Variant Subunit Specificity in the Quaternary Structure of *Artemia* Hemoglobin

Cassandra J. Vandenberg, Charles M. Matthews, and Clive N. A. Trotman

Department of Biochemistry, University of Otago, Dunedin, New Zealand

The brine shrimp *Artemia* has three extracellular hemoglobins (Hbs) that are developmentally expressed and exhibit distinct oxygen-binding characteristics (Heip, Moens, and Kondo 1978; Heip et al. 1978). These Hbs are composed of two polymers, each of which comprises nine covalently linked globin domains. Although the cDNA sequences of two nine-domain globins from *Artemia* have been published, there is evidence for the existence of further expressed globin genes (Manning, Trotman, and Tate 1990). In the present study extensive analysis at the cDNA and genomic levels was performed in order to determine the globin gene copy number in *Artemia*. Sequence and Southern analysis suggest that four Hb polymers (T1, T2, C1, and C2) are expressed in *Artemia*. In addition, there is also at least one globin pseudogene. Protein sequencing of the native Hbs revealed that there are limitations on which two polymers can associate. The composition of the Hbs has been determined to be: Hb I, C1C2; Hb II, C1T2; and Hb III, T1T2. These pairings allow the levels of the three *Artemia* Hbs to be regulated independently by polymer expression alone, therefore explaining the previously inconsistent developmental and hypoxia-induced expression patterns.

Introduction

The *Artemia* hemoglobin (Hb) molecule is a large dimer of M_r 320,000, containing a total of 18 globin domains. Three different dimers occur physiologically (Heip et al. 1978) and can be separated electrophoretically (Moens, 1982). The heterodimer form (Hb II) contains subunits known as the T (translated cDNA) and C (corresponding) polymers. These differ by 15.5% at the DNA level and 11.7% at the amino acid level (Matthews, Vandenberg, and Trotman 1998). The two homodimers, CC (Hb I) and TT (Hb III), are also present physiologically. Each single polymer is a concatenation of nine globin domains differing individually in length but exactly in register in the alignment between polymers, including a number of idiosyncrasies such as a unique deletion of one amino acid between residues C1 and C4 in the ninth domain. Logically, the nine-domain concatenation arose from successive gene fusion events (Jellie, Tate, and Trotman 1996; Matthews, Vandenberg, and Trotman 1998). Duplication of an ancestral nine-domain polymer leading to the present T and C forms happened about 60 MYA and evidently later than 85 MYA because *Parartemia zietziana* has the gene for a single nine-domain globin (Coleman, Matthews, and Trotman 2001).

Hb constitutes up to 2% of the *Artemia* hemolymph protein (Gilchrist 1955). The three dimers have different oxygen affinities and kinetic properties that may be of physiological importance through differential expression in response to salinity and oxygen tension (Heip et al. 1978; Decleir et al. 1980; Spicer and El-Gamal 1999). How differential expression is regulated is not transparent because the distribution of T and C subunits among the three polymers present in the hemolymph often does not approximate the obvious 1:2:1 ratio of TT:CT:CC.

The present work describes the finding that each of the subunits, T and C, is coded on replicate but slightly variant genes so that different forms T1, T2, and C1, C2 are expressed. However, the variants are not freely interchangeable in the formation of dimers, and the three physiologically recognized Hbs, Hb I, Hb II, and Hb III, are composed of, respectively, C1C2, C1T2, and T1T2. This provides for an additional level of regulation by enabling, for instance, the expression of T1 to control the level of Hb III, whereas the expression of T2 could control the level of Hb II.

The aligned amino acid sequences (cDNA translations) of *Artemia* C1 and T1 globins have been published previously (Matthews, Vandenberg, and Trotman 1998), and their alignment with *Parartemia* globin has been published in the Molecular Biology and Evolution website (see Coleman, Matthews, and Trotman 2001).

Materials and Methods

Library Screening

Artemia cDNA (Matthews, Vandenberg, and Trotman 1998) and genomic libraries (Matthews and Trotman 1998) were screened for globin clones by hybridization with C1 and T1 polymer cDNA probes as described previously (Matthews and Trotman 1998).

DNA Sequencing

Polymer-specific primers were used to amplify selected regions of positive clones using the polymerase chain reaction (PCR). All sequencing was performed on uncloned PCR products after purification using a QIAquick PCR purification kit (QIAGEN). Sequencing was performed using dRhodamine terminator chemistry and an ABI Prism 377 automated sequencer.

Southern Analysis

Genomic DNA (10 μg) was digested with *Hin*dIII and fractionated by electrophoresis at 30 V for 12 h on

Key words: hemoglobin, *Artemia*, protein sequence, protein structure, evolution.

Address for correspondence and reprints: Clive N. A. Trotman, Biochemistry Department, University of Otago, Box 56, Dunedin, New Zealand. E-mail: clive.trotman@stonebow.otago.ac.nz.

Mol. Biol. Evol. 19(8):1288–1291. 2002

a 1% agarose, TAE gel. Before transfer, agarose gels were pretreated (depurination, denaturation, and neutralization) according to the instructions provided with the Hybond-N membrane (Amersham). The DNA was then Southern blotted to the membrane and fixed by baking at 80°C for 12 h. The Southern blots were hybridized with radiolabeled PCR products at 60°C for ≥12 h in 4× SSPE, 0.1% tetrasodium pyrophosphate, 0.5 mg/ml heparin, and 0.5% SDS. Membranes were washed twice with 2× SSPE, 0.5% SDS for 10 min each, then with 1× SSPE, 0.1% SDS for 15 min. Before a second round of hybridization, the membranes were stripped by pouring boiling 0.1% SDS onto the Hybond-N membrane and allowing the solution to cool to room temperature. This step was repeated if necessary.

Artemia Hb II

Artemia Hb II was kindly provided by C. Marshall (Department of Biochemistry, University of Otago). This Hb had been purified by elution from nondissociating gels as detailed previously (Marshall et al. 1986).

Cleavage of Hb II

Cleavage reactions were performed in siliconized tubes. *Artemia* Hb II (100 μg) in 70% formic acid was incubated at 37°C for 20 h. The reaction mixture was dried in a Speed Vac, resuspended in SDS-PAGE loading buffer, and neutralized by adding 1 M Tris base until the sample turned from yellow to blue. The cleavage products were then separated by SDS-PAGE (Kolbe et al. 1984).

Protein Sequencing

The protein to be sequenced was run on an SDS-PAGE gel, transferred to the PVDF membrane, stained with Coomassie Blue (0.025% Coomassie Blue in 40% methanol) for 30–60 s, and destained (50% methanol) for 2 h. The protein bands were excised and amino-terminals sequenced using an Applied Biosystems gas phase peptide sequencer.

Protein and DNA Sequence Compilation and Analysis

Translation functions were performed using the program NLDNA (Stockwell 1987). The homology of new DNA sequences was determined using HOMED (Stockwell 1988) or GCG (Genetics Computer Group Inc.) sequence analysis package, version 7.3.

Results and Discussion

Globin Gene Copy Number

The sequence analysis of globin cDNA clones has confirmed the existence of two distinct classes of the T polymer clone. The original globin cDNA clones isolated had been grouped into two classes based on the restriction enzyme digest analysis (Manning et al. 1990); however, the sequence of only one clone had been determined (Manning, Trotman, and Tate 1990). The sequencing of a clone belonging to the other class has shown that the two classes of T polymer, which have been named T1 and T2, differ by approximately 3% at the coding level. Comparison of the sequence of 36 genomic clones (Matthews, 1998) with the cDNA sequences of T1, T2, and C1 (Matthews, Vandenberg, and Trotman 1998) allowed these clones to be grouped into five distinct classes. Three of these corresponded to T1, T2, and C1 polymer genes. A fourth category of clone was designated C2, a previously undescribed C polymer gene. The fifth type of clone appeared to be a pseudogene of the T polymer type. This gene was assigned as a pseudogene based on the position of a stop codon at A6 of domain 9 and the presence of nonconservative substitutions that would not be expected in a functional gene. The most prominent example is the substitution of a glycine at position A12 in domain 9 in place of the highly conserved tryptophan found in all except domain 5, where it is a phenylalanine in polymers T and C.

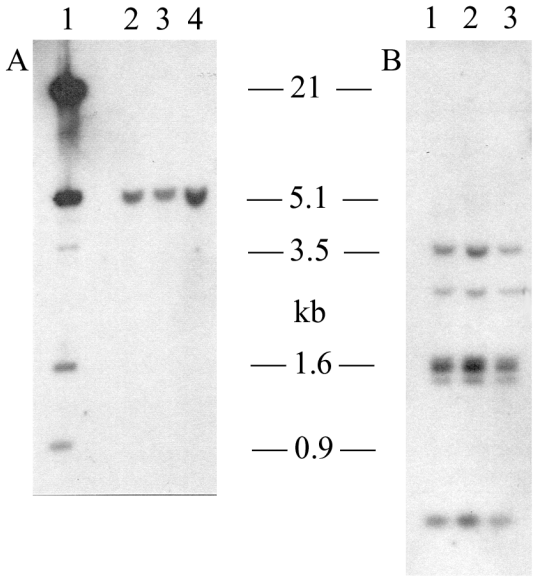

FIG. 1.—The Southern analysis of the *Artemia* globin gene copy number. (*A*) The enzyme-restricted *Artemia* genomic DNA hybridized with a polymer-specific probe (C1 polymer, 1.2 kb intron probe). Lane 1, *Eco* RI-*Hin*dIII–digested lambda DNA; lanes 2–4, 10-μg *Hin*dIII–digested *Artemia* genomic DNA. The single band indicates that digestion is complete. (*B*) The Southern blot in (*A*) was stripped and hybridized with a probe corresponding to the ninth exon of the T1 and C1 polymer genes. Lanes 1–3, 10-μg *Hin*dIII–digested *Artemia* genomic DNA.

The Southern analysis of *Hin*d III–digested genomic DNA produced five well-separated bands (fig. 1) that have been interpreted as corresponding to five genes: T1, T2, C1, C2, and one T polymer pseudogene. The probe corresponded to the ninth exons of both the T1 and C1 polymer genes, which contain no *Hin*dIII sites. The different intensities of the bands presumably reflects the level of mismatch between the probe and target sequence. Thus, it appears that there are five distinct globin genes in *Artemia*, four functional and one nonfunctional. However, there is the possibility that two or more positive genomic fragments of similar size are running at the same position.

Sequence Analysis

It is proposed that duplication of an ancestral nine-domain gene gave rise to the ancestral C and T polymer genes. Further events of gene duplication have occurred more recently to produce the T1, T2 and C1, C2 polymer genes. The alignment of the three sequenced polymers is presented as *Supplementary Material* (see MBE website at www.molbiolevol.org). The overall difference at the amino acid level between the T1 and C1 polymers and the T2 and C1 polymers is 11.7% and 11.2%, respectively. Based on the vertebrate globin divergence calibration (1% per 5 Myr; Dickerson and Geiss 1983, p. 97) the separation of the T and C polymers has been dated at approximately 60 MYA (Matthews, Vandenberg, and Trotman 1998). Similarly, the 3% sequence divergence of T1 and T2 polymers suggests they separated approximately 15 MYA; however, the low resolution of the small degree of divergence limits the reliability of such an estimate. The assumption that the *Artemia* and vertebrate globins have evolved at a similar rate is also debatable, although the remarkable coincidence between the divergence date of *Artemia* and *Parartemia* from a common ancestor using the same calibration (85 MYA), and the geological separation of New Zealand from Australia, lend some support to the calibration. A further confounder is that one of the duplicates may have experienced an accelerated rate of substitution after the gene duplication event (Ohta 1994). From the available C2 sequence the difference between C1 and C2 is about 7.5%, which suggests that C1 and C2 are more distant than T1 and T2, and indeed there is no reason for C and T to have duplicated at the same time.

The coding regions of the T1 and C1 polymer genes show on average 84% identity at the DNA level. However, no obvious homology exits between the intron sequences of these genes. Partial sequences from all 22 introns of the T1 and C1 polymers were compared and the identity found to range from 26% to 38%, with much of the observed match likely to be contributed to by the high AT content of the introns (61%–79%). In contrast, the identity between the T1 and T2 introns is much greater, 84%–95%. At the coding level the T1 and T2 polymers are only about 3% different; therefore, the introns, with an average change of 10.7 substitutions per 100 sites, have a substitution rate about 3.5 times that of the coding sequences.

Protein Sequencing

Sequencing of the native *Artemia* Hbs indicates there are limitations on which two polymers can associate. Amino-terminal sequencing of Hb I showed that it is composed of two forms, one having arginine at the fourth position whereas the other has leucine. These two sequences were present in equal quantities and correspond to the C1 and C2 polymers. Only one sequence is obtained by this method for Hb III because the T1 and T2 polymers are identical for the first 148 residues. Interestingly, amino-terminal sequencing of the heterodimer, Hb II, gave only the C1 polymer sequence and the T polymer sequence (as for Hb III). The C2 polymer is not involved in the formation of Hb II. The cleavage of Hb II followed by sequencing of the fragments allowed the T polymer composition to be determined. Of the eight sequences generated five were C1, two were T2, and one was not polymer specific. The assignment of the sequences obtained by Marshall (1988) from cyanogen bromide and trypsin fragments of Hb II gives similar results: four C1, one T2, three not specific. No sequences have been identified that are unique to T1; however, because not all the fragments have been sequenced, the involvement of T1 in Hb II has not been completely ruled out. These results indicate that Hb II is composed of C1 and T2.

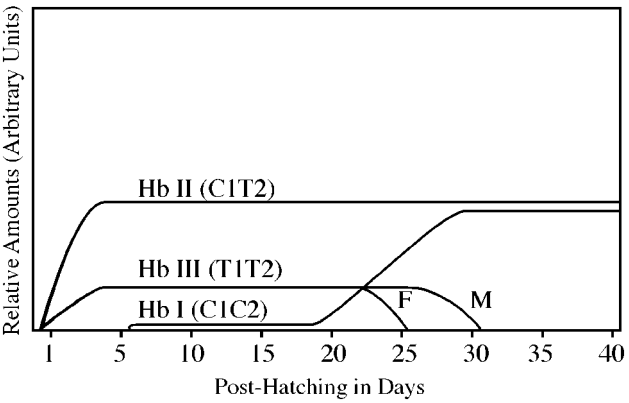

FIG. 2.—The developmental expression pattern of the *Artemia* Hb (adapted from Heip et al. 1978).

The existence of four distinct polymers and their restricted pairing to form the three *Artemia* Hbs may resolve the question of how *Artemia* achieves both the developmental and environmental expression patterns that are observed. If only the polymer combinations C1C2 (Hb I), C1T2 (Hb II), and T1T2 (Hb III) exist, then the Hbs can be independently regulated by polymer expression alone. If the affinities between the two T polymers and between the two C polymers exceed that between C1 and T2, then the level of Hb I can be controlled by the expression of C2 and Hb III by T1. Production of Hb II is then dependent on excess C1 and T2. Under hypoxic conditions all three Hbs are present, which would not be possible if the affinity between C1 and T2 were greater than between T1 and T2 and between C1 and C2. However, we cannot completely rule out the possibility that under some developmental stages other combinations of polymers form. Similarly, we cannot exclude the possibility of the other uncharacterized globin genes being functional in the brine shrimp *Artemia*.

Globin Gene Expression

The developmental expression pattern of *Artemia* Hb, seen in figure 2, can now be described in terms of the induction and inactivation of specific genes. The first genes to be expressed posthatching are C1 and T2, resulting in the appearance of Hb II (Heip et al. 1978). A few hours later T1 is induced, and Hb III can then be detected. Not until 7–9 days posthatching does the induction of C2 occur, which is seen by the Hb I expression. The level of T1,

and therefore Hb III, decreases at 20–25 days because this gene is inactive in adult shrimp under normal conditions. Under hypoxic conditions all the globin genes are upregulated as seen by an increase in the level of all three Hbs (Heip, Moens, and Kondo 1978). The greatest increase observed is for Hb III, which is induced in adult shrimp lacking Hb III, i.e., T1 is induced. Differential up-regulation of Hb polymers by hypoxia has been observed in *Daphnia magna* (Kimura et al. 1999). A cluster of four Hb genes has been identified in *D. magna* with putative hypoxia responsive elements (HREs) located in the intergenic regions (Kimura et al. 1999). The analysis of the genomic sequence of the *Artemia* globin genes is expected also to reveal putative HREs.

Why Four Globin Genes in *Artemia*?

Gene duplication is a powerful mechanism whereby new copies of an old gene may arise (Ohno 1970, pp. 71–82). Duplicate genes can have different fates. An increase in copy number may be driven by a need for rapid production of the gene product, for example in rRNA genes (Ridley 1993, pp. 241–242). In other cases, duplicate genes may diverge from each other in their coding regions and may acquire new functions. It appears that *Artemia* has four Hb genes, not for the purpose of increasing expression (although this may be an additional role during hypoxia) but for functional specialization. The three native Hb molecules produced from these genes clearly have different physiological functions as evidenced by their different oxygen binding characteristics (D'Hondt et al. 1978) and what appears to be tight control over the expression of the amounts and proportions of each.

Supplementary Material

The GenBank accession numbers of the Hb cDNA sequences referred to are: *Parartemia*, AF258616; *Artemia* globin C, AF104216; and *Artemia* globin T, AF104217. The amino acid alignment of the *Artemia* Hb polymers T1, T2, and C1 is provided on the website of Molecular Biology and Evolution. An alignment of the *Parartemia* and *Artemia* T and C Hb amino acid sequences was provided in the website previously (Coleman, Matthews, and Trotman 2001).

Acknowledgment

We are grateful for financial support from the Marsden Fund administered by The Royal Society of New Zealand.

LITERATURE CITED

COLEMAN, M., C. M. MATTHEWS, and C. N. A. TROTMAN. 2001. A multimeric hemoglobin of the Australian brine shrimp *Parartemia*. Mol. Biol. Evol. **18**:570–576.

DECLEIR, W., J. VOS, F. BERNAERTS, and C. VAN DEN BRANDEN. 1980. The respiratory physiology of *Artemia* sp. Pp. 137–145. *in* G. PERSOONE, O. ROELS, and E. JASPERS, eds. The brine shrimp *Artemia*, Vol. 2. Universa Press, Wetteren, Belgium.

D'HONDT, J., L. MOENS, J. HEIP, A. D'HONDT, and M. KONDO. 1978. Oxygen-binding characteristics of three extracellular haemoglobins of *Artemia salina*. Biochem. J. **171**:705–710.

DICKERSON, R. E., and I. GEISS. 1983. Hemoglobin: structure, function, evolution, and pathology. Benjamin Cummings, Menlo Park.

GILCHRIST, B. M. 1955. Haemoglobin in *Artemia*. Proc. R. Soc. B **143**:136–146.

HEIP, J., L. MOENS, M. JONIAU, and M. KONDO. 1978. Ontogenetical studies on extracellular hemoglobins of *Artemia salina*. Dev. Biol. **64**:73–81.

HEIP, J., L. MOENS, and M. KONDO. 1978. Effect of concentrations of salt and oxygen on the synthesis of extracellular hemoglobins during development of *Artemia salina*. Dev. Biol. **63**:247–251.

JELLIE, A. M., W. P. TATE, and C. N. A. TROTMAN. 1996. Evolutionary history of introns in a multidomain globin gene. J. Mol. Evol. **42**:641–647

KIMURA, S., S. TOKISHITA, T. OHTA, M. KOBAYASHI, and H. YAMAGATA. 1999. Heterogeneity and differential expression under hypoxia of two-domain hemoglobin chains in the water flea, *Daphnia magna*. J. Biol. Chem. **274**:10649–10653.

KOLBE, H. V., D. COSTELLO, A. WONG, R. C. LU, and H. WOHLRAB. 1984. Mitochondrial phosphate transport. Large scale isolation and characterization of the phosphate transport protein from beef heart mitochondria. J. Biol. Chem. **259**:9115–9120.

MANNING, A. M., R. J. POWELL, C. N. A. TROTMAN, and W. P. TATE. 1990. Developmental expression and cDNA cloning of globin genes from the brine shrimp, *Artemia*. New Biol. **2**:77–83.

MANNING, A. M., C. N. A. TROTMAN, and W. P. TATE. 1990. Evolution of a polymeric globin in the brine shrimp *Artemia*. Nature **348**:653–656.

MARSHALL, C. J. 1988. Structures of the haemoglobins of the brine shrimp, *Artemia*: an examination of their properties and synthesis. Doctoral dissertation, University of Otago, Dunedin, New Zealand.

MARSHALL C. J., J. F. CUTFIELD, C. N. A. TROTMAN, and W. P. TATE. 1986. Purification of the haemoglobins I and III from the brine shrimp *Artemia*. Biochem. Int. **12**:693–700.

MATTHEWS, C. M. 1998. Evolutionary changes in the haemoglobin genes of *Artemia*. Doctoral thesis, University of Otago, Dunedin, New Zealand.

MATTHEWS, C. M., and C. N. A. TROTMAN. 1998. Ancient and recent intron stability in the *Artemia* hemoglobin gene. J. Mol. Evol. **47**:763–771.

MATTHEWS, C. M., C. J. VANDENBERG, and C. N. A. TROTMAN. 1998. Variable substitution rates of the 18 domain sequences in *Artemia* hemoglobin. J. Mol. Evol. **46**:729–733.

MOENS, L. 1982. The extracellular haemoglobins of *Artemia* sp: a biochemical and ontological study. Meded. K. Acad. Wet. Lett. Schone Kunsten Belg. Kl. Wet. **44**:1–21.

OHNO, S. 1970. Evolution by gene duplication. Springer-Verlag, Berlin.

OHTA, T. 1994. Further examples of evolution by gene duplication revealed through DNA sequence comparisons. Genetics **138**:1331–1337.

RIDLEY, M. 1993. Evolution. Blackwell Scientific, Massachusetts.

SPICER, J. I., and M. M. EL-GAMAL. 1999. Hypoxia accelerates the development of respiratory regulation in brine shrimp—but at a cost. J. Exp. Biol. **202**:3637–3646.

STOCKWELL, P. A. 1987. DNA sequence analysis software. Pp. 19–45. *in* C. T. RAWLINGS and M. J. BISHOP, eds. Nucleic acid and protein sequence analysis, a practical approach. IRL Press, Washington, D.C.

———. 1988. HOMED: a homologous sequence editor. Trends Biochem. Sci. **13**:322–324.

WILLIAM TAYLOR, reviewing editor

Accepted March 25, 2002