is a unique gene product expressed only in *B. virescens*.

INTRODUCTION

The bivalves belonging to the family Arcidae contain abundant hemoglobins in circulating erythrocytes [9], but their compositions differ remarkably in different clams. For example, *Anadara trapezia* contains homodimeric (γ_2) and tetrameric ($\alpha_2\beta_2$) hemoglobins [1], *Barbatia reeveana* contains tetramer ($\alpha_2\beta_2$) and polymer consisting of unusual 34 kDa two-domain chain (2D) [3], and the congeneric clam *Barbatia virescens* has only a heterodimeric hemoglobin (I-II) [8]. In addition, we found that *Barbatia lima* subsp. collected from Amami-Oshima, Japan, expresses three types of hemoglobins: homodimer (δ_2), tetramer ($\alpha_2\beta_2$) and polymer of 2D and δ chains (Suzuki *et al.*, in preparation). One of the most remarkable feature of *Anadara* (*Scapharca*) homodimeric and tetrameric hemoglobins is that the subunit assembly is “back-to-front” relative to vertebrate hemoglobin [4]. This means that the tetrameric assembly has been acquired independently in vertebrates and molluscs.

As a first step to make clear the complex molecular evolution of clam hemoglobins, we are analyzing the primary structure of each constituent chain of hemoglobins of *B. virescens* and *B. lima*. Here we report the cDNA-derived amino acid sequence of chain II of *B. virescens* heterodimeric hemoglobin.

MATERIALS AND METHODS

The chain II of the hemoglobin of *Barbatia virescens* was isolated

as described previously [8]. *Barbatia* chain II was carboxymethylated and digested with trypsin, pepsin and *S. aureus* V-8 protease, respectively [6], and the peptides were separated by a reverse-phase chromatography. The column (Cosmosil 5C₁₈-300, 2.5 × 150 mm, Nacalai Tesque) was eluted with a linear gradient of acetonitrile in 0.1% trifluoroacetic acid (TFA) at a flow rate of 1 ml/min. Some peptides were purified further by rechromatography. Peptides were sequenced with the manual Edman method [6]. The peptic peptide P-3 with amino acid composition of Ser₁, Glu₁, Pro₁, Ala₃, Ile₁ and Lys₁ was further digested with acylaminoacid releasing enzyme (Takara), before sequencing [8].

mRNA was isolated from the erythrocytes of *B. virescens* with a FastTrack mRNA Isolation Kit (Invitrogen). The single stranded cDNA was synthesized with avian reverse transcriptase using oligo-dT adaptor of 5'GGATCCGAATTCCTCCGGGT₁₇ as a primer. The 3' half of the cDNA of the chain II was first amplified for 30 cycles, each consisting of 0.5 min at 94°C for denaturation, 0.5 min at 45°C for annealing and 1 min at 72°C for primer extension, by polymerase chain reaction (PCR) [5]. As a enzyme, Taq DNA polymerase (Promega) was used. The primers used are the oligo-dT adaptor described above and the redundant oligomer α (20 mer, a mixture of 32) of 5'GC(AC)AT(CA)TA(CT)CT(AC)ATGTA(TC)GC, based on the amino acid sequence of Ala-Ile-Tyr-Leu-Met-Tyr-Ala. Codon usage for Ala, Ile and Leu of *B. virescens* chain I was used to reduce redundancy. The 500 bp products thus amplified were subcloned in the *Sma*I site of pUC18 and sequenced by dideoxy chain termination method with Sequenase Ver.2.0 (United States Biochemical), BcaBEST DNA sequencing Kit (Takara) and non-RI Uniplex DNA sequencing Kit (Millipore).

The 5' half of the cDNA was amplified as follows. The single-stranded cDNA was newly synthesized with the non-redundant primer *b* (24 mer): 5'CTCTTTAAATGTCATATGCAAAG. This is complementary to the sequence 364–387 shown in Figure 1, and the poly A tail was added to the 3' end with a terminal deoxynucleotidyl transferase. Then the 5' half of the cDNA was

Accepted April 28, 1995

Received March 14, 1995

amplified for 30 cycles, each consisting of 0.5 min at 94°C for denaturation, 0.5 min at 55°C for annealing and 1 min at 72°C for primer extension, using the oligo-dT adaptor and the non-redundant primer *c* (24 mer) of 5'CAATATCTTCTACGATGCTAGGGT, complementary to the sequence 333–356 shown in Figure 1. The 350 bp products thus amplified were subcloned in the *Sma*I site of pUC18 and sequenced.

RESULTS AND DISCUSSION

We have succeeded in amplifying the cDNA encoding *B. virescens* chain II by PCR. The complete nucleotide sequence of 690 bp was constructed by two PCR fragments amplified separately (Fig. 1). There was no sequence discrepancy in the overlapping region. The open reading frame is 477 nucleotides in length and encodes a protein with 158 amino acid residues, of which 120 amino acids underlined in Figure 1 were identified directly by the protein sequencing and/or amino acid analyses of the peptides obtained from digestions with trypsin, *S. aureus* V8 protease and pepsin.

No N-terminal amino acid residue of the intact chain II was detected by Edman sequencing, suggesting that the N-terminus is blocked, as in the case of the two-domain chain of *B. lima* [8]. So we isolated the N-terminal peptic peptide P-2, digested with acylamino acid releasing enzyme and sequenced. Thus we confirmed that the initiation Met is

cleaved away, and the mature protein begins with the blocked Ser. The molecular mass for the protein was calculated to be 17605 Da.

The cDNA-derived amino acid sequence of *B. virescens* heterodimeric chain II was aligned with those of *B. virescens* chain I [7] and *B. lima* subsp. tetrameric $\alpha\beta$ chains (Suzuki *et al.*, in preparation) in Figure 2, using the algorithm of Feng and Doolittle [2]. The percent identity between the chains is summarized in Table 1. The chain II shows the highest homology (42%) with chain I, but shows lower homology (32–35%) with each of the α and β chains of *B. lima*. This sequence homologies indicate that *B. virescens* chains I and II do not correspond to *B. lima* α and β chains. Namely the heterodimeric hemoglobin is a unique gene product expressed only in *B. virescens*.

In *Anadara* (*Scapharca*) tetrameric and homodimeric hemoglobins, E and F helices are shown to participate in important inter-subunit interactions [4]. Consistent with this, three constituent chains of *Anadara* hemoglobins are highly-homologous especially in E and F helices: 24 out of 30 residues are identical as shown in Figure 2. The overall similarities between *Anadara* chains are 48–55%. If we compare the amino acid sequences of *B. virescens* chains I and II with that of *Anadara* consensus sequence [1], only 13–14 residues (underlined in Fig. 2) are conserved. Thus it is unlikely that *B. virescens* heterodimeric hemoglobin forms a

GCCTTACTCCCTTTCTGAGAAAGACACTTATTTAGATCACCGCTGCAGAATTAGAGCCTAAGAATGAGTAAACC	75
M S K P	4
AGCTGAAGCCATTGCTGCTGTGACACAGCCAGACGTCAAAGCAGCACTCAAGTCCTCATGGGGTCTACTGGCTCC	150
A E A I A A V T Q P D V K A A L K S S W G L L A P	29
CAATAAAAAGAAGTATGGTGTGAGTTAATGTGCAAGTTATTTTCATTGCACAAAGATACCGCTCGATACTTCGA	225
N K K K Y G V E L M C K L F S L H K D T A R Y F E	54
ACGGATGGGGAAGCTTGATGCAAAACAATGCAGGAAGTAACCGAGAGTTGATCGGACACGCTATTTACCTTATGTA	300
R M G K L D A N N A G S N R E L I G H A I Y L M Y	79
→ TGCAATTGAATCATTTGTGGATCAACTTGACGACCCCTAGCATCGTAGAAGATATTGGTCGTAACCTTTGCATATAG	375
A I E S F V D Q L D D P S I V E D I G R N F A Y R	104
ACATTTAAAGAGGGGTATTGGAAGCAAATCATTTTCGGTGATCCTCGACACACTTGATCAATTTCTCGGAGATTC	450
H L K R G I G S K S F S V I L D T L D Q F L G D S	124
ATTAGGAGTTAACTACACCGCAGAAAGTTAAAGATGCATGGGAGAAACTTGTTAAAGTTATCTTGGCCCTTCTTGA	525
L G V N Y T A E V K D A W E K L V K V I L A L L D	154
TGATGAGAAAAAGTGAATGAAATGACATTAAATAATACAAATGTTTCATAACATAAGAAACATTGTCTGCAAAAT	600
D E K K *	158
GCTCTCTTCTCTGCGCAAAACTGCAAACTTTTAAATATCTTATTGTTAATTTATGTGTTCTTGTCTCTTGA	675
AAAAATTTTAAGTTG	690

FIG. 1. Nucleotide and derived amino acid sequences of cDNA of chain II of *B. virescens* hemoglobin. Arrows *a* and *b* indicate the primers used in the PCR amplification. Polyadenylation signal (AATAAA) are boxed. The amino acid residues determined by protein sequencing are underlined.

	1		60
Bv.I	PSVAAVSAVT	NKDVAQEIKDTWPFLKKDMVANGISLMIKMFECTPESLEF	
Bv.II	SKPAEATAAVT	QPDVKAALKSSWGLLAPNKKKYGVELMCKLFSLHKDTARY	
Bl.β	MSFEEAALEVTGSEKIKEDLRLTWGILSNELEDTGVTLMTLFLKMEPGSKAR		
Bl.α	MADEEQIAKAVFKEAKALCDSDAAKENLRSSFILANDIESSGLQLMKNLFTSHPDTLGY		
		* * *	
	61		120
Bv.I	FKRMGDL	SNPKANRKLRGHGSSLMYGIMAFVEQLDPSDLTDVAGKFAFNHVNRRNISS	
Bv.II	FERMGKLDANNAGSNRELIGHAIYLMYAIESFVDQLDDPSIVEDIGRNFAYRHLLKRGIGS		
Bl.β	FGRFGNI	DSGMGRDKLRGHSITLMYALQNFMDSLDNTEKLRVVDKFAVNHRRIRKISA	
Bl.α	FARFKDV	SLGKQNWQLRGHAITLMYALMNFVDALDEPDRCLKCVLKFVNNHNRISIP	
	* *	* * *	* * *
Anadara consensus	NSKLRGH.ITL.YAL..F.D	VEKFVNH.I.	
	----- E -----		----- F -----
	121		167
Bv.I	FEFQWALVPLLEVLRLRERLGRNRYRQETEEAWTKLVSVIQATLDDRKR		
Bv.II	KSFVILDITLDQFLGDSLGN YTAEVKDAWEKLVKVLALLDDEKK		
Bl.β	SEFGWIMKPIREVLMERMG QFYDPSFVDAWGKLIGVVQASLAREQ		
Bl.α	QVFGKINGPMDLLKQRMG KYYNRETANAWKQLVGVVQAALTTPN		
	*	* *	* *

FIG. 2. Alignment of the cDNA-derived amino acid sequence of *B. virescens* heterodimeric chain II with those of *B. virescens* chain I, *B. lima* tetrameric α and β chains, and with the *Anadara* consensus sequences in E and F helices. Invariable residues conserved in all chains are marked by asterisks. This alignment was obtained with the algorithm of Feng and Doolittle [2].

TABLE 1. Matrix for sequence homologies (percent identity) between the polypeptide chains of *B. virescens* and *B. lima* hemoglobins

	Bv-II	Bl- α	Bl- β
Bv-I	41.9	36.4	36.4
Bv-II		34.6	32.0
Bl- α			43.2

Keys: Bv-I and II, *Barbatia virescens* chains I and II; Bl- α and β , *Barbatia lima* α and β chains.

“back-to-front” subunit assembly as in the case of *Anadara* hemoglobins.

Why are the different hemoglobins expressed in closely related species *B. virescens* and *B. lima*? We suggest that this is due to the reason that the blood clam hemoglobin plays physiologically less important role compared with that of vertebrate hemoglobins. Most of bivalves do not express hemoglobins [9], but they must have retained hemoglobin genes probably as a pseudo-form. Thus, various type of clam hemoglobins, such as those of homodimer, heterodimer, tetramer, two-domain globin and extracellular 18-20-domain globin [9], would be recognized as relics of molecular evolution.

REFERENCES

- Como PF, Thompson EOP (1980) Multiple hemoglobins of the bivalve mollusc *Anadara trapezia*. Aust J Biol Sci 33: 643–652
- Feng DF, Doolittle RF (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic tree. J Mol Evol 25: 351–360
- Naito Y, Riggs CK, Vandergon TL, Riggs AF (1992) Origin of a “bridge” intron in the gene for a two-domain globin. Proc Natl Acad Sci USA 88: 6672–6676
- Royer Jr WE, Love WE, Fenderson FF (1985) Cooperative dimeric and tetrameric clam haemoglobins are novel assemblages of myoglobin folds. Nature 316: 277–280
- Saiki RK, Gelfand D, Stoffel S, Scharf S, Higuchi R, Horn G, Mullis K, Erlich H (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491
- Suzuki T (1986) Amino acid sequence of myoglobin from the mollusc *Dolabella auricularia*. J Biol Chem 261: 3692–3699
- Suzuki T, Nakamura A, Satoh Y, Inai C, Furukohri T, Arita T (1992) Primary structure of chain I of the heterodimeric hemoglobin from the blood clam *Barbatia virescens*. J Protein Chem 11: 629–633
- Suzuki T, Shiba M, Furukohri T, Kobayashi M (1989) Hemoglobins from the two closely related clams *Barbatia lima* and *Barbatia virescens*. Comparison of their subunit structure and N-terminal sequence of the unusual two-domain chain. Zool Sci 6: 269–281
- Terwilliger RC, Terwilliger NB (1985) Molluscan hemoglobins. Comp Biochem Physiol 81B: 255–261