Aust. J. Biol. Sci., 1986, **39**, 109–15

Genetic Variation of the Dimeric Haemoglobin of the Bivalve Mollusc *Anadara trapezia*

R. G. Mann, W. K. Fisher, A. T. Gilbert and E. O. P. Thompson

School of Biochemistry, University of New South Wales, P.O. Box 1, Kensington, N.S.W. 2033.

Abstract

The bivalve mollusc *A. trapezia* has two haemoglobins, a tetrameric major haemoglobin, and a dimeric minor haemoglobin, the latter having two identical chains that are different from the chains in the tetramer. Genetic variation in the dimer results in two different haemoglobins, HbIIa and HbIIb, and it is known that the relative proportions of these two polymorphic forms vary with latitude along the eastern coastline of Australia. The HbIIb variant is more common at higher latitudes where water temperature may act as selecting agent. Comparative peptide mapping and amino acid analysis of peptides have shown that in HbIIa an aspartyl residue replaces the seryl residue found in HbIIb at residue 64, position E2 in the E helix. The E and F helices have recently been shown to be highly conserved in arcid globins and to be involved in subunit contacts in the cooperative dimers.

Introduction

In an earlier paper (Como and Thompson 1980*a*) the separation of the multiple haemoglobins of the bivalve mollusc *Anadara trapezia* was described. The major haemoglobin is tetrameric, of the $\alpha_2\beta_2$ type. Two minor polymorphic haemoglobins, HbIIa and HbIIb, are also present. These are dimeric with identical chains that are different from the α- and β-chains of the tetrameric haemoglobin.

The amino acid sequence of the 153 residues of the α-chain (Como and Thompson 1980*b*), the 151 residues of the β-chain (Gilbert and Thompson 1985) and the 146 residues of the globin IIb chain (Fisher *et al.* 1984) have been reported. Dimeric as well as tetrameric haemoglobins have been found in other arcid molluscs, and the amino acid sequence of the dimeric globin from *A. broughtonii* (Furuta and Kajita 1983) and from *Scapharca inaequivalvis* (Petruzzelli *et al.* 1985) have been determined. The sequences show strong homology with the sequence of globin IIb from *A. trapezia*, particularly in the E and F helices. Overall, there were only 16 differences between the amino acid sequences of these three dimeric globin chains: *A. trapezia v. A. broughtonii* showed 15 differences, *v. S. inaequivalvis* 16 differences, and *A. broughtonii v. S. inaequivalvis* only three differences.

Only *A. trapezia* shows two forms of the dimeric haemoglobin. The HbIIa is more common at lower latitudes (further north of Sydney) than HbIIb, which is more prevalent in Sydney waters. These two forms are genetically controlled (Nicol and O'Gower 1957), and it was shown (O'Gower and Nicol 1968; Dixon 1976) that the gene frequency followed a latitudinal cline along the south-eastern coastline of Australia. The discontinuities in the cline correspond with disjunctions between water masses off the east coast of Australia, which suggests that water temperature is acting as a selecting agent.

0004-9417/86/020109$02.00

Amino acid analyses (Como and Thompson 1980*a*) showed only minor differences in composition between globin IIa and IIb. In this paper we present the results of peptide-mapping studies of globins IIa and IIb which suggest that they are identical except for an amino acid difference at residue 64.

Materials and Methods

Haemoglobin Purification

Specimens of *A. trapezia* from estuaries in the Sydney area were bled and the haemoglobins isolated and fractionated by gel filtration as previously described (Como and Thompson 1980*a*). The dimeric haemoglobin fraction was titrated to pH 6·5 and diluted to reduce the sodium ion concentration to below 0·05 M before loading on a column (22 by 4·2 cm) of carboxymethyl-Sephadex equilibrated with a pH 6·5 buffer containing 0·02 M Na_2HPO_4–0·01 M KCN–0·001 M EDTA adjusted from higher pH with phosphoric acid. The column was washed with starting buffer until a stable base line was established and then eluted with similar buffers adjusted to pH 7, 8 and 9. Haemoglobin IIa eluted first, in the pH 8 buffer, while pH 9 buffer eluted the haemoglobin IIb. To reduce cross-contamination, which was most likely in the haemoglobin IIa fraction, the fractions were titrated to pH 6·5, diluted and rechromatographed on a freshly equilibrated column. The fractions were examined by Cellogel electrophoresis (Nicol and O'Gower 1967) before bulking for recovery.

Globin Purification

The globins were prepared and purified by chromatography on carboxymethylcellulose in 8 M urea–thiol buffers to preserve the cysteinyl residues as described previously (Fisher *et al.* 1980). The globins were either carboxymethylated before recovery from the urea buffer or recovered unmodified after dialysis and freeze-drying.

Performic acid oxidation was carried out as described previously (Thompson 1954) with the addition of phenol (1 mg/ml reagent) to minimize halogenation of tyrosine residues as used in the hydrolysis of proteins for the same purpose (Sanger and Thompson 1963).

For amino acid analysis and peptide isolation there is advantage in using performic acid-oxidized globin which gives more accurate values for cysteine and methionine in their oxidized forms of cysteic acid and methionine sulfone, and avoids the production of peptides with partially oxidized *S*-carboxymethylcysteine residues or methionine sulfoxide residues.

Peptide Isolation, Amino Acid Analysis and Sequence Determination

Tryptic or other enzyme digests were fractionated by peptide mapping using paper ionophoresis–chromatography (10–15 mg) or high-performance liquid chromatography as described previously (Fisher *et al.* 1984), and identified by amino acid analysis and *N*-terminal residue determination in comparison with peptides known from the previous work on HbIIb (Fisher *et al.* 1984). For efficient tryptic digestion, the globins were dissolved in 8 M urea–1% (w/v) NH_4HCO_3 and diluted to 2 M urea before adding the trypsin. After 3 h at 37°C the digest was freeze-dried. The peptides were solubilized by adding water until the original volume was obtained for loading on a μbondapak C18 HPLC column.

Results

Amino Acid Composition of Globin IIa

In the previous paper (Como and Thompson 1980*a*), which reported comparative analyses of globins IIa and IIb, the mean differences were in lysine (13·3 and 14·0 residues respectively), aspartic acid (20·8 and 20·2) and glycine (11·5 and 10·5). Subsequent work on the amino acid sequence of globin IIb (Fisher *et al.* 1984) showed that there were 15 residues of lysine, 20 residues of aspartic acid and 11 residues of glycine. The analytical value extrapolated for serine (8·3) gave one residue less than the amino acid sequence (9 residues) required.

When fresh preparations of *S*-carboxymethylated (SCM-) globin IIb and IIa were analysed in parallel, the results were very similar to each other and in agreement with the amino acid sequence of globin IIb with the exception of lysine values, which

were again low by one residue, methionine which gave low values with progressive losses with longer times of hydrolysis, and leucine which was approximately 0·6 residues higher in globin IIa.

It became obvious that it would be difficult by amino acid analysis to detect the true variation between haemoglobins IIa and IIb which results in a charge difference. Such a charge difference could be the result of an Asp-Asn or Glu-Gln change or vice-versa, which would not be detectable by amino acid analysis. A change of this nature would be detectable by peptide mapping since it would affect the electrophoretic mobility of the peptide.

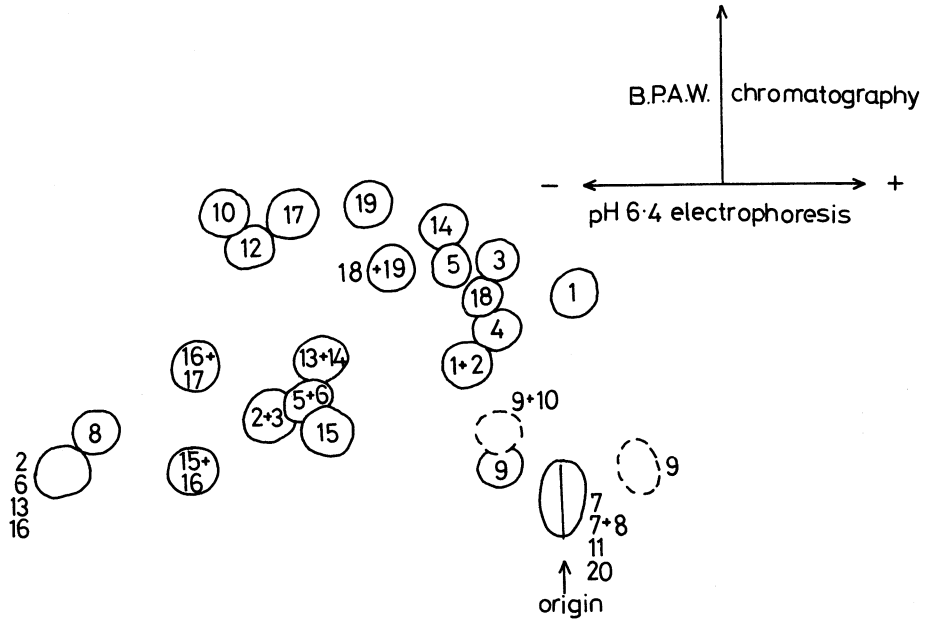
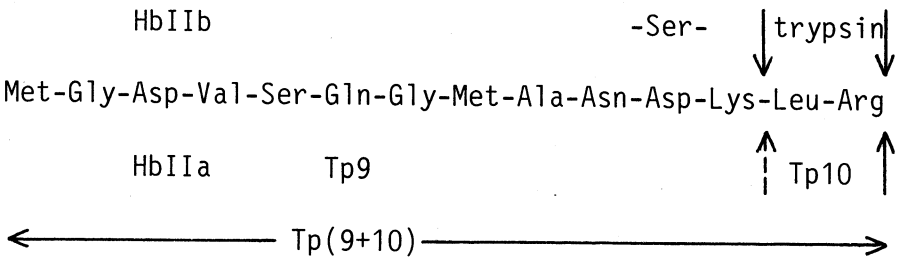


Fig. 1. Peptide map of a tryptic digest of performic acid-oxidized globin IIb of *A. trapezia*. Peptides marked in a dotted line indicate the changed positions of Tp9 and Tp(9 + 10) found in globin IIa. Electrophoresis at pH 6·4 was followed by chromatography with butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12 v/v) as indicated. The peptides are given the number corresponding to their position in the chain.

Peptide Mapping

The peptide map for SCM-globins IIb was reported in the supplementary material to a previous paper (Fisher *et al.* 1984). When samples of SCM-globin IIa were mapped in parallel with SCM-globin IIb, almost identical peptide patterns were obtained after developing with ninhydrin. If the HbIIa or HbIIb was oxidized with performic acid before tryptic digestion, clearer patterns with stronger ninhydrin colours were obtained, since additional partially oxidized peptides containing *S*-carboxymethyl-cysteine sulfoxide and methionine sulfoxide were absent. Fig. 1 shows a similar composite tryptic peptide map for the oxidized globins. The most obvious change was in the vicinity of the origin corresponding to the position of Tp9. This resulted in a peptide from globin IIa which was acidic at pH 6·4 in contrast to Tp9 from globin IIb which was neutral. Amino acid analysis of this material gave a composition with one residue more of aspartic acid and one residue less of serine.

In addition to this new spot there was an additional stronger spot in the neutral region running slightly faster in the chromatography solvent than Tp9 from HbIIb. This analysed as the composite peptide Tp(9 + 10) containing the amino acids leucine and arginine in addition to those normally present as components of Tp9. The composite peptide is the result of the substitution of an aspartic acid residue for the serine residue preceding the lysine in Tp9. The difference is summarized below:



The acidic charge of an aspartyl residue is known to inhibit the complete hydrolysis of an immediately adjacent lysyl or arginyl residue by trypsin (Lehmann and Huntsman 1974). A similar sequence, aspartyllysine, at the *C*-terminal end of Tp18 results in the formation of Tp(18 + 19), as well as the separate peptides.

There are three lysyllysyl, one lysylarginyl and one arginyllysyl sequences in HbIIa and HbIIb. Composite tryptic peptides are usually obtained for the expected tryptic peptides preceding or following these adjoining basic amino acid residues. The presence of those composite peptides assists in the comparison of the sequences of globin IIa and IIb since it indicates that it is the same basic residue in each case, and that there has been no deletion of a residue. In the case of globin IIb the composite peptides Tp(1 + 2), Tp(2 + 3), Tp(5 + 6), Tp(13 + 14) were reported (Fisher *et al.* 1984). In the comparative examination of globin IIa and IIb, Tp(7 + 8), Tp(15 + 16), Tp(16 + 17) and Tp(18 + 19) were also detected.

In the fractionation experiments it was not appreciated at first that there was considerable cross-contamination of globin IIa and IIb. This was improved by rechromatography as described in the methods. The less-purified samples of HbIIa always showed some Tp9 with the composition of the peptide known to be formed from HbIIb. Even with rechromatography, however, there was always some of the Tp9 peptide formed from HbIIb in the HbIIa digests, despite the absence of a band in the HbIIb region during Cellogel ionophoretic analysis of the HbIIa. This is possibly attributable to some deamidation of the HbIIb during isolation, which would change the electrophoretic and chromatographic properties to make them similar to those of HbIIa. The most effective separation of HbIIa and HbIIb occurs at pH values of 8 to 9 and it is known from earlier reports that some deamidation is possible in this pH range (Edmundson 1965; Satterlee *et al.* 1969; Romero-Herrera and Lehmann 1974; Fisher and Thompson 1979; Fisher *et al.* 1980). Indeed, in the earlier paper on the amino acid sequence of HbIIb (Fisher *et al.* 1984) a variant peptide Ch5 was reported with *N*-terminal aspartic acid at residue number 42 (B15) as well as one with *N*-terminal asparagine.

High-performance liquid chromatography (HPLC) was also used to fractionate tryptic digests of globin IIa and IIb under the same conditions as used previously (Fisher *et al.* 1984). The peaks were collected separately and a sample subjected to *N*-terminal group identification by the dansyl method. The peak fractions known to contain Tp9 from the earlier work were collected, dried and mapped on paper

to give Tp9 in the case of HbIIb, and Tp(9 + 10) in the case of HbIIa, as the principal peptides near the origin. The peptide Tp(9 + 10) was a major peptide reported by Petruzzelli *et al.* (1985) from tryptic digests of *S. inaequivalvis* dimeric globin, where aspartyllysylleucyl is the sequence not hydrolysed by trypsin. The HPLC fractions were particularly useful for the isolation of the large tryptic peptides which were insoluble during tryptic digestion but soluble in the reconstituted 8 M urea solution used to load the digest on HPLC columns. It was known from the previous work that Tp7, Tp11 and Tp20 eluted in different peak fractions which can be identified by *N*-terminal group analysis. The amino acid analyses of these fractions were sufficient to indicate that there are no significant amino acid differences between these peptide sequences in HbIIa and HbIIb. For the least-purified insoluble peptide, Tp11, which showed additional impurities containing arginine, glycine and alanine, partial digestion with pepsin and peptide mapping was used to identify peptide sequences in HbIIa that were identical in amino acid composition to those previously found in HbIIb (Fisher *et al.* 1984).

Amino Acid Sequence of the Variant Peptide

The amino acid sequence of the HbIIa peptide Tp9 was shown by stepwise phenyl isothiocyanate degradation in a gas-phase sequencer to be the same as that from HbIIb apart from the substitution of an aspartyl residue for a seryl residue (residue 64 in the globin sequence).

Discussion

The difference between HbIIa and HbIIb has been shown to involve a substitution in HbIIa of an aspartyl residue at residue 64, E2 in the helical notation, for the seryl residue in HbIIb. This substitution resulted in a composite tryptic peptide Tp(9 + 10) being obtained as the major product of digestion rather than Tp9 and Tp10 due to the inhibiting effect of an aspartyl residue on tryptic digestion when it precedes an arginyl or lysyl residue in a peptide chain.

No detailed sequence examination, apart from amino acid composition and *N*-terminal residue, of all the other tryptic peptides was made, so that it is an assumption that they are identical in sequence in HbIIa and HbIIb. The dangers inherent in this assumption were noted by Nolan and Margoliash (1968) who listed three different *C*-terminal tetrapeptide sequences for cytochrome *c* involving the same four amino acids. However, in the case of HbIIa and HbIIb we are dealing with a genetic variant not a species difference.

The change from serine to aspartic acid necessitates two base changes in a codon, which sets it apart from changes associated with abnormal human haemoglobins which involve single base changes. The change might therefore be seen to be consistent with a selection process that results in more favourable oxygenation characteristics that become beneficial to the mollusc under different conditions. The work of O'Gower and Nicol (1968) and Dixon (1976) suggests that the selecting agent could be water temperature associated with the movement of water masses off the east coast of Australia. The lack of polymorphism in arcids, whose distribution is restricted to warm tropical or Mediterranean waters compared with *A. trapezia* whose distribution extends from the tropics to the much cooler waters of Southern Australia, is probably related to the relatively stable water temperatures that occur in these other areas.

Arcid molluscs are generally recognized as being a species associated with tropical waters (Habe 1965). They once extended around the coast of Australia, and a change in climatic conditions lead to their elimination from the cooler waters that resulted (Macpherson and Gabriel 1962). In these changing conditions it should be expected that the polymorphic variant having survival advantage would be the one with the differences from species surviving in warm waters. This would suggest that HbIIb with serine at residue 64 is the variant with survival advantage.

In the amino acid sequence of two other dimeric haemoglobin chains, those of *A. broughtonii* (Furuta and Kajita 1983) and *S. inaequivalvis* (Petruzzelli *et al.* 1985), the residue 64 is aspartic acid as in HbIIa. Interestingly, the work on *S. inaequivalvis* was stimulated by greater physiological adaptation and survival of this species compared with the native *Venus gallina* in the periodic diminished oxygenation due to algal growth in the middle Adriatic Sea (Chiancone *et al.* 1981). In the other tetrameric α- and β-globin chains of *A. trapezia* there is serine at residue 64 as in HbIIb.

The occurrence of the variant amino acid residue at position 64, E2 in the helical nomenclature, places it in the E and F helical regions that are involved in the dimer intersubunit contacts (Royer *et al.* 1985). It is the only residue (in HbIIb) that disturbs the 100% homology between the E and F helices of the three dimeric haemoglobins. It is this E and F helical contact that must be responsible for the co-operative dimer in these molluscs, which is quite different to vertebrate haemoglobin which does not give co-operative dimers (Antonini 1967; Mills *et al.* 1976). In forming dimeric and tetrameric haemoglobins in molluscs it appears that the chains are assembled 'back-to-front' compared with vertebrate haemoglobins (Royer *et al.* 1985) which have the E and F helices exposed on the surface of the molecule. In line with the different role of the E and F helical residues, it has been noted by Petruzzelli *et al.* (1985) that with respect to both the haemoglobin and the myoglobin chains from vertebrates there are several additional hydrophobic residues. If the strong conservation of the E and F helices in the molluscs is an indication of their importance for oxygenation reactions, then a substitution of an amino acid in the E helix might be seen as having a marked influence on these oxygenation characteristics. What the nature of the influence is must await the more detailed information that will come from the crystal structure analysis of the dimeric haemoglobin (Royer *et al.* 1985) at higher resolution.

Acknowledgments

This work was supported in part by the Australian Research Grants Scheme. We thank Mr S. S. Moore for the confirmation of the sequence of the variant peptide.

References

Antonini, E. (1967). Hemoglobin and its reaction with ligands. The alpha-beta dimer appears to be the important unit in hemoglobin function. *Science (Wash., D.C.)* **158**, 1417–25.

Chiancone, E., Vecchini, P., Verzili, D., Ascoli, F., and Antonini, E. (1981). Dimeric and tetrameric hemoglobins from the mollusc *Scapharca inaequivalvis*. Structural and functional properties. *J. Mol. Biol.* **152**, 577–92.

Como, P. F., and Thompson, E. O. P. (1980*a*). Multiple haemoglobins of the bivalve mollusc *Anadara trapezia*. *Aust. J. Biol. Sci.* **33**, 643–52.

Como, P. F., and Thompson, E. O. P. (1980*b*). Amino acid sequence of the α-chain of the tetramic haemoglobin of the bivalve mollusc *Anadara trapezia*. *Aust. J. Biol. Sci.* **33**, 653–64.

Dixon, P. I. (1976). The biology of *Anadara trapezia* with particular reference to its haemoglobin. *Malacol. Rev.* **9**, 139.

Edmundson, A. B. (1965). Amino-acid sequence of sperm whale myoglobin. *Nature (Lond.)* **205**, 883–7.

Fisher, W. K., Gilbert, A. T., and Thompson, E. O. P. (1984). Amino acid sequence of the globin IIb chain of the dimeric haemoglobin of the bivalve mollusc *Anadara trapezia*. *Aust. J. Biol. Sci.* **37**, 191–203.

Fisher, W. K., Koureas, D. D., and Thompson, E. O. P. (1980). Myoglobins of cartilaginous fishes II. Isolation and amino acid sequence of myoglobin of the shark *Mustelus antarcticus*. *Aust. J. Biol. Sci.* **33**, 153–67.

Fisher, W. K., and Thompson, E. O. P. (1979). Myoglobin of the shark *Heterodontus portusjacksoni*: Isolation and amino acid sequence. *Aust. J. Biol. Sci.* **32**, 277–94.

Furuta, H., and Kajita, A. (1983). Dimeric hemoglobin of the bivalve mollusc *Anadara broughtonii*: Complete amino acid sequence of the globin chain. *Biochemistry* **22**, 917–22.

Gilbert, A. T., and Thompson, E. O. P. (1985). Amino acid sequence of the β-chain of the tetrameric haemoglobin of the bivalve mollusc, *Anadara trapezia*. *Aust. J. Biol. Sci.* **38**, 221–36.

Habe, T. (1965). The Arcid subfamily Anadarinae in Japan and its adjacent areas. *Bull. Natl Sci. Mus. (Tokyo)* **8**, 71–85.

Lehmann, H., and Huntsman, R. G. (1974). 'Man's Haemoglobins.' 2nd Edn. (North-Holland Publishing Co.: Amsterdam.)

Macpherson, J. H., and Gabriel, C. J. (1962). 'Marine Molluscs of Victoria.' (Melbourne University Press.)

Mills, F. C., Johnson, M. L., and Ackers, G. K. (1976). Oxygen-linked subunit interactions in human hemoglobin: experimental studies on the concentration dependence of oxygeneration curves. *Biochemistry* **15**, 5350–62.

Nicol, P. I., and O'Gower, A. K. (1967). Haemoglobin variation in *Anadara trapezia*. *Nature (Lond.)* **216**, 684.

Nolan, C., and Margoliash, E. (1968). Comparative aspects of primary structures of proteins. *Annu. Rev. Biochem.* **37**, 727–90.

O'Gower, A. K., and Nicol, P. I. (1968). A latitudinal cline of haemoglobins in a bivalve mollusc. *Heredity* **23**, 485–91.

Petruzzelli, R., *et al.* (1985). Amino acid sequence of the cooperative homodimeric hemoglobin from the mollusc *Scapharca inaequivalvis* and topology of the intersubunit contacts. *FEBS Letts.* **184**, 328–32.

Romero-Herrera, A. E., and Lehmann, H. (1974). The amino acid sequence of human myoglobin and its minor fractions. *Proc. R. Soc. Lond.* **B186**, 249–79.

Royer, W. E., Love, W. E., and Fenderson, F. F. (1985). The cooperative dimeric and tetrameric clam hemoglobins are novel assemblages of myoglobin folds. *Nature (Lond.)* **316**, 277–80.

Sanger, F., and Thompson, E. O. P. (1963). Halogenation of tyrosine during acid hydrolysis. *Biochim. Biophys. Acta* **71**, 468–71.

Satterlee, L. D., Lillard, H. S., and Snyder, H. E. (1969). Separation of microheterogeneous components of bovine metmyoglobin and investigation of possible causes of the microheterogeneity. *Life Sci.* **8**, 871–84.

Thompson, E. O. P. (1954). Modification of tyrosine during performic acid oxidation. *Biochim. Biophys. Acta* **15**, 440–1.

Manuscript received 5 February 1986, accepted 14 April 1986