STUDIES ON MONOTREME PROTEINS

III.* AMINO ACID SEQUENCE OF THE α- AND β-GLOBIN CHAINS OF THE MINOR HAEMOGLOBIN FROM THE ECHIDNA, *TACHYGLOSSUS ACULEATUS ACULEATUS*

By E. O. P. THOMPSON,† W. K. FISHER,† and R. G. WHITTAKER†

[Manuscript received 10 May 1973]

Abstract

Echidnas exhibit polymorphism in their haemoglobins. The minor component in animals from the mainland of Australia is Hb-IIA, which can be identified by electrophoretic migration rate and separated by chromatography on DEAE-Sephadex. This minor component has α- and β-globin chains which were only partially separated by chromatography on carboxymethylcellulose columns. By tryptic digestion of the partially purified chains, followed by fractionation by paper ionophoresis and chromatography, the component peptides from the α- and β-chains were isolated and their amino acid sequences determined.

The β-chain was identical in amino acid sequence with the β-chain from the major haemoglobin component, and the α-chain varied in only nine positions from the α-chain of the major haemoglobin component, Hb-IB, whose amino acid sequences were previously reported. From the number of changes in sequence, its presence in all echidnas, and the proportion of the minor component, it must be the product of non-allelic α-chain genes.

There were no changes in "contact site" residues in the α-chain of the minor haemoglobin compared with those present in the major haemoglobin.

I. INTRODUCTION

In Parts I and II of this series (Whittaker *et al.* 1972, 1973) the fractionation of the haemoglobins and of the globin chains from the south-eastern Australian echidna *Tachyglossus aculeatus aculeatus* (Cooper *et al.* 1973) was reported, together with the amino acid sequences of the α- and β-globin chains of the major haemoglobin component, Hb-IB.

In the present paper the amino acid sequences of the α- and β-globin chains from the minor haemoglobin, Hb-IIA, are reported.

II. MATERIALS AND METHODS

The haemoglobins were separated by chromatography on DEAE-Sephadex (Whittaker *et al.* 1972). The globin from Hb-IIA was fractionated on carboxymethylcellulose columns eluted with a gradient of sodium (phosphate) ion from 0·01 to 0·05M in 8M urea–mercaptoethanol buffers, and carboxymethylated as previously described. An alternate method was to first carboxymethylate the

* Part II, *Aust. J. biol. Sci.*, 1973, **26**, 877–88.

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

mixture of α- and β-globin chains (150 mg) in 5 ml of 8M urea containing 150 μl of mercaptoethanol plus 0·15 ml of 3M Tris buffer, pH 8·5, which had been under nitrogen at 37°C for 1 hr. Iodoacetic acid (0·4 g) in 1 ml of 3M Tris buffer was reacted with the reduced protein mixture for 15 min at 37°C before dialysing against water and recovering the *S*-carboxymethylated chains by freeze-drying. In this case fractionation on carboxymethylcellulose columns in 8M urea solution was performed in the absence of thiols.

The methods of digestion with trypsin, chymotrypsin, and thermolysin, fractionation of peptides, and sequence determination by the dansyl–Edman procedure were the same as in the previous papers (Whittaker *et al.* 1972, 1973). As before, the residues determined by the dansyl–Edman method are printed in *italic* fount.

III. Results

The tryptic peptides are numbered from the *N*-terminus and correspond exactly in size with those previously reported for Hb-IB.

(*a*) *Fractionation of Globin Chains*

The separation of the α- and β-chains from the globin of Hb-IIA was only partial compared with the separation of globin chains from Hb-IB (Fig. 1). No improvement in resolution was found if the fractionation was carried out on the *S*-carboxymethylated chains. The α-chain from Hb-IIA is eluted more rapidly than that from Hb-IB and tends to merge with the β-chain elution peak. Variation in the gradient used to effect separation was tried without success.

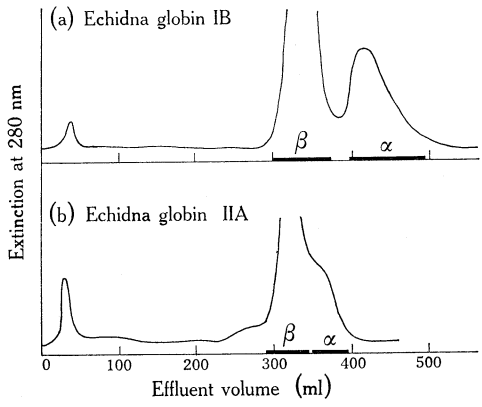

Fig. 1.—Continuous recording of the extinction at 280 nm of the effluent from a linear gradient elution of echidna globins IB (*a*) and IIA (*b*) from carboxymethylcellulose equilibrated with 8M urea–thiol buffer. Initial buffer was 8M urea–0·05M mercaptoethanol–0·01M Na^+ (phosphate)–0·001M EDTA at pH 6·7; flow rate 96 ml/hr; column size 1·8 cm diam. by 12 cm; limit buffer 250 ml, containing 8M urea–0·05M mercaptoethanol–0·05M Na^+ (phosphate)–0·001M EDTA. Approximately 150 mg of globin was loaded in each case. The fractions pooled for sequence studies are indicated by bars on the abscissae.

(*b*) *Amino Acid Sequence of the β-Chain*

The fraction of *S*-carboxymethylated globin (Fig. 1*b*) that was predominantly β-chain was digested with trypsin and, after removal of the insoluble "core" by centrifugation, dried *in vacuo*. The peptides soluble at pH 6·4 in pyridine–acetate buffer were separated from the insoluble peptide βTp2,3 by centrifugation and fractionated by peptide mapping. The β-chain peptides were in the same relative positions (Fig. 2) as the β-chain peptides from Hb-IB (Whittaker *et al.* 1972). All peptides were sequenced by the dansyl–Edman procedure and gave results in agreement with the sequence in Hb-IB for peptides βTp1, βTp4, βTp6, βTp7,8, βTp9A, βTp9B, βTp10A, βTp10B+11,12A, βTp12C, βTp13, βTp14, and βTp15. The

sequence of βTp2,3 was determined by sequencing the peptides obtained after thermolysin digestion of the fraction soluble at pH 8·7 but insoluble at pH 6·4. The results were identical to those previously reported for the Hb-IB peptide, confirming both the sequence and the identification of the amide residues due to asparagine.

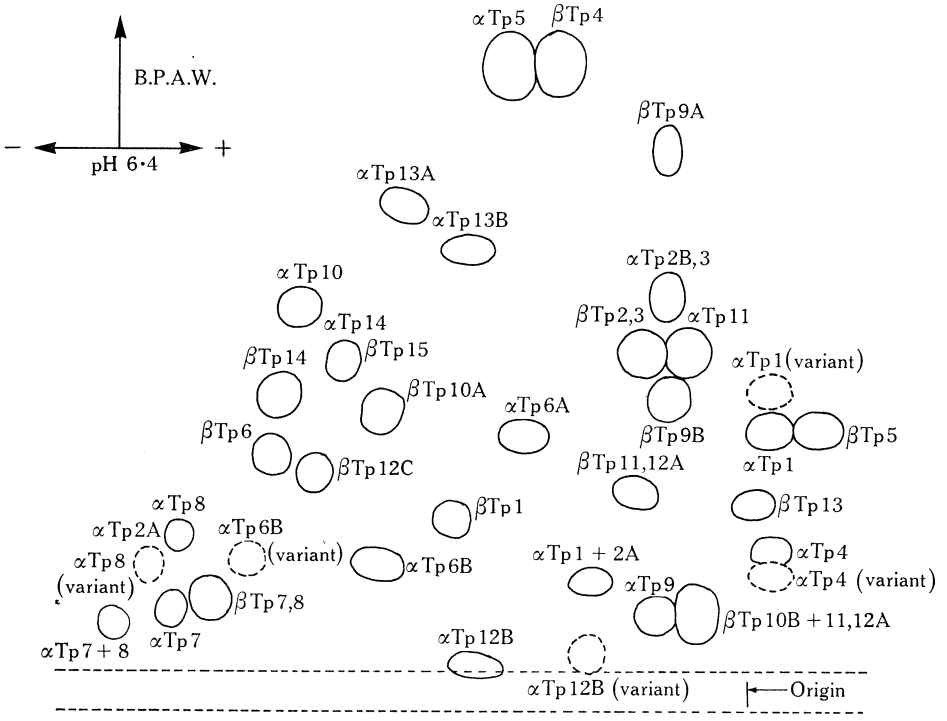

Fig. 2.—Composite peptide map of the α- and β-chain tryptic peptides from both Hb-IB and Hb-IIA. The peptides were fractionated by ionophoresis at pH 6·4 followed by chromatography with butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12 v/v; B.P.A.W.) as indicated. The peptides are given the identification number corresponding to their position in the chain (see Whittaker *et al.* 1972, 1973). The β-chain peptides from both haemoglobins were the same. The α-chain peptides from Hb-IIA which differed in ionophoretic or chromatographic rate from the corresponding Hb-IB peptides are shown by dotted lines.

The *N*-terminal sequence of βTp5 was successfully determined for 11 residues. The peptide was also cleaved by cyanogen bromide to give the *C*-terminal peptide *Gly-Asn-Ala*-Lys. Further samples of βTp5 were hydrolysed with thermolysin and the peptide fractionated by paper ionophoresis–chromatography. The peptides *Ala-Asp*-Ala and *Val-Met-Gly-Asn* were identified to provide the data for the complete sequence being identical with that of the peptide from Hb-IB.

The peptide βTp12B was isolated and purified from the insoluble "core" from the tryptic digestion. The core was redissolved in 5% formic acid and heated in a boiling water-bath for 5 min, then adjusted to pH 8·7 with ammonium bicarbonate solution and digested again with trypsin. The precipitate was centrifuged and washed with fresh ammonium bicarbonate solution. Amino acid analysis showed the

following composition, which is similar to that for the βTp12B peptide from Hb-IB: $Arg_{1.2}Asp_{1.4}Gly_{0.9}Ala_{1.1}Val_{3.7}Leu_{3.0}$, with a background of other amino acids representing contaminating peptides which included βTp4. The βTp12B fraction was sufficiently pure for sequence analysis. For the dansyl–Edman method best results were obtained if the peptide was dissolved in 50% propanol in order to sample for the dansylation reaction, and then coupled with phenyl isothiocyanate after adding 5% *N*-ethylmorpholine as a buffer. The degradation was carried out for five residues to give the sequence *Leu-Gly-Asx-Val-Leu*. Digestion of another sample with thermolysin followed by fractionation of the peptides on paper showed the presence of (Leu,Ala,Arg) and *Ala-Arg* as basic peptides, with (Leu,Gly,Asn) as a neutral spot. This is sufficient to show that the peptide has the same sequence as the corresponding peptide from Hb-IB.

(*c*) *Variable Amino Acid Residues in β-Chains*

Although the amino acid sequence of the β-chain in the Hb-IIA studied here was identical with that previously reported for Hb-IB, it should be noted that in a β-chain from Hb-IB from a different animal evidence for changes in the residues at two positions has been obtained. During fractionation of the globin chains on carboxymethylcellulose using urea–thiol buffers some resolution of the β-chains into two peaks was obtained.

The first change was in βTp1 where an additional peptide with a different electrophoretic mobility was evident. The sequence was shown to be

Gly
or
Val-His-Leu-Ser-Glu-Ser-Glu-Lys
1 5 8

with glutamic acid replacing the glycine at position 5 in the normal peptide. This change would explain a more rapid elution of the β-chain from carboxymethylcellulose columns.

In the same animal there were two different βTp14 peptides with similar rates on ionophoresis at pH 6·4 but which were separated by chromatography and were found to have the sequences

Ser
or
Leu-Val-Ser-Gly-Val-Ala-His-Ala-Leu-Ala-His-Lys
133 144

These changes are most likely the result of allelic genes in this particular animal.

(*d*) *Amino Acid Sequence of the α-Chain*

The peptide maps of the soluble tryptic peptides of the α-chain were similar to those from the α-globin of Hb-IB, although two peptides were in obviously different positions. These were αTp6B (yellow with ninhydrin) which had a higher mobility at pH 6·5 than the corresponding peptide from Hb-IB, and αTp12B which had a reduced mobility (see Fig. 2). Free lysine was present but free arginine was not, in contrast to its formation during tryptic digestion of the α-chain of Hb-IB.

Similar ionophoretic mobilities or positions on a peptide map are not conclusive evidence of identical amino acid sequences. Amino acid composition may indicate changes in amino acids with uncharged side chains. Thus amino acid composition changes were detected in αTp1, αTp4, αTp6B, αTp8, αTp9, and αTp12B. Peptides αTp2B,3, αTp5, αTp6A, αTp7, αTp10, αTp11, αTp12A, αTp13A, αTp13B, and αTp14 had amino acid sequences which were exactly the same as those from Hb-IB α-chain. The peptides with different amino acid sequences are discussed in turn.

αTp1

Peptide αTp1 differed in amino acid composition from the corresponding peptide from Hb-IB by having arginine in place of lysine. The sequence was shown to be the same except for this change in the *C*-terminal residue.

αTp4

Peptide αTp4 differed in amino acid composition from the corresponding peptide from Hb-IB by having a residue of aspartic acid replacing one glutamic acid residue. The peptide was digested with chymotrypsin to give two major peptides, one corresponding to the *N*-terminal sequence (residues 17–24) which contained the aspartic acid, and the other, which gave a yellow colour with ninhydrin, corresponding to the *C*-terminal sequence. The dansyl–Edman procedure gave the sequence of both peptides and located the aspartic acid residue at position 23. Both peptides were acidic at pH 6·4, indicating that, as in Hb-IB α-chain, all the glutamic and aspartic acid residues were present as such and not as glutaminyl or asparaginyl residues.

αTp6B

Peptide αTp6B had a changed ionophoretic rate compared with the peptide from Hb-IB, being more basic. Amino acid analysis showed the presence of histidine and arginine with no lysine or glutamic acid as previously. The dansyl–Edman method showed that histidine replaces glutamine at residue 54, the arginine replacing lysine as the *C*-terminal residue.

αTp8

Residue 61, which was arginine in Hb-IB α-chain, is lysine (αTp8) in Hb-IIA α-chain. This was shown by the absence of free arginine in the tryptic digest and by the presence of αTp7,8, which contained additional lysine compared with αTp7.

αTp9

Peptide αTp9 is the largest peptide, with 29 residues, in tryptic digests of either α- or β-chain. It usually does not appear as a definite spot on a peptide map, due to its size and to the valine *N*-terminal residue, which gives a weak colour with ninhydrin. It is most readily isolated as the fastest band when tryptic digests of either α- or α- + β-chain are subjected to gel filtration on Sephadex G50 in 1% ammonium

bicarbonate. It gives no absorption at 280 nm and precedes the first peak that absorbs at this wavelength. Amino acid analysis of this fraction gave the following results:

$$Lys_{1\cdot2}His_{2\cdot5}Asp_{5\cdot8}Thr_{1\cdot7}Ser_{1\cdot7}Gly_{2\cdot0}Ala_{4\cdot5}Val_{2\cdot0}Met_{0\cdot9}Leu_{3\cdot9}Phe_{1\cdot0}$$

which, compared with the analysis of αTp9 from Hb-IB, shows extra residues of valine, aspartic acid, and glycine replacing one serine and two alanine residues.

Digestion with thermolysin followed by peptide mapping gave a series of peptides. Sequence determinations showed that these peptides mostly corresponded with those given by Hb-IB αTp9 and the different residues were readily identified. The sites of cleavage are indicated below, and the residues which are different in Hb-IB are shown above the sequence to indicate the changes:

62 (Ala) (Ser) 80
Val-Ala-Asp-Ala-Leu-Thr-Thr-Ala-Val-Gly-His-Phe-Asn-Asp-Met-Asp-Gly-Ala-Leu-
Th1 ↑ Th2 ↑ Th3 ↑ Th4A ↑ Th4B ↑

(Ala) 90
*Ser-Asp-Leu-Ser-Asp-Leu-His-Ala-His-*Lys
Th5 ↑ Th6 ↑ Th7 ↑ Th8

Peptides Th1, Th2, Th6, Th7, and Th8 are identical in position on the peptide map and in sequence to those from Hb-IB. Peptide Th3 had a valine residue as residue 2 in place of an alanine residue. The replacement of alanine by aspartic acid at position 82 resulted in peptide Th5 being identical to Th6. No neutral peptide corresponding to Th5 from Hb-IB was detected.

The enzyme gave an additional split in the peptide Th4, cleaving the peptide bond preceding the methionine residue. The mobility of Th4A at pH 6·4 indicated one acidic residue while Th4B, being acidic, showed the presence of aspartic acid. The substitution of glycine for the serine residue at position 78 was proved by the dansyl–Edman method. After removing two residues of Th4A by the Edman method a small amount of aspartic acid was found on ionophoresis at pH 1·8, suggesting that the sequence is -Asn-Asp- as it was in αTp9 from Hb-IB.

Confirmatory evidence for the sequence was obtained by chymotryptic digestion and peptide mapping. The following were the main points of splitting:

Val-Ala-Asp-Ala-Leu-Thr-Thr-Ala-Val-Gly-His-Phe-Asn-Asp-Met-Asp-Gly-Ala-Leu-
↑ ↑ ⋮

Ser-Asp-Leu-Ser-Asp-Leu-His-Ala-His-Lys
↑ ⋮

αTp12A

This peptide is insoluble and was isolated from the "core" of the tryptic digest after solubilization in buffer at pH 3·1, heat denaturation, and redigestion with trypsin. The insoluble material was centrifuged and washed well with ammonium bicarbonate solution followed by 60% pyridine.

The dansyl–Edman method gave the sequence as found for Hb-IB α-chain for 10 residues. Digestion with thermolysin gave the peptides *Ala*-Arg and Leu-Ala-Arg, which completed the sequence and showed it to be identical to the corresponding

Hb-IB sequence. In addition, peptides which gave (Leu,Ala,His,SCM-Cys) and (Phe,Leu,Val) on hydrolysis were obtained.

αTp12B

Peptide αTp12B had a changed ionophoretic mobility that indicated an extra acidic residue. The peptide was digested with thermolysin then peptide mapped. The sequence and points of cleavage are shown below:

112 (Ala) 127
His-His-Pro-Glu-Glu-Phe-Thr-Pro-Ser-Ala-His-Ala-Ala-Met-Asp-Lys
Th1 ↑ Th2 ↑ Th3 ↑ Th4

Only peptide Th1 was different from the corresponding peptides from Hb-IB, the Hb-IB residue being shown in parentheses. This peptide was no longer basic at pH 6·4, being slightly acidic, probably due to incomplete ionization of the histidine residues. When sequenced there was overlap in the dansylamino acids, due to the partial removal of histidine residues during coupling as reported previously (Whittaker *et al.* 1973). Glutamic acid was the major dansylamino acid for residues 4 and 5.

The complete sequence of the amino acid residues of the α-IIA chain is shown in Table 1. The residues that are different in α-IB are shown above the variant residue.

IV. Discussion

The sequence of amino acid residues in the α-chain of echidna Hb-IIA differs at nine positions from that of Hb-IB. Although a clear separation of the α- and β-chains was not obtained, and amino acid compositions of the separate chains were not determined, the tryptic peptides were separated by peptide mapping and identified by their close correspondence in position with the peptides obtained from Hb-IB. The variant peptides were easily detected when they differed in ionophoretic rate. The nine changes in amino acid residues in Hb-IIA α-chain (Table 1) result in an increased net negative charge at pH 6·7, which accounts for its more rapid elution from carboxymethylcellulose, and hence its overlapping with the β-chain, compared with the later elution of the α-chain of Hb-IB.

The peptides were characterized directly by amino acid sequencing, rather than by amino acid analysis followed by sequencing, since analysis facilities were limited. Evidence was obtained for each position in the chains and in the case of the β-chains the data showed that this chain is common to Hb-IB and Hb-IIA. For the α-chains the changes correspond to minimum base changes of one residue in the messenger RNA. Many of the changes in Hb-IIA are to residues that have been detected in these positions in other species (Dayhoff 1972).

None of the residues that are different in Hb-IB and Hb-IIA α-chains are "contact" amino acids (Perutz 1969) involved in α_1–β_1 or α_1–β_2 interactions.

The echidna, in exhibiting polymorphism in its haemoglobins which is a result of changes in nine residues in the α-chain with retention of a common β-chain,*

* In the absence of direct evidence Cooper *et al.* (1973) mistakenly attributed the polymorphism to variation in the β-chain (see also corrigendum, p. 1444).

resembles other species such as the sheep (16 changes in β-chain; Boyer *et al.* 1967), goat (four changes in α-chain; Huisman *et al.* 1968), mouse (three changes in α-chain; Hilse and Popp 1968), horse (one change in α-chain; Kilmartin and Clegg 1967), and human (one change in γ-chain; Schroeder *et al.* 1968) where non-allelic genes have been found. Hb-IIA is common to all echidnas found on the mainland of Australia (Cooper *et al.* 1973) and is a minor component. Its functional significance is not known.

TABLE 1

AMINO ACID SEQUENCE OF THE α-CHAIN OF ECHIDNA HAEMOGLOBIN IIA

Residues are numbered from the *N*-terminus and the positions of cleavage are indicated by arrows. A dotted arrow indicates limited cleavage. The nine residues which differ in the Hb-IB sequence are shown above the variant position

```
                5     Lys           10                    15                  20
Val-Leu-Thr-Asp-Ala-Glu-Arg-Lys-Glu-Val-Thr-Ser-Leu-Trp-Gly-Lys-Ala-Ser-Gly-His-Ala-Glu-
           αTpl          ⇡   ↑          αTp2B,3              ↑
                        αTp2A

Glu    25                30                    35                  40
Asp-Tyr-Gly-Ala-Glu-Ala-Leu-Glu-Arg-Leu-Phe-Leu-Ser-Phe-Pro-Thr-Thr-Lys-Thr-Tyr-Phe-Ser-
              αTp4               ↑                 αTp5                ↑

45                  50              Gln 55 Lys           60 Arg              65
His-Met-Asp-Leu-Ser-Lys-Gly-Ser-Ala-His-Val-Arg-Ala-His-Gly-Lys-Lys-Val-Ala-Asp-Ala-Leu-
       αTp6A          ↑        αTp6B       ↑     αTp7     ⇡   ↑
                                                          αTp8

           Ala ·                75          Ser    80    Ala         85
Thr-Thr-Ala-Val-Gly-His-Phe-Asn-Asp-Met-Asp-Gly-Ala-Leu-Ser-Asp-Leu-Ser-Asp-Leu-His-Ala-
                                       αTp9

  90                  95                    100                 105                 110
His-Lys-Leu-Arg-Val-Asp-Pro-Val-Asn-Phe-Lys-Leu-Leu-Ala-His-Cys-Phe-Leu-Val-Val-Leu-Ala-
     ↑ αTp10 ↑          αTp11          ↑                     αTp12A

             Ala                 120                125                 130
Arg-His-His-Pro-Glu-Glu-Phe-Thr-Pro-Ser-Ala-His-Ala-Ala-Met-Asp-Lys-Phe-Leu-Ser-Arg-Val-
 ↑                             αTp12B                          ↑   αTp13A    ↑

    135             140
Ala-Thr-Val-Leu-Thr-Ser-Lys-Tyr-Arg
     αTp13B          ↑ αTp14
```

The variation in amino acids in positions 5 and 138 of the β-chain reported here for blood from a particular animal is probably evidence of allelic genes for the β-chain as observed in other species such as cattle (Schroeder *et al.* 1967), sheep (Boyer *et al.* 1967), and kangaroos (Thompson and Air 1971).

V. ACKNOWLEDGMENTS

This work was supported in part by the Australian Research Grants Committee. The authors are indebted to Dr. M. Griffiths and Dr. D. W. Cooper for samples of echidna blood.

VI. References

BOYER, S. H., HATHAWAY, P., PASCASIO, F., BORDLEY, J., ORTON, C., and NAUGHTON, M. A. (1967).—*J. biol. Chem.* **242**, 2211.

COOPER, D. W., VANDEBERG, J. L., GRIFFITHS, M. E., and EALEY, E. H. M. (1973).—*Aust. J. biol. Sci.* **26**, 605.

DAYHOFF, M. O. (1972).—"Atlas of Protein Sequence and Structure." (National Biochemical Research Foundation: Silver Spring.)

HILSE, K., and POPP, R. A. (1968).—*Proc. natn. Acad. Sci. U.S.A.* **61**, 930.

HUISMAN, T. H. J., BRANDT, G., and WILSON, J. B. (1968).—*J. biol. Chem.* **243**, 3675.

KILMARTIN, J. V., and CLEGG, J. B. (1967).—*Nature, Lond.* **213**, 269.

PERUTZ, M. F. (1969).—*Proc. R. Soc.* B **173**, 113.

SCHROEDER, W. A., HUISMAN, T. H. J., SHELTON, J. R., SHELTON, J. B., KLEIHAUER, E. F., DOZY, A. M., and ROBBERSON, B. (1968).—*Proc. natn. Acad. Sci. U.S.A.* **60**, 537.

SCHROEDER, W. A., SHELTON, J. R., SHELTON, J. B., ROBBERSON, B., and BABIN, D. R. (1967).—*Archs Biochem. Biophys.* **120**, 124.

THOMPSON, E. O. P., and AIR, G. M. (1971).—*Aust. J. biol. Sci.* **24**, 1199.

WHITTAKER, R. G., FISHER, W. K., and THOMPSON, E. O. P. (1972).—*Aust. J. biol. Sci.* **25**, 989.

WHITTAKER, R. G., FISHER, W. K., and THOMPSON, E. O. P. (1973).—*Aust. J. biol. Sci.* **26**, 877.