Amino acid sequence of the cooperative homodimeric hemoglobin from the mollusc *Scapharca inaequivalvis* and topology of the intersubunit contacts

Raffaele Petruzzelli, Bianca M. Goffredo, Donatella Barra, Francesco Bossa, Alberto Boffi, Daniela Verzili, Franca Ascoli° and Emilia Chiancone*

CNR Center of Molecular Biology, Institute of Biochemistry, University of Rome 'La Sapienza', Piazzale de Aldo Moro 5, 00185 Rome, and °*Department of Cell Biology, University of Camerino, 62032 Camerino (Mc), Italy*

Received 3 April 1985

The dimeric hemoglobin (HbI) from *Scapharca inaequivalvis* is highly homologous to the other known dimeric Arcid hemoglobins. The sequence has a distinctive hydrophobicity profile in the region corresponding to the E and F helices with respect to both the hemoglobin and myoglobin chains from vertebrates due to the presence of several additional hydrophobic residues. The characteristic topology of the E and F helices is conserved in all the known sequences of Arcid hemoglobins including that of the so-called α chain of the tetrameric component from *Anadara trapezia*. The rationale for this conservation lies in the unusual assembly of Arcid hemoglobins where the E and F helices are involved in the interdimeric contact. It is suggested that the extra hydrophobic residues play a major role in the assembly of the basic dimeric unit in these hemoglobins.

Molluscan hemoglobin *Amino acid sequence* *Hemoglobin assembly*

1. INTRODUCTION

The dimeric (HbI) and tetrameric (HbII) hemoglobins contained in the red cells of Arcid molluscs have been the subject of numerous investigations in recent years. The properties of both components are remarkably constant in all the species (*Anadara broughtonii, A. satowi, A. trapezia, A. senilis, Scapharca inaequivalvis*) studied to date. The tetrameric HbII is assembled from two polypeptide chains into an $\alpha_2\beta_2$ structure, while HbI is a homodimer constructed with a third kind of chain. Both proteins have a fairly low oxygen affinity ($p_{1/2} \sim 10$ Torr), bind oxygen in a cooperative fashion and lack the alkaline Bohr effect [1–5]. Upon removal of oxygen HbII polymerizes, while HbI does not undergo ligand-linked association-dissociation reactions [4,6].

The properties of the dimeric HbI just outlined provide the reason for our interest in this protein as the simplest and hence ideal system for the study of cooperative phenomena in hemoglobins. Detailed measurements of ligand binding to *S. inaequivalvis* HbI have shown that cooperativity has the same equilibrium and kinetic basis as in tetrameric vertebrate hemoglobins [7–9]. However, it was suggested that the detailed structural basis for cooperativity in these dimers could differ from that of vertebrate hemoglobins; for example, the breakage of salt bridges cannot be invoked for the molluscan protein given the pH independence of its functional properties. Moreover, a preliminary analysis of the *S. inaequivalvis* HbI amino acid sequence in terms of the distribution of polar and apolar residues in the regions corresponding to the intersubunit contacts in vertebrate hemoglobins indicated that dimers of the type $\alpha_1\beta_1$ or $\alpha_1\beta_2$ would not be very stable [10]. Indeed, ac-

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

cording to the very recent low resolution X-ray data of Royer et al. [11], the two *S. inaequivalvis* hemoglobins appear to be made of myoglobin-like subunits, whose assembly is 'back-to-front' with respect to that of vertebrate hemoglobins. Thus, the E and F helices form an extensive intersubunit contact in the dimeric HbI, while they are exposed to solvent in vertebrate hemoglobins. The tetrameric HbII is a dimer of dimers.

As a contribution to the understanding of the unusual assembly of Arcid hemoglobins this paper reports the complete amino acid sequence of the dimeric *S. inaequivalvis* HbI and a comparison of the hydrophobicity profile of relevant polypeptide chain segments with those of the corresponding regions in vertebrate myoglobin and hemoglobin chains. The comparison brings out that the E and F helices in *S. inaequivalvis* HbI contain several additional hydrophobic residues with respect to vertebrate globins. This topology is shared by all known sequences of Arcid hemoglobins and is likely to dictate the assembly of the basic dimeric unit in these proteins.

2. MATERIALS AND METHODS

The purification of *S. inaequivalvis* HbI and the globin preparation were performed according to [4,12]. A globin sample (10 mg) was carboxymethylated with radioactive iodoacetate as described [13] and then digested with 0.3 mg trypsin (TPCK-treated, Worthington) in 0.1 M ammonium bicarbonate for 5 h at 37°C. Digestion was stopped by acidification with acetic acid followed by lyophilization; peptides were purified by high-performance liquid chromatography (HPLC) on a reverse-phase column (Brownlee Labs, RP-300, 10 μm). The digestion mixture was applied in several aliquots onto the column which was developed in 50 min with a gradient of 10–66.5% (v/v) acetonitrile in 0.2% trifluoroacetic acid generated in a Beckman model 420 instrument at a flow rate of 1.2 ml/min. The absorbance of the effluent was monitored both at 220 and 280 nm using a Beckman model 165 variable wavelength detector. Further purification of the peptides, when necessary, was achieved by repetition of the HPLC procedure with gradients of a different shape. The manual dansyl-Edman procedure was the degradation technique adopted for the sequence determination of the peptides and was complemented by carboxypeptidase digestions; details of these and other analytical procedures have been published [13].

The hydrophobicity profiles were calculated by using the OMH (optimal matching hydrophobicity) scale, derived on the assumption that families of proteins that share the same type of folding do so because they have the same pattern of residue hydrophobicities along their amino acid sequences [14].

3. RESULTS AND DISCUSSION

HPLC on a macroporous reverse-phase column adequately resolved the peptides obtained after tryptic digestion of the carboxymethylated globin from *S. inaequivalvis* HbI. As illustrated in fig.1, most peptides were sufficiently pure to be analysed by the above-mentioned techniques, a few

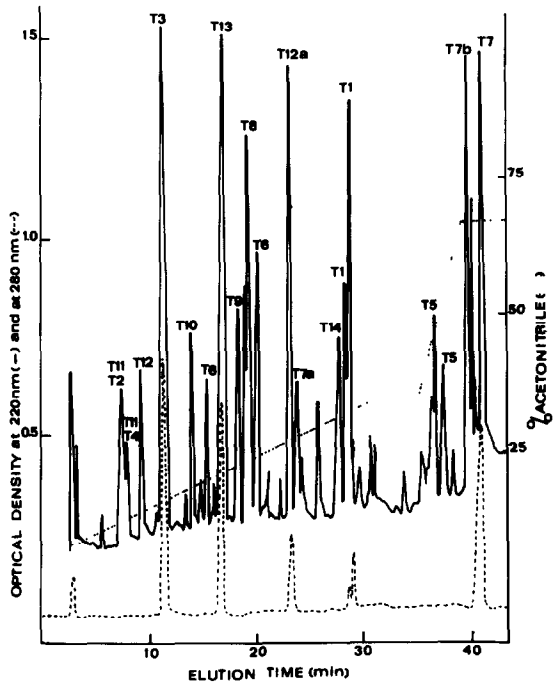

Fig.1. Reverse-phase high performance liquid chromatography of the tryptic digest of carboxymethylated *S. inaequivalvis* HbI globin. Conditions of analysis are reported in the text. The numbers above the peaks refer to the peptides subsequently found in the peaks themselves.

necessitated further purification by a second HPLC step under properly modified gradient conditions. The sequences of all the tryptic peptides are presented in fig.2, ordered according to an unambiguous alignment with the highly homologous structures of *A. broughtonii* and *A. trapezia* [15,16]. The number of differences in the 3 dimeric hemoglobins amounts to at most 15 in 146 amino acids. Of greater interest is the comparison between the sequences of all the dimeric Arcid hemoglobins with that of the so-called α chain from the tetrameric HbII of *A. trapezia* [17]. Such a comparison brings out the presence of highly invariant segments along the polypeptide chains in the region corresponding to the E and F helices (72 and 100% homology, respectively), as observed in [16]. The rationale for the conservation of this part of the chain is now clear on the basis of the X-ray data; it lies in the involvement of the E and F helices in the dimer intersubunit contacts.

The dimeric structure of *S. inaequivalvis* HbI is extremely stable towards dilution, high salt concentrations and changes in pH in the range 5–9 [4,18]. Hence, salt bridges and/or hydrogen bonding are likely to play a less important role in the packing of the two subunits than in the case of vertebrate hemoglobins. Conversely, hydrophobic contacts should be more important in the dimeric clam hemoglobins than in the vertebrate ones. These considerations led us to compare the hydrophobicity profile of the E and F helices in *S. inaequivalvis* HbI, where they form the intersubunit contact, and in the vertebrate hemoglobin and myoglobin chains, where they are exposed to solvent. Fig.3a shows that significant differences occur in *S. inaequivalvis* HbI with respect to the latter proteins: several side chains that are polar in vertebrate globins are substituted by hydrophobic ones in the clam protein, i.e., E9, E13, EF9, EF11, F6, F9. The spacing beween the additional hydrophobic residues in the helical regions is such that their side chains are facing the same side of the helix, the side that is on the exterior of the molecule in vertebrate hemoglobins and myoglobins [19]. Two of these residues occur in the middle of the E helix, approximately where the dyad axis lies that relates the two subunits within the dimer; the other two are towards the end of the F helix where it lies across the E helix of the partner subunit [11].

Our procedure, and hence the assignment of the additional peaks of hydrophobicity to residues that participate in the intersubunit contacts, is validated by the comparisons of other regions of the molecule. Fig.3b and c shows that in several

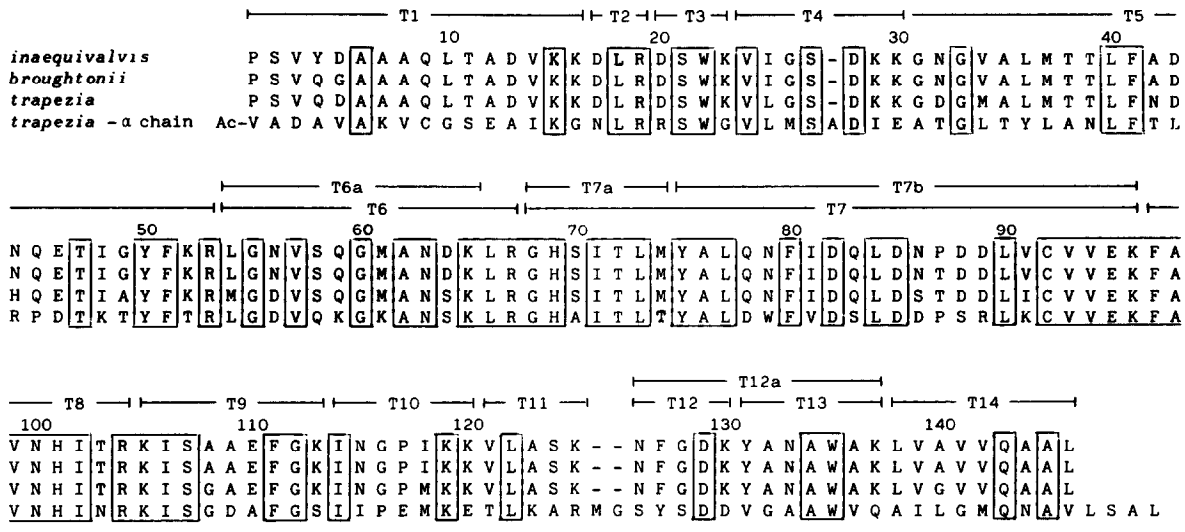

Fig.2. Comparison of the amino acid sequence of *S. inaequivalvis* HbI with those of other Arcid dimeric hemoglobins and of the α chain from the tetrameric *A. trapezia* hemoglobin. Identical residues are indicated by boxes. Segments above the *S. inaequivalvis* HbI sequence denote peptides derived by tryptic digestion of the carboxymethylated globin.

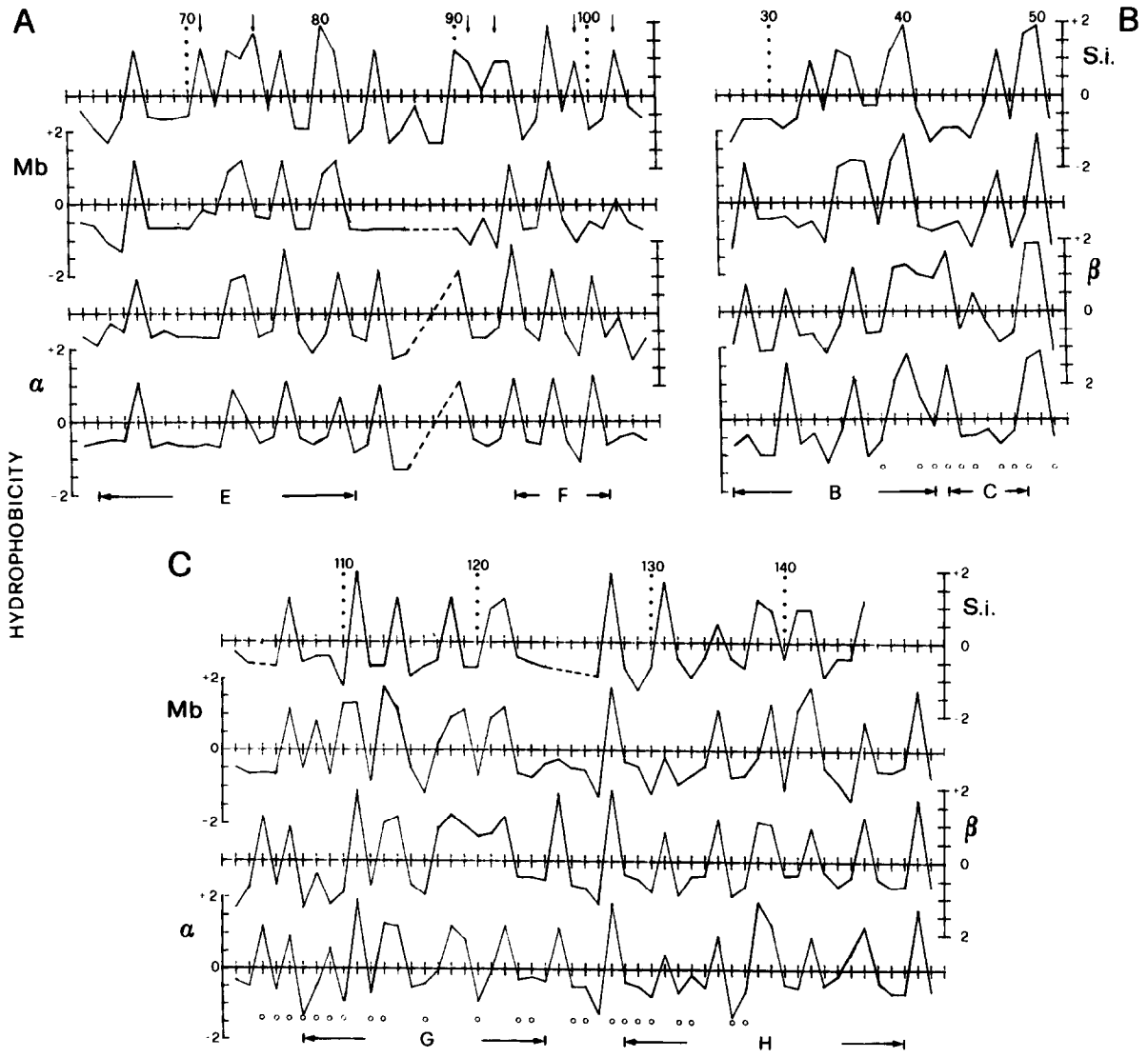

Fig.3. Comparison of the hydrophobicity profiles of *S. inaequivalvis* HbI globin with those of vertebrate hemoglobin and myoglobin chains. The hydrophobicity values for the α, β, and myoglobin chains are average values obtained by means of the OMH (optimal matching hydrophobicity) scale for the selection of sequences given by Dickerson and Geis [20]. The helix notations are those of vertebrate hemoglobins, the numbers those of the *S. inaequivalvis* HbI sequence. (○) Residues involved in intersubunit contacts in vertebrate hemoglobins [20]; (↓) additional hydrophobic residues present in *S. inaequivalvis* HbI. (A) E–F regions; (B) B–C regions; (C) G–H regions.

critical positions at the $\alpha_1\beta_1$ contacts (involving B, G and H helices and the GH corner) and at the $\alpha_1\beta_2$ contacts (involving mainly helices C and G, and the FG corner) the hydrophobicity pattern of *S. inaequivalvis* HbI differs from that of hemoglobin chains and is similar to that of myoglobin. The resemblance with myoglobin is particularly striking in the C helix and FG corner. At the level of the GH corner the comparison is rendered difficult by the deletions present in the clam hemoglobin; in the H helix the critical residues corresponding to 140 and 141 in the α chains and 145 and 146 in the β chains are lacking.

In conclusion, the analysis of the amino acid se-

quence and hydrophobicity profile of *S. inaequivalvis* HbI indicates that the E and F helices are the only regions in the molecule that are characterized by a distinctive hydrophobicity pattern with respect to both the myoglobin and hemoglobin chains. The additional hydrophobic residues are conserved in all the known dimeric Arcid hemoglobins and in the so-called α chain of the tetrameric component from *A. trapezia* [17]. It may be anticipated that the same topology will be found also in the third kind of chain on Arcid hemoglobins as a major factor in the stabilization of their basic dimeric unit. The next step in the understanding of the structure of these proteins will be the elucidation of how the unusual dimer interface operates in the achievement of cooperativity.

ACKNOWLEDGEMENTS

The authors wish to thank Professor K.H. Winterhalter for his help and advice, and Mr M. Sanchioni for skillful technical assistance.

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, 1722–1730.
[3] Djangmah, J.S., Gabbott, P.A. and Wood, E.J. (1978) Comp. Biochem. Physiol. 60B, 245–250.
[4] Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F. and Antonini, E. (1981) J. Mol. Biol. 152, 577–592.
[5] Como, P.F. and Thompson, E.O.P. (1980) Aust. J. Biol. Sci. 33, 643–652.
[6] Furuta, H., Ohe, M. and Kajita, A. (1981) Biochim. Biophys. Acta 668, 448–455.
[7] Antonini, E., Chiancone, E. and Ascoli, F. (1982) in: Structure and Function Relationships in Biochemical Systems (Bossa, F. et al. eds) pp.67–73.
[8] Ikeda-Saito, M., Yonetani, T., Chiancone, E., Verzili, D. and Antonini, E. (1983) J. Mol. Biol. 170, 1009–1018.
[9] Antonini, E., Ascoli, F., Brunori, M., Chiancone, E., Verzili, D., Morris, R.J. and Gibson, Q.H. (1984) J. Biol. Chem. 259, 6739–6738.
[10] Boffi, A., Ascoli, F., Hermans, J. and Chiancone, E. (1984) 8th Int. Biophys. Congr., Bristol, Abstr., p.46.
[11] Royer, W.E. jr, Love, W.E. and Fenderson, F.F. (1985) submitted.
[12] Verzili, D., Santucci, R., Ikeda-Saito, M., Chiancone, E., Ascoli, F., Yonetani, T. and Antonini, E. (1982) Biochim. Biophys. Acta 704, 215–220.
[13] Barra, D., Schininà, M.E., Simmaco, M., Bannister, J.V., Bannister, W.H., Rotilio, G. and Bossa, F. (1984) J. Biol. Chem. 259, 12595–12601.
[14] Eisenberg, D. (1984) Annu. Rev. Biochem. 53, 595–623.
[15] Furuta, H. and Kajita, A. (1983) Biochemistry 22, 917–922.
[16] Fisher, W.K., Gilbert, A.T. and Thompson, E.O.P. (1984) Aust. J. Biol. Sci. 37, 191–203.
[17] Como, P.F. and Thompson, E.O.P. (1980) Aust. J. Biol. Sci. 33, 653–664.
[18] Gattoni, M., Verzili, D., Chiancone, E. and Antonini, E. (1983) Biochim. Biophys. Acta 743, 180–185.
[19] Fermi, G. and Perutz, M.F. (1981) Atlas of Molecular Structure in Biology. 2. Hemoglobin and Myoglobin, Clarendon Press, Oxford.
[20] Dickerson, R.E. and Geis, I. (1983) Hemoglobin: Structure, Function, Evolution, and Pathology, pp.68–69, Benjamin/Cummings, Menlo Park, CA.