Volume 259, number 1, 133-136 FEB 07939 December 1989

Scapharca hemoglobins, type cases of a novel mode of chain assembly and heme-heme communication

Amino acid sequence and subunit interactions of the tetrameric component

Raffaele Petruzzelli#, Alberto Boffi, Donatella Barra, Francesco Bossa, Franca Ascoli* and Emilia Chiancone

CNR Center of Molecular Biology, Department of Biochemical Sciences, University La Sapienza, 00185 Rome, #Department of Biology and **Department of Experimental Medicine and Biochemical Sciences, University Tor Vergata, 00173 Rome, Italy*

Received 1 November 1989

The sequences of the A and B chains of the *Scapharca inaequivalvis* tetrameric hemoglobin (HbII) are reported. They are homologous to the corresponding chains of other Arcid hemoglobins. Moreover, a comparison of the present data with the sequence of the *S. inaequivalvis* dimeric hemoglobin (HbI), for which high-resolution X-ray data are available, allows the identification of the residues that direct the assembly of the two components and of those that may be involved in the interdimeric contacts.

Molluscan hemoglobin; Amino acid sequence; Hemoglobin assembly

1. INTRODUCTION

Recent studies on the dimeric (HbI) and tetrameric (HbII) hemoglobins contained in the erythrocytes of Arcid molluscs (*Anadara broughtonii, A. satowi, A. trapezia, A. senilis, Scapharca inaequivalvis*) have shown that they display remarkably constant structural and functional properties with unique features with respect to the vertebrate proteins [1-5]. Arcid hemoglobins are assembled from three distinct polypeptide chains characterized by the myoglobin fold; two of them, IIA and IIB, associate into HbII, a heterotetramer, while the third, chain I, forms HbI, a homodimer. It is in the assembly of the polypeptide chains that Arcid hemoglobins are unique: thus, low resolution X-ray data on the two *S. inaequivalvis* hemoglobins have shown that the heme-carrying E and F helices are not on the outside of the molecule as in vertebrate hemoglobins, but form the intersubunit contact in HbI and in the structural unit of HbII, a dimer of heterodimers [6]. The unique assembly of *S. inaequivalvis* hemoglobins endows these molecules with novel functional characteristics in that it allows cooperativity in ligand binding to be displayed also by the homodimeric HbI [3,5]. This fact represents an important exception to the rule based on the behavior of human hemoglobin that cooperativity in hemoglobins requires the presence of a tetrameric molecule built from unlike chains [7].

The present paper reports the primary structure of the A and B chains of HbII from *S. inaequivalvis*. These are compared with the sequences of dimeric and tetrameric Arcid hemoglobins and are discussed in the light of the very recent high resolution crystal structure of carbonmonoxy *S. inaequivalvis* HbI [8]. The analysis brings out distinctive features of the chains that are likely to direct the assembly of the two hemoglobin components.

Correspondence address: E. Chiancone, Dipartimento di Scienze Biochimiche, Universita, La Sapienza, P. le Aldo Moro 5, 00185 Rome, Italy

The nucleotide sequence(s) presented here has (have) been submitted to the EMBL/GenBank database under the accession number no. Y07524.

2. MATERIALS AND METHODS

The purification of *S. inaequivalvis* HbII was performed according to [5]; globin was prepared by the acid-acetone method [9]. The two globin chains were separated by chromatography on CM-cellulose in the presence of 8 M urea and 2-mercaptoethanol as described in [4].

Trypsin (code TRTPCK) and chymotrypsin (code CDI) were from Worthington Biochemical, HPLC-grade acetonitrile from Farmitalia Carlo Erba and sequence-grade reagents from Applied Biosystems.

S-Carboxymethylated globin chains were digested with trypsin and chymotrypsin under the conditions described in [10]. Peptide mixtures were separated by reverse phase HPLC [10]. Amino acid analyses and manual dansyl-Edman degradations were performed as in [10]. Automated Edman degradation was carried out using an Applied Biosystems model 470A gas phase sequencer equipped with an Applied Biosystems model 120A PTH-analyzer. Cleavage of the Asp-Pro bond was performed according to [11]. The sequence of the chymotryptic peptide comprising the first 26 amino acids of the IIB chain was determined by automated Edman degradation after deblocking of its amino terminus as described in [10].

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/89/$3.50 © 1989 Federation of European Biochemical Societies

3. RESULTS AND DISCUSSION

The complete amino acid sequence of the IIA and IIB chains of *S. inaequivalvis* is presented in figs 1 and 2, respectively, along with the corresponding sequences of other Arcid hemoglobins, namely *A. trapezia* [12,13] and *A. broughtonii* [14] and with the sequence of *S. inaequivalvis* HbI [15], to which the sequence numbering is referred throughout this work. The sequence of the *S. inaequivalvis* IIA and IIB chains was deduced from the analysis of the complete set of tryptic peptides; overlaps were obtained by means of chymotryptic peptides. Cleavage of the IIA chain at the unique Asp-Pro bond provided sequence information from residue 87 to residue 126, with no interference from the N-terminal fragment of the two chains which remained blocked. At any rate, the high similarity with the corresponding chains of *A. trapezia* and *A. broughtonii* renders the need for overlapping sequences less stringent. The sequence of the N-terminal tryptic peptide of the IIA chain was inferred from the amino acid composition by homology with the *A. trapezia* sequence. The N-terminal sequence of the IIB chain was determined on the relevant chymotryptic peptide after deblocking of its amino terminus [10]. The nature of the blocking groups was not determined.

The IIA chain is composed of 149 residues and contains only two histidines, the heme ligands, and two cysteine residues, one of which, located at the beginning of the F helix, is conserved in all Arcid hemoglobin chains. The IIB chain has 151 residues; it contains two histidines, the heme ligands, and the conserved cysteine.

Inspection of the available sequences of the three Arcid chains given in figs 1 and 2, namely chain I, which forms the homodimeric HbI and chains IIA and IIB which build the tetrameric HbII, allows the following phylogenetic considerations. A pairwise comparison of the amino acid residue positions brings out, on the one hand, the species relationships, *S. inaequivalvis* appearing more distantly related to *A. trapezia* (percentage identity: 80.0 for the IIA chains and 92.1 for the IIB chains) than to *A. broughtonii* (percentage identity 93.3 for the IIA chains), and on the other hand, that the two tetramer chains are more closely related to each other (64.3% identity) than to the dimer one (about 52% identity for either tetramer chain). The latter observation suggests that in the evolution of the Arcid globin tree, a first gene duplication gave rise to the dimer chain, while the two tetramer chains originated later upon a second gene duplication.

In view of the unique assembly of Arcid hemoglobins, it is perhaps more interesting to combine the sequence data with the three-dimensional structural information on the two *S. inaequivalvis* hemoglobins [6] for inferring the elements in the primary structure that direct the assembly of the two hemoglobin components. To this end, the first set of residues to be considered is that forming the subunit interface in the dimeric protein, as revealed by the recently solved high resolution structure of carbonmonoxy *S. inaequivalvis* HbI [8]. In HbI, the subunit interface comprises 10 amino acid residues belonging to the E and F helices and, in addition, one residue belonging to the B helix, Lys 30, which forms a salt bridge with Asp 89 from the neighboring subunit. In this region of the molecule about one third of the residues is changed in the tetramer chains (table 1). The substitutions do not involve the two residues, Lys 96 and Asn 100, that are in direct contact with the heme propionates of the neighboring subunit and provide the novel means of heme-heme communication first described in *S. inaequivalvis* HbI [8]. Moreover, the changes that do occur in the tetramer chains with respect to the dimer one are such as to allow the formation of similar interfaces in HbI and in the IIA-IIB dimeric building module of the tetrameric protein. In this respect, the substitution of Lys 30 with a residue of different charge, Asp in both

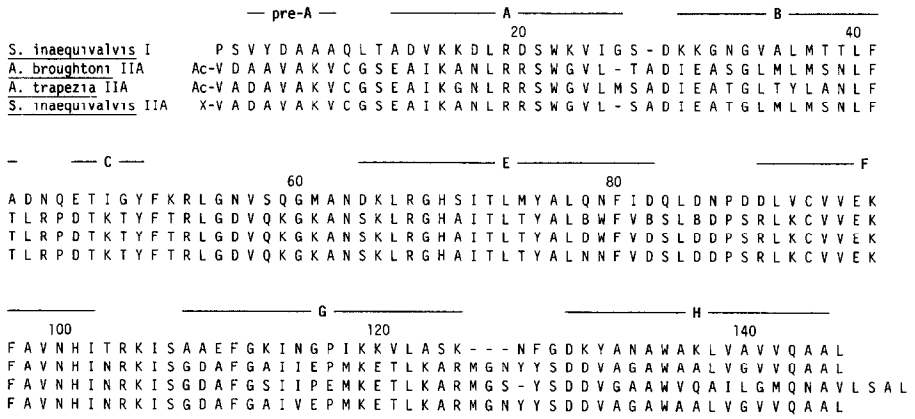

Fig.1. Comparison of the amino acid sequence of the *S. inaequivalvis* IIA chain with those of the same chain in *A. trapezia* [12] and *A. broughtonii* [14]. The numbering and the helix notation refer to the sequence of *S. inaequivalvis* homodimeric HbI [15], which is likewise reported.

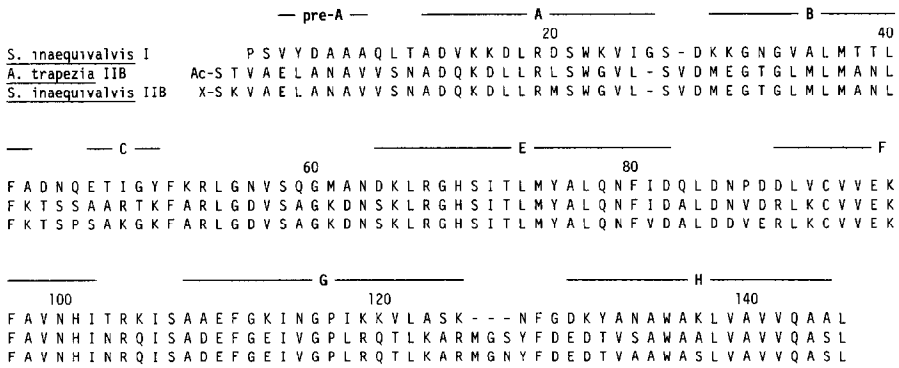

Fig.2. Comparison of the amino acid sequence of the *S. inaequivalvis* IIB chain with that of the same chain in *A. trapezia* [13]. The numbering and the helix notation refer to the sequence of *S. inaequivalvis* homodimeric HbI [15], which is likewise reported.

tetramer chains, may appear surprising. However, it is accompanied by a second compensating substitution at position 89, which is occupied by an Arg residue; thus in HbII a salt bridge can still be formed, but with a sign reversal, a difference that is likely to play a major role in preventing the hybridization of dimer and tetramer chains. At the EF corner, in addition to residue 89, residue 88 is substituted as well in both tetramer chains; these changes are likely to alter the local tertiary structure slightly and to result in the small differences in the microenvironment of the proximal histidine (His 101) revealed by the resonance position of the $N_\delta H$ group [16].

Table 1

Amino acid residues involved in contacts across the interface of homodimeric *S. inaequivalvis* HbI [8] in comparison with the residues occupying the same positions in the A and B chains of Arcid tetrameric HbII

	Globin Chain		
Position	I	IIA	IIB
30	**Lys**	**Asp**	**Asp**
53	Arg	Arg	Arg
64	**Asp**	**Ser**	**Ser**
67	Arg	Arg	Arg
68	Gly	Gly	Gly
71	Ile	Ile	Ile
72	Thr	Thr	Thr
75	Tyr	Tyr	Tyr
78	**Gln**	**Asn, Asp**	**Gln**
79	**Asn**	**Asn, Trp**	**Asn**
82	Asp	Asp	Asp
83	**Gln**	**Ser**	**Ala**
88	**Asp**	**Ser**	**Glu, Asp**
89	**Asp**	**Arg**	**Arg**
92	Cys	Cys	Cys
93	Val	Val	Val
96	Lys	Lys	Lys
97	Phe	Phe	Phe
100	Asn	Asn	Asn

Variable residues are boldfaced. Residues 30 and 89, that are involved in a salt bridge at the interface of HbI [8], display a sign reversal in both tetramer chains (the implications of these substitutions in terms of the assembly process are discussed in the text)

The other regions of the chains to be considered in the framework of the assembly of the two hemoglobin components are those involved in the interdimer contact in HbII; on the basis of the low resolution structure [6], these regions comprise the A helix and the AB and GH corners. All the Arcid tetramer chains sequenced thus far share a three-residue insertion at the level of the GH corner that may be involved in the interdimer contacts; in addition, in the tetramer chains there are several substitutions at the level of the A helix, some of which result in a sign reversal (e.g. 12 Ala →Glu in the IIA chains; 16 Lys→Ala or Asp, 17 Glu→Asn or Leu, 20 Asp→Arg or Met in the IIA and IIB chains, respectively; 23 Lys→Gly in both tetramer chains). However, the relative importance of these changes in the assembly process cannot be established in the absence of a high-resolution structure of HbII.

In conclusion, consideration of the sequences of Arcid globins in evolutionary terms has provided information concerning the phylogenetic relationships among the three different globin chains and among the various species considered. From a structural point of view, the availability of the high-resolution structure of the dimeric protein allowed us to pinpoint changes in amino acid residues that are located at the dimer interface and are likely to direct the assembly of the two hemoglobin components.

REFERENCES

[1] Ohnoki, S., Mitomi, Y., Hata, R. and Satake, K. (1973) J. Biochem. (Tokyo) 73, 717–725.
[2] Furuta, H., Ohe, M. and Kajita, A. (1977) J. Biochem. (Tokyo) 82, 1723–1730.
[3] Djangmah, J.S., Gabbott, P.A. and Wood, E.J. (1978) Comp. Biochem. Physiol, 60B, 245–250.
[4] Como, P.F. and Thompson, E.O.P. (1980) Aust. J. Biol. Sci. 33, 643–652.
[5] Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F. and Antonini, E. (1981) J. Mol. Biol. 152, 577–592.
[6] Royer, W.E. jr, Love, W.E. and Fenderson, F.F. (1985) Nature 316, 277–280.
[7] Perutz, M.F. (1979) Annu. Rev. Biochem. 48, 327–386.

[8] Royer, W.E. jr, Hendrickson, W.A. and Chiancone, E., J. Biol. Chem., in press.
[9] Verzili, D., Rosato, N., Ascoli, F. and Chiancone, E. (1988) Biochim. Biophys. Acta 954, 108–113.
[10] Petruzzelli, R., Barra, D., Sensi, L., Bossa, F. and Brunori, M. (1989) Biochim. Biophys. Acta 995, 255–258.
[11] Schinina', M.E., De Biase, D., Bossa, F. and Barra, D. (1988) J. Prot. Chem. 7, 284–286.
[12] Como, P.F. and Thompson, E.O.P. (1980) Aust. J. Biol. Sci. 33, 653–664.
[13] Fisher, W.K., Gilbert, A.T. and Thompson, E.O.P. (1984) Aust. J. Biol. Sci. 37, 191–203.
[14] Furuta, H. and Kajita, A. (1986) in: Invertebrate Oxygen Carriers (Linzen, B. ed.) pp. 117–120, Springer, Berlin.
[15] Petruzzelli, R., Goffredo, B.M., Barr, D., Bossa, F., Boffi, A., Verzili, D., Ascoli, F. and Chiancone, E. (1985) FEBS Lett. 184, 328–332.
[16] Inubushi, T., Yonetani, T. and Chiancone, E. (1988) FEBS Lett. 235, 87–92.