Aust. J. Biol. Sci., 1976, **29**, 57–72

Studies on Monotreme Proteins VII.* Amino Acid Sequence of Myoglobin from the Platypus, *Ornithorhynchus anatinus*

W. K. Fisher and E. O. P. Thompson

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

Abstract

Myoglobin isolated from skeletal muscle of the platypus contains 153 amino acid residues. The complete amino acid sequence has been determined following cleavage with cyanogen bromide and further digestion of the four fragments with trypsin, chymotrypsin, pepsin and thermolysin. Sequences of the purified peptides were determined by the dansyl–Edman procedure.

The amino acid sequence showed 25 differences from human myoglobin and 24 from kangaroo myoglobin. Amino acid sequences in myoglobins are more conserved than sequences in the α- and β-globin chains, and platypus myoglobin shows a similar number of variations in sequence to kangaroo myoglobin when compared with myoglobin of other species. The date of divergence of the platypus from other mammals was estimated at 102 ± 31 million years, based on the number of amino acid differences between species and allowing for multiple mutations during the evolutionary period. This estimate differs widely from the estimate given by similar treatment of the α- and β-chain sequences and a constant rate of mutation of globin chains is not supported.

Introduction

In earlier papers concerned with the amino acid sequences of the α- and β-globin chains of the echidna (Whittaker *et al.* 1972, 1973) and the platypus (Whittaker and Thompson 1974, 1975) attempts were made to calculate the date of divergence of the monotremes from the other mammals. The estimates differed widely for the α- and β-globin chains (Whittaker and Thompson 1975), although an earlier divergence of monotremes compared with marsupials seemed likely. No evidence was found to support the theory of Gregory (see Griffiths 1968) that the monotremes are more closely related to the marsupials.

The present paper is concerned with the amino acid sequence of myoglobin from skeletal muscle of the platypus. The complete amino acid sequence has been determined following cleavage with cyanogen bromide and further digestion of the four fragments with trypsin, chymotrypsin, pepsin and thermolysin. Sequences of the purified peptides were determined by the dansyl–Edman procedure and the sequence of myoglobin was deduced from the sequences of the overlapping peptides and the strong homology with other myoglobins of known amino acid sequence.

The sequence is compared with that of myoglobins from other species, and estimates of the date of divergence of the platypus from other mammals have been made using the number of differences in amino acid sequence. There are no significant fossil remains of the monotremes and the estimate of the time of divergence using myoglobin sequences is compared with similar estimates made with α- and β-globin comparisons (Whittaker and Thompson 1974, 1975).

* Part VI, *Aust. J. Biol. Sci.*, 1975, **28**, 353–65.

Materials and Methods

Preparation of Platypus Myoglobin

The skeletal muscle (500 g) from two animals that had been stored frozen and subsequently thawed was cut into small pieces and homogenized for 2 min in a Waring blender with 1200 ml 70% saturated $(NH_4)_2SO_4$ solution [436 g $(NH_4)_2SO_4$, 10 ml 0·1M EDTA, 36 g $Na_2HPO_4.12H_2O$ per litre adjusted to pH 6·5 with H_3PO_4 (Hapner *et al.* 1968)]. The homogenate was centrifuged at 11 700 *g* for 1 h at 4°C, the supernatant filtered through glass wool and Whatman No. 1 paper, and 10 mg $K_3Fe(CN)_6$ was added per 100 ml filtrate. The oxidized filtrate was concentrated by pressure dialysis (Berggard 1961) and dialysed against water and several changes of phosphate buffer, pH 6·4 (98 ml 0·5M KH_2PO_4, 10 ml 0·15M KCN, 10 ml 0·1M EDTA and approximately 17 ml 0·5M $Na_2HPO_4.12H_2O$ per litre).

The dialysed solution was centrifuged at 30 000 *g* for 10 min at 4°C and the solution loaded on a 72 by 5·5 cm column of sulphopropyl-Sephadex (C50) equilibrated and eluted with the phosphate buffer at pH 6·4. Fractions were read at 280 and 540 nm, and the haem-containing peaks concentrated by pressure dialysis. Each peak was examined by electrophoresis at pH 8·6 on cellulose acetate before removing the haem groups by acid–acetone precipitation at −20°C (Anson and Mirsky 1930) using 1·5 ml 10M HCl per 100 ml acetone. The precipitated apomyoglobin and globin fractions were washed with cold acetone, dissolved and dialysed against water, and freeze-dried and stored at −20°C. Each haem-free fraction was examined by cellulose acetate electrophoresis in 8M urea buffer at pH 8·6 (Nash and Thompson 1974).

Cyanogen Bromide Cleavage

Reaction with cyanogen bromide (Gross and Witkop 1962) was carried out by dissolving the apomyoglobin (300 mg) in 100% formic acid and diluting to 70% formic acid (10 ml). After cooling to 4°C, cyanogen bromide (300 mg) was added and the mixture allowed to react at 4°C for 24 h. Cyanogen bromide was removed by freeze-drying and the products separated on a column of Sephadex G50 (68 by 2·3 cm) equilibrated with 10% formic acid. Tubes were read at 280 nm and the pooled fractions recovered by freeze-drying from dilute aqueous solution.

Amino Acid Sequence Methods

The methods of digestion with proteolytic enzymes, of high-voltage ionophoresis, peptide mapping, amino acid analysis and sequence determination by the dansyl–Edman procedure were the same as previously described (Air and Thompson 1969, 1971; Beard and Thompson 1971). Additional analyses for sulphydryl groups followed the method of Ellman (1959), and for tryptophan the spectrophotometric method of Beaven and Holiday (1952).

Estimation of Dates of Divergence

The myoglobin sequence of the platypus was used to estimate a date of divergence from the other mammalian groups of known myoglobin amino acid sequence in the same manner previously reported for the α- and β-chain sequences (Whittaker and Thompson 1974, 1975).

As before, the original method of Air *et al.* (1971) and Thompson and Air (1971) was modified to accommodate the palaeontological estimates of the dates of divergence listed by Whittaker and Thompson (1975) or suggested by Romero-Herrera *et al.* (1973) for the primates or other mammals, and the PAM units of Dayhoff (1972) which statistically allow for hidden mutations.

The species compared were human (Romero-Herrera and Lehmann 1974*c*), chimpanzee (Romero-Herrera and Lehmann 1972*a*), gibbon (Romero-Herrera and Lehmann 1971), baboon and macaque (Romero-Herrera and Lehmann 1972*b*), treeshrew (Romero-Herrera and Lehmann 1974*b*), marmoset, woolly monkey and squirrel monkey (Romero-Herrera and Lehmann 1973*a*), potto (Romero-Herrera and Lehmann 1975), lemur and galago (Romero-Herrera and Lehmann 1973*b*), badger (Tetaert *et al.* 1974), horse (Dautrevaux *et al.* 1969), sheep (Han *et al.* 1972), ox (Han *et al.* 1970), porpoise and harbour seal (Bradshaw and Gurd 1969), dolphin (Kluh and Bakaradjieva 1971), sperm whale (Edmundson 1965), kangaroo (Air and Thompson 1971) and chicken (Deconinck *et al.* 1975).

Because of the importance of correct assignment of amide residues to the number of differences between species, the nature of residue 122 in kangaroo myoglobin has been checked and subsequently revised to an aspartyl residue. Checking seemed warranted following the revised results on residue 122 in sperm whale and horse myoglobin (Romero-Herrera and Lehmann 1974*a*), and the finding of aspartic acid in platypus. The similar mobility of peptides C2Ch7 (molecular weight

731, $m_{Lys} = 0{\cdot}39$) and C2Ch10 (molecular weight 760, $m_{Lys} = 0{\cdot}38$, Table 5) in the kangaroo work (Air and Thompson 1971) suggested an error in the original allocation. For this checking, tryptic peptide Tp19A comprising residues 119–132 was isolated for further digestion with chymotrypsin.

Results

As in previous papers the CNBr fragments are numbered from the *N*-terminal (CNBr1 to CNBr4). Tryptic (Tp), chymotryptic (Ch), peptic (Pe) and thermolysin (Th) peptides from each CNBr fragment are then numbered from the *N*-terminal of that fragment but using C instead of CNBr (e.g. C1Tp1, C2Th1). Subtilisin BPN′ is abbreviated Su.

Tryptic and chymotryptic peptides were also isolated from digests of myoglobin (Mb). The tryptic peptides are numbered from the *N*-terminal following the nomenclature for human myoglobin and the suggestions of Gerald and Ingram (1961). The symbols A and B have been used where extra lysyl residues have occurred in platypus myoglobin. When a resistant Lys–Pro bond is present two numbers have been given to the large peptide in human myoglobin and when lysine or arginine have not occurred in platypus myoglobin the peptide has been given all the numbers of the corresponding human peptides.

Residues identified as dansyl derivatives are set in *italic* fount.

Fractionation of Platypus Myoglobin

As was the case with kangaroo myoglobin (Air and Thompson 1971), the 70% ammonium sulphate extract of platypus muscle contains large quantities of protein in addition to myoglobin. Separation of this mixture on SP-Sephadex is shown in Fig. 1. A large peak of protein is not retained by the ion-exchanger and three haem-containing peaks are eluted after these. The major peak (2) is myoglobin whilst the most retarded peak (3) is haemoglobin. This is shown by the cellulose acetate electrophoresis experiments (Fig. 2). Predominantly single bands are given by peaks 2 and 3 in the absence of urea but in the presence of urea after removal of the haem the globin from peak 3 gives two bands corresponding to the α- and β-chains of platypus haemoglobin, whereas peak 2 gives a single band of apomyoglobin. The nature of peak 1 has not been investigated apart from its electrophoresis. There are at least four components in this minor fraction, which could reflect incomplete oxidation with $K_3Fe(CN)_6$, as well as changes in the myoglobin due to storage of the skeletal material or deamidation during the isolation procedures (cf. Romero-Herrera and Lehmann 1974*c*).

Chromatography of the major peak in the form of apomyoglobin on CM-cellulose in 8M urea–thiol buffers (Air and Thompson 1971) did not show significant heterogeneity by the elution curve measured at 280 nm or by amino acid analysis of the recovered material. The amino acid analyses of the major peak after hydrolysis under vacuum in 6M HCl at 105°C for 24, 48, 72 and 96 h are given in Table 1. No cysteinyl residues could be detected by the Ellman method. The *N*-terminal sequence by the dansyl–Edman procedure was Gly-Leu-.

Cyanogen Bromide Cleavage of Platypus Apomyoglobin

The amino acid analyses of platypus apomyoglobin (Table 1) show three methionyl residues, hence cyanogen bromide cleavage would be expected to yield four fragments.

The results of gel filtration of the treated material is shown in Fig. 3. Four fractions were bulked as indicated. The fractionation was similar to that previously obtained from cyanogen bromide-treated kangaroo myoglobin (Air and Thompson 1971) and this combined with amino acid analyses (Table 2) and peptide maps of chymotryptic digests suggested that fraction 1 consisted of CNBr1 and CNBr2 still linked together,

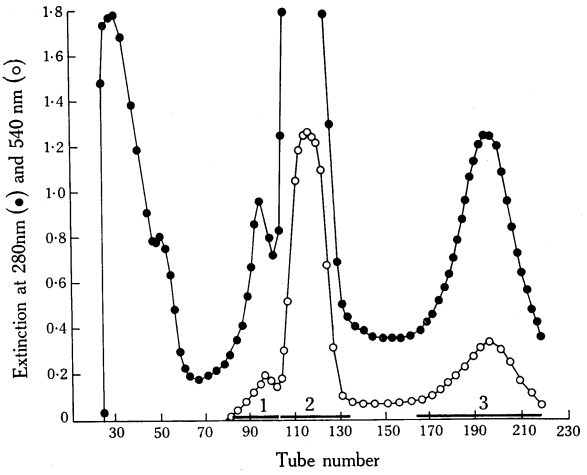

Fig. 1. Purification of platypus metmyoglobin on a column of SP-Sephadex (72 by 5·5 cm). The column was equilibrated and run in a K^+, Na^+ phosphate buffer, pH 6·4, ionic strength 0·1. The material loaded was the extract from 70% saturated ammonium sulphate treatment of muscle after oxidation with $K_3Fe(CN)_6$, concentration and dialysis against the chromatography buffer. Flow rate was 20 ml/h, fraction size 8 ml. Fractions bulked are shown by bars.

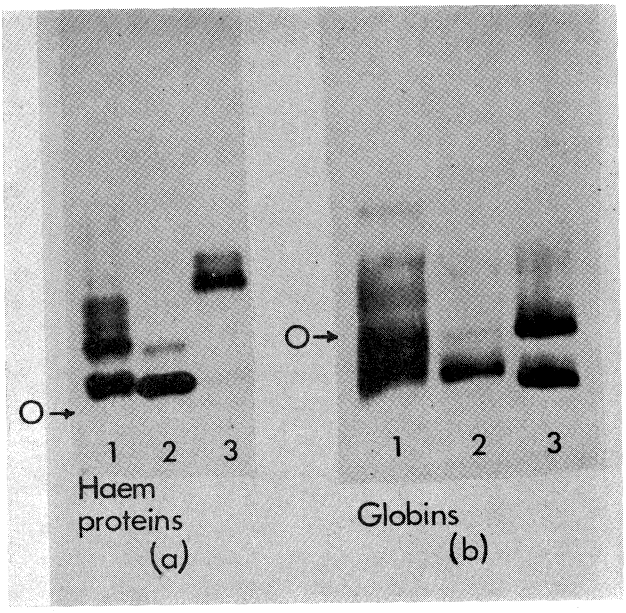

Fig. 2. Cellulose acetate electrophoresis at pH 8·6 (with migration towards the anode) of (*a*) the haem-containing fractions pooled from chromatograms of platypus metmyoglobin extracts. After removal of the haem, the resulting fractions (*b*) were run on cellulose acetate at pH 8·6 in buffers containing 8M urea (with migration towards the anode). In each case the origin O is shown.

or alternatively aggregated together by non-covalent forces. The *N*-terminal sequence of fraction 1 supported the former suggestion.

Fraction 2 with little absorption of light at 280 nm, but containing the largest weight of material, was CNBr2. Fraction 3 with strong 280 nm extinction was CNBr1, whilst fraction 4 contained a mixture of CNBr3 and CNBr4. The *N*-terminal

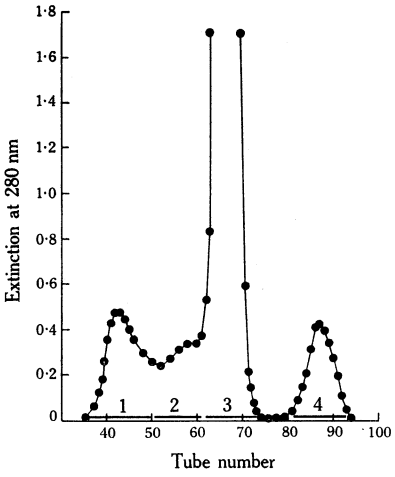

Fig. 3. Gel filtration of a cyanogen bromide digest of platypus apomyoglobin on a column of Sephadex G50 (68 by 2·3 cm) run in 10% formic acid. Flow rate was 20 ml/h, fraction size 3 ml. Fractions bulked are shown by bars.

Table 1. Amino acid composition of platypus myoglobin

Values are given as moles per mole of myoglobin calculated assuming a phenylalanine content of eight residues. Samples were hydrolysed at 105°C for various times under vacuum with 6M HCl containing 1 mg/ml phenol. The values have been averaged, extrapolated to zero time of hydrolysis for unstable amino acids or selected at the upper limit for slow release. Also shown is the amino acid composition determined by sequence

Amino acid	Hydrolysis for: 24 h	48 h	72 h	96 h	Mean or extrapolated value	Sequence
Lysine	20·1	21·0	20·9	21·4	20·8	21
Histidine	7·6	7·9	7·8	8·1	7·8	8
Arginine	2·0	2·0	2·0	2·2	2·0	2
Aspartic acid	10·4	10·4	10·3	9·6	10·2	10
Threonine	6·0	5·9	5·7	5·6	(6·1)	6
Serine	6·5	6·2	5·9	5·6	(6·8)	7
Glutamic acid	20·1	19·9	20·3	20·0	20·1	20
Proline	3·1	2·9	2·8	2·9	2·9	3
Glycine	15·3	15·9	15·3	15·3	15·4	15
Alanine	15·6	16·1	15·7	15·7	15·8	15
Valine	5·5	6·2	6·3	6·5	6·5	6
Methionine	3·0	3·0	3·1	3·0	3·0	3
Isoleucine	4·8	5·6	6·0	6·2	6·2	7
Leucine	16·6	17·1	17·2	17·2	17·2	18
Tyrosine	1·9	1·9	2·0	2·0	2·0	2
Phenylalanine	8·0	8·0	8·0	8·0	8·0	8
Tryptophan[A]					1·5	2
Total						153

[A] Determined by the spectrophotometric method of Beaven and Holiday (1952).

sequences by the dansyl–Edman method were Gly-Leu- for fraction 1, Lys-Ala- for fraction 2, Gly-Leu- for fraction 3 and two sequences Gly-Lys- and Ala-Ala- for fraction 4.

Table 2. Amino acid compositions of CNBr fractions of platypus apomyoglobin

The four fractions were isolated by gel filtration as described in the text, then hydrolysed in vacuum with 6M HCl at 105°C for times up to 150 h. Values are given as moles per mole based on the expected content of phenylalanine residues, and are means unless corrected for losses on hydrolysis (Ser, Thr) or slow rate of liberation (Val, Leu, Ile) in the usual way. Also shown are the amino acid compositions obtained from the amino acid sequence of CNBr fragments. n.d., Not determined

Amino acid	Fraction 1	CNBr1+ CNBr2	Fraction 2	CNBr2	Fraction 3	CNBr1	Fraction 4	CNBr3+ CNBr4
Trp	n.d.	2	n.d.	0	n.d.	2		
Lys[A]	16·1	18	10·1	11	6·8	7	3·0	3
His[A]	7·5	8	5·1	6	2·1	2		
Arg	1·0	1			0·9	1	0·9	1
Asp	8·4	8	4·1	4	4·1	4	1·9	2
Thr	6·4	6	3·0	3	2·9	3		
Ser	7·1	7	5·3	6	1·2	1		
Glu[B]	16·9	17	8·3	8	9·4	9	3·4	3
Pro	3·2	3	1·2	1	2·0	2		
Gly	12·0	12	5·1	5	7·1	7	3·0	3
Ala	12·7	12	11·9	12	0·7	0	3·1	3
Val[A]	5·2	6	2·0	2	4·2	4		
Hse	+	1–2	+	1	+	1	+	1
Ile[A]	4·8	7	4·7	6	1·1	1		
Leu[A]	14·0	16	7·5	8	7·9	8	2·0	2
Tyr	0·9	1	0·9	1	0	0	1·1	1
Phe	5·0	5	2·0	2	3·0	3	2·9	3
Totals		131		76		55		22

[A] There are several bonds resistant to hydrolysis in CNBr1 (2 Val–Leu, Ile–Arg) and CNBr2 (2 Val–Leu, Ile–Leu, Ile–Ile, Ile–Lys, Ile–His) that necessitate extensive hydrolysis.
[B] No special measures were taken to convert homoserine lactone to homoserine or to separate any homoserine from glutamic acid.

Amino Acid Sequence of CNBr1

The sequential analysis of CNBr1 was based on the alignment of overlapping peptides obtained by digestion with trypsin and the further digestion of the insoluble 'core' peptide material from a tryptic digest with chymotrypsin and pepsin. Supporting peptides were identified in chymotryptic and tryptic digests of myoglobin. The peptide map of tryptic peptides is shown in Fig. 4 and the overlapping peptides in Table 3. Tryptophan-containing peptides were identified by the Ehrlich reaction and the position of tryptophan inferred from the specificity of chymotrypsin and the absence of a dansyl derivative in the degradations. The tryptophyl residues were located in positions common to other myoglobins.

As seen in Table 3, the core from the tryptic digest contained a large amount of undigested or partially digested material that was soluble at the pH of digestion (pH 8·7) but not soluble in the limited amount of pH 6·4 buffer used to dissolve the digest before paper ionophoresis or in subsequent washing. Peptides Tp1A and

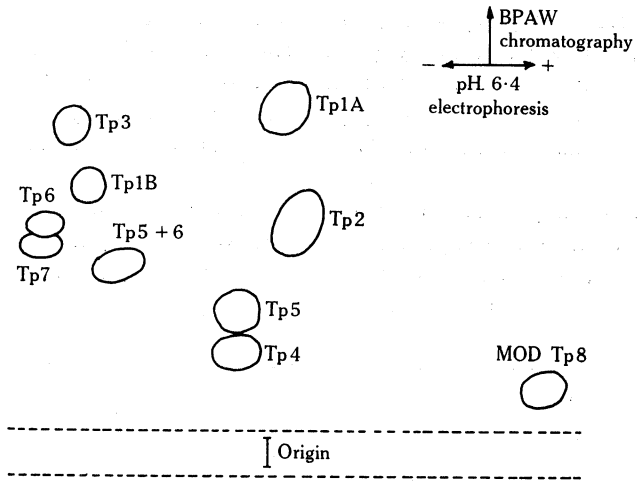

Fig. 4. Peptide map of the tryptic peptides of the CNBr1 fragment of platypus apomyoglobin. Ionophoresis at pH 6·4 was followed by chromatography with butanol–pyridine–acetic acid–water buffer (15 : 10 : 3 : 12 v/v) as indicated. The peptides are given the identification number corresponding to their position in the chain.

Table 3. Peptides isolated from enzymic digests of platypus apomyoglobin or its CNBr1 fragment
Tp, Ch, Pe and Th are tryptic, chymotryptic, peptic and thermolysin peptides respectively. Residues in *italic* fount were identified by the dansyl–Edman procedure. An arrow indicates a point of enzymic hydrolysis

	5	10 15
C1-Tp	*Gly-Leu-Ser-Asp-Gly-Glu*-Trp-Gln-Leu-Val-Leu-Lys-*Val*-Trp-*Gly-Lys-Val-Glu*-	
	Tp1A	↑ Tp1B ↑
Mb-Ch	*Gly Leu Ser Asp Gly Glu* Trp↑	↑*Lys Val* Trp↑*Gly Lys Val Glu*
C1-Tp 'core'-Ch	*Gly Leu Ser Asp Gly Glu* Trp↑	↑*Lys Val* Trp↑*Gly Lys Val Glu*
C1-Tp 'core'-Pe	*Gly Leu Ser Asp Gly Glu* Trp *Gln*↑*Leu Val Leu*↑*Lys Val* Trp *Gly Lys Val Glu*	
C1-Tp 'core'-Pe	*Gly Leu Ser Asp Gly Glu* Trp↑*Gln Leu Val Leu*↑	↑Gly Lys↑
	20 25 30	35
C1-Tp	*Gly-Asp-Leu-Pro-Gly-His-Gly-Gln-Glu-Val-Leu-Ile*-Arg-*Leu-Phe-Lys-Thr*-His-*Pro*-	
	Tp2	↑ Tp3 ↑
Mb-Ch	*Gly Asp Leu* Pro Gly His Gly Gln Glu Val Leu↑	↑*Lys Thr* His Pro
C1-Tp 'core'-Ch	*Gly Asp Leu Pro Gly His Gly Gln Glu Val Leu*↑*Ile* Arg↑	↑*Lys Thr* His *Pro*
C1-Tp 'core'-Pe	*Gly Asp Leu Pro Gly His Gly Gln Glu Val Leu*↑*Ile* Arg *Leu*↑*Phe Lys Thr* His *Pro*	
C1-Tp 'core'-Pe	↑*Ile* Arg↑	
	40 45 50	55
C1-Tp	*Glu-Thr-Leu-Glu*-Lys-*Phe-Asp-Lys-Phe-Lys-Gly-Leu-Lys-Thr-Glu-Asp-Glu-Hse*	
	Tp4 ↑ Tp5 ↑ Tp6 ↑ Tp7 ↑	Tp8*N*
Mb-Ch	Glu Thr Leu↑	↑*Asp Lys Phe*↑*Lys Gly Leu*↑*LysThr Glu Asp Glu Met*↑
C1-Tp 'core'-Ch	*Glu Thr Leu Glu* Lys↑*Phe Asp* Lys↑	↑*Thr Glu Asp* Glu Hse
C1-Tp 'core'-Pe	*Glu Thr Leu Glu Lys*↑*Phe Asp* Lys↑	
C1-Tp 'core'-Pe		

Tp1B were the least soluble tryptic peptides. In earlier work on kangaroo myoglobin (Air and Thompson 1971), sheep myoglobin (Han *et al.* 1972) and human myoglobin (Romero-Herrera and Lehmann 1974*c*) this same area 1–16 of CNBr1 was insoluble or only partially soluble, but there is a lysyl rather than asparaginyl residue at position 12 in platypus, so that two peptides Tp1A and Tp1B were obtained from platypus myoglobin.

The ordering of the peptides in CNBr1 was facilitated by the extensive homology with other myoglobins but is supported by deductive reasoning from the overlapping sequences and peptides Glu-Lys-Phe (from a chymotryptic digest of CNBr1),
41 43

Phe-Lys-Gly and *Leu-Lys*-Thr-Glu-Asp-Glu-Hse (from a thermolysin digest of
46 48 49 55
CNBr1).

Amino Acid Sequence of CNBr2

As found in previous work (Air and Thompson 1971), chymotrypsin is the preferred enzyme for fragmentation of CNBr2. The region corresponding to residues 98–118 was obtained as an insoluble 'core', a region of the molecule that is also insoluble in tryptic digests of myoglobins (cf. Romero-Herrera and Lehmann 1974*c*). This core was further digested with pepsin and thermolysin. Additional information on the CNBr2 56–131 region was obtained from peptides isolated from tryptic and chymotryptic digestion of apomyoglobin.

A peptide map of the soluble peptides from a chymotryptic digest of CNBr2 is given in Fig. 5. In Table 4 the peptides isolated and sequenced from the various digests are listed. The ordering of the peptides in CNBr2 was facilitated by the extensive homology with other myoglobins but is supported by deductive reasoning from the overlapping peptides and known enzyme specificities.

Amino Acid Sequence of CNBr3 and CNBr4

The amino acid sequences of these fragments were given by the peptides obtained by chymotryptic digestion of fraction 4 (Fig. 2) which contained both CNBr3 and CNBr4. In conjunction with the sequences of tryptic peptides from digests of apomyoglobin, the ordering of the peptides was clear and in agreement with the strong homology to other myoglobins. The relevant peptides are shown below.

```
                132          135                 140
C3              Gly-Lys-Ala-Leu-Glu-Leu-Phe-Arg-Asn-Asp-Hse
                    C3Ch1   ↑  C3Ch2  ↑     C3Ch3

Mb-Tp                  Ala-Leu-Glu-Leu-Phe-Arg-Asn-Asp-Met-Ala-Ala-Lys
                               Tp20        ↑        Tp21+22A

                    145                 150
C4              Ala-Ala-Lys-Tyr-Lys-Glu-Phe-Gly-Phe-Gln-Gly
                    C4Ch1   ↑  C4Ch2  ↑     C4Ch3

Mb-Tp                  Tyr-Lys-Glu-Phe-Gly-Phe-Gln-Gly
                       Tp22B ↑          Tp23
```

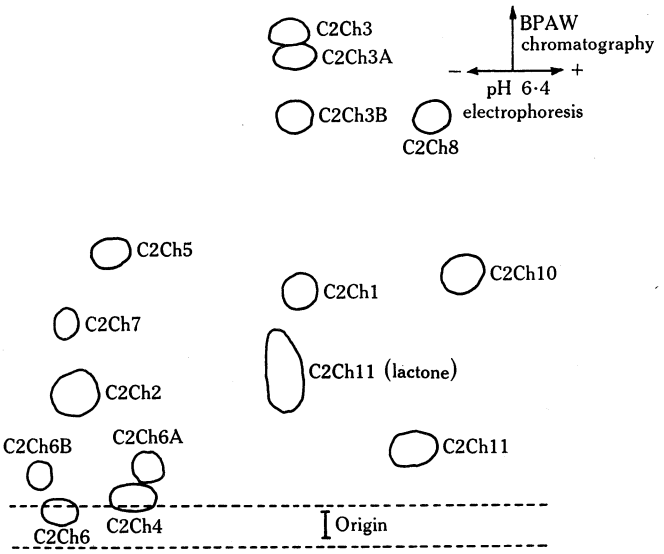

Fig. 5. Peptide map of the chymotryptic peptides of the CNBr2 fragment of platypus apomyoglobin. Fractionation as in Fig. 4. The peptides are given the identification number corresponding to their position in the chain.

Table 4. Peptides isolated from enzymic digests of platypus apomyoglobin or its CNBr2 fragment

Symbols are as in Table 3

	56	60	65	70	75
C2-Ch	*Lys-Ala-Ser-Ala-Asp-Leu-Lys-Lys*-His-*Gly-Gly-Thr-Val-Leu-Thr-Ala-Leu-Gly-Asn-Ile-*				
		C2Ch1 ↑	C2Ch2	↑	C2Ch3
Mb-Tp	↑*Ala Ser Ala Asp Leu* Lys Lys↑			Tp 'core'	
		80	85	90	95
C2-Ch	*Leu-Lys-Lys-Lys-Gly-Gln*-His-*Glu-Ala-Glu-Leu-Lys-Pro-Leu-Ala-Gln-Ser*-His-*Ala-Thr-*				
	↑	C2Ch4	↑	C2Ch5 ↑	C2Ch6
Mb-Tp	↑Lys *Lys Gly* Gln His *Glu Ala Glu* Leu↑Lys Pro Leu Ala Gln Ser His Ala Thr				
		100	105	110	115
C2-Ch	*Lys*-His-*Lys-Ile-Ser*-Ile-Lys-Phe-*Leu-Glu-Tyr*-Ile-Ser-Glu-Ala-Ile-Ile-His-Val-Leu-				
	↑	C2Ch7	↑ C2Ch8 ↑		C2Ch9
Mb-Tp	Lys↑His Lys↑*Ile Ser Ile Lys*↑			Tp 'core'	
C2 'core'-Pe	↑*Lys Ile Ser*↑		↑*Leu Glu Tyr*↑		↑*Ala Ile Ile* His *Val Leu*
			↑Leu Glu↑	↑*Ile Ser Glu*↑	
				↑*Tyr Ile Ser Glu*↑*Ala Ile Ile* His *Val Leu*	
C2 'core'-Th	*Lys Ile Ser*↑*Ile Lys Phe*↑*Leu Glu Tyr*↑*Ile Ser Glu Ala*↑*Ile Ile* His Val↑*Leu*				
		120	125	130	
C2-Ch	Gln-Ser-Lys-His-*Ser-Ala-Asp-Phe-Gly-Ala-Asp-Ala-Gln-Ala-Ala*-Hse				
		↑ C2Ch10 ↑		C2Ch11	
Mb-Tp	↑His Ser *Ala Asp Phe* Gly Ala Asp Ala Gln Ala Ala Met Gly Lys↑				
C2 'core'-Pe	*Gln Ser Lys His*↑*Ser Ala Asp Phe*↑				
	Gln Ser Lys His↑				
C2 'core'-Th	*Gln Ser* Lys↑				

The two peptides in fraction 4 (Fig. 2) could be separated by paper ionophoresis at pH 6·4 or pH 1·8 and identified by their amino acid compositions and *N*-terminal sequences using the dansyl–Edman procedure. CNBr3 gave a yellow colour with ninhydrin, due to the *N*-terminal glycyl residue.

Table 5. Amino acid compositions of peptides from platypus myoglobin

Peptides were isolated by paper ionophoresis at pH 6·4 and chromatography in butanol–acetic acid–water–pyridine. Hydrolysates were prepared with 6M HCl at 110°C for 24–48 h. Values are not corrected for losses during hydrolysis and are given as moles per mole of peptide with values from sequence data in parentheses. Detection by ninhydrin usually results in low recovery of the *N*-terminal amino acid. The enzyme used to give the peptide is indicated together with the residue numbers (cf. Tables 3 and 4)

Amino acid	Ch (62–69)	Ch (70–76)	Ch (77–89)	Tp (99–102)	Th (101–103)	Pe (106–109)	Ch-Pe (110–119)
Lys	2·0 (2)		3·9 (4)	1·0 (1)	0·6 (1)		1·0 (1)
His	1·0 (1)		1·0 (1)				1·8 (2)
Asp		1·0 (1)					
Thr	1·0 (1)	0·9 (1)					
Ser				0·8 (1)		0·8 (1)	1·0 (1)
Glu			3·2 (3)			1·1 (1)	1·2 (1)
Pro			0·8 (1)				
Gly	2·2 (2)	1·1 (1)	1·1 (1)				
Ala		1·1 (1)	1·2 (1)				0·7 (1)
Val	1·1 (1)						1·2 (1)
Ile		1·0 (1)		2·1 (2)	1·0 (1)	1·0 (1)	1·6 (2)
Leu	1·0 (1)	2·0 (2)	1·8 (2)				1·2 (1)
Tyr						0·4 (1)	
Phe					1·0 (1)		

Amino Acid Sequence of Platypus Myoglobin

In the deduction of the sequence of platypus myglobin it was not always possible, due to limited facilities, to quantitatively analyse each peptide. Paper ionophoresis at pH 1·8 (Dreyer and Bynum 1967) was used extensively. This does not unequivocally resolve the Val, Ser, Ile, Leu area but all of these amino acids are clearly identified as dansyl derivatives. A large number of dansyl–Edman degradations were performed and most of the residues were allocated by repeated identification of dansyl derivatives using different peptides containing the same segment of amino acids (Tables 3 and 4). Careful checking of peptides from the chymotryptic core region of CNBr2 was necessary because of the difficulty in achieving complete hydrolysis of isoleucyl and valyl bonds, and the effect of this on the amino acid composition of both CNBr1 and CNBr2 for valine, isoleucine and leucine. In these cases selected peptides were quantitatively analysed and the results are shown in Table 5.

Assignment of amide groups to side chains in CNBr1 etc. was made on the basis of ionophoretic mobility at pH 6·4 (Sanger *et al.* 1955; Ambler 1963; Offord 1966) and the peptides used in these assignments are listed in Table 6.

The complete amino acid sequence of platypus myoglobin is shown in Table 7. There is good agreement between the amino acid composition of myoglobin and its cyanogen bromide fragments as found by analysis (Tables 1 and 2) and by sequence.

Revision of Kangaroo Sequence for Residue 122

The tryptic peptide Tp19A was not isolated in the original work, and it is one residue shorter than the corresponding peptide from platypus. It was isolated from a peptide map and had a mobility, $m_{Asp} = 0 \cdot 12$, that was similar to the Tp19 spot from platypus and indicated only a fractional net negative charge at pH 6·4. The

Table 6. Assignment of side chain amide and acidic groups in platypus myoglobin

These groups were established wherever possible from the ionophoretic mobility of small peptides at pH 6·4 (Sanger *et al.* 1955; Offord 1966). The residue numbers of the peptides used are given in parentheses, and whether they are neutral (n), negatively (−) or positively (+) charged; the enzyme used for digestion is also given. Some of the peptides were minor products not otherwise discussed in the text. The running pH of the paper is less than the pH of the soaking buffer (pH 6·4) and the histidyl residues carry almost a full positive charge (75–90%)

Residue No.	Residue		Peptide and net charge
4	Asp	(−)	Tp1A (1–12)
6	Glu	(−)	Tp1A (1–12)
8	Gln	(n)	Pe (8–11)
18	Glu	(−)	Ch (17–24)
20	Asp	(−)	Ch (17–24)
26	Gln	(n)	Ch (25–26), Su (19–26)
27	Glu	(−)	Su (27–29)
38	Glu	(n)	Tp4 (35–42)
41	Glu	(n)	Tp4 (35–42)
44	Asp	(n)	Tp5 (43–45)
52	Glu	(−)	C1-Tp (51–55) $m_{Asp} = 0 \cdot 9$
53	Asp	(−)	C1-Tp (51–55) $m_{Asp} = 0 \cdot 9$
54	Glu	(−)	C1-Tp (51–55) $m_{Asp} = 0 \cdot 9$
60	Asp	(n)	Tp9 (57–62), C2Ch1 (56–61)
74	Asn	(n)	C2Ch3 (70–76)
81	Gln	(+)	Su (79–81)
83	Glu	(−)	Su (80–85)
85	Glu	(−)	Th (84–86), (n) Su (84–89)
91	Gln	(+)	C2Ch (90–93)
105	Glu	(−)	C2Ch8 (104–106)
109	Glu	(−)	C2-Pe (107–109)
116	Gln	(+)	C2-Th (115–118)
122	Asp	(−)	C2Ch10 (120–123)
126	Asp	(−)	C2Ch11-Th (124–128)
128	Gln	(n)	C2Ch11-Th (127–128)
136	Glu	(−)	C3Ch2 (136–138)
140	Asn	(n)	C3Ch3 (139–142), Tp (140–145), C3Ch3 (141–142),
141	Asp	(−)	residues 139, 140 removed by C_6H_5NCS
148	Glu	(n)	C4Ch2 (147–149)
152	Gln	(n)	C4Ch3 (150–153)

peptide contains an *N*-terminal histidyl residue and its p*K* (imidazoyl) will probably be lowered by the positively charged neighbouring amino group. In addition there is one lysyl residue so that at least two negatively charged side chains of acidic amino acid residues would be necessary to give an overall net negative charge. Chymotryptic digestion followed by peptide fractionation gave a neutral spot at pH 6·4 which

gave a yellow colour with ninhydrin, corresponding in composition to the sequence Gly-Ala-Asp-Ala-Gln-Ala-Ala-Met-Lys and an acidic peptide corresponding in
124 132

composition to His-Ala-Gly-Asp-Phe. This peptide had a fractional net negative
119 123

charge at pH 6·4, $m_{Asp} = 0{\cdot}22$, due to the incomplete charge on the imidazoyl group. This data confirms that residue 122 is aspartic acid rather than asparagine in kangaroo myoglobin.

Table 7. Amino acid sequence of platypus myoglobin

Residues are numbered from the *N*-terminal residue and the positions of cleavage by CNBr are indicated together with the positions of cleavage of CNBr1 by trypsin and CNBr2 by chymotrypsin

```
                5                  10                  15                  20
Gly-Leu-Ser-Asp-Gly-Glu-Trp-Gln-Leu-Val-Leu-Lys-Val-Trp-Gly-Lys-Val-Glu-Gly-Asp-Leu-Pro-
                    C1Tp1A                      ↑    C1Tp1B  ↑
        25                  30                  35                  40
Gly-His-Gly-Gln-Glu-Val-Leu-Ile-Arg-Leu-Phe-Lys-Thr-His-Pro-Glu-Thr-Leu-Glu-Lys-Phe-Asp-
  C1Tp2                          ↑  C1Tp3  ↑          C1Tp4                  ↑  C1Tp5
 45                  50                  55                  60                  65
Lys-Phe-Lys-Gly-Leu-Lys-Thr-Glu-Asp-Glu-Met-Lys-Ala-Ser-Ala-Asp-Leu-Lys-Lys-His-Gly-Gly-
   ↑C1Tp6↑  C1Tp7  ↑       C1Tp8N      ↑         C2Ch1         ↑          C2Ch2
                                      CNBr
             70                  75                  80                  85
Thr-Val-Leu-Thr-Ala-Leu-Gly-Asn-Ile-Leu-Lys-Lys-Lys-Gly-Gln-His-Glu-Ala-Glu-Leu-Lys-Pro-
                ↑       C2Ch3          ↑              C2Ch4                  ↑  C2Ch5
     90                  95                 100                 105                 110
Leu-Ala-Gln-Ser-His-Ala-Thr-Lys-His-Lys-Ile-Ser-Ile-Lys-Phe-Leu-Glu-Tyr-Ile-Ser-Glu-Ala-Ile-Ile-
   ↑              C2Ch6          ↑       C2Ch7        ↑  C2Ch8  ↑              C2Ch9
         115                 120                 125                 130
His-Val-Leu-Gln-Ser-Lys-His-Ser-Ala-Asp-Phe-Gly-Ala-Asp-Ala-Gln-Ala-Ala-Met-Gly-Lys-Ala-
                        ↑   C2Ch10   ↑              C2Ch11              ↑
                                                                      CNBr
135                 140                 145                 150
Leu-Glu-Leu-Phe-Arg-Asn-Asp-Met-Ala-Ala-Lys-Tyr-Lys-Glu-Phe-Gly-Phe-Gln-Gly
      CNBr3                    ↑              CNBr4
                             CNBr
```

In studying hydrolysates of peptides from a tryptic digest of kangaroo myoglobin, an acidic peptide was located that analysed $His_{0{\cdot}8}\,Arg_{0{\cdot}9}\,Asp_{2{\cdot}0}\,Thr_{0{\cdot}9}\,Glu_{2{\cdot}0}\,Gly_{3{\cdot}0}\,Val_{2{\cdot}9}\,Ile_{1{\cdot}0}\,Leu_{1{\cdot}0}$, as though the lysine detected at residue 26 in the previous work had been replaced by glutamine as in the platypus. The previous assignment was based on the analysis and sequence of two peptides C1Tp2 and C1Tp3 and was supported by the analysis of the CNBr1 fragment. Dansyl–Edman degradation showed that the sequence in the new myoglobin peptide was Val-Glu-Thr-Asp-Val-
17

Gly-Gly-His-Gly-Gln-Asp-Val-Leu-Ile-Arg with a valine replacing glutamic acid at
31

residue 21 in addition to the glutamine for lysine at residue 26. No myoglobin from the early work was available and it is possible that this variation in sequence is due to individual animal genotype. Further animals must be investigated before the normal sequence is defined.* It is of some interest that residue 21 is usually hydrophobic (isoleucine, leucine or valine) and only valine mutates to glutamic acid by a single base change. This B2 helix position is also usually occupied by a hydrophobic residue in α- and β-globins but there is a glutamyl residue in horse β-chain.

Table 8. Matrices of differences between myoglobins

The matrix of observed differences between myoglobin sequences compared over 153 residues is shown in the upper triangle. PAM units (per 100 links) are shown in the lower triangle

	Man	Chimpanzee	Gibbon	Baboon	Macaque	Tree shrew	Marmoset	Potto	Woolly monkey	Squirrel monkey	Sport lemur	Galago	Badger	Horse	Sheep	Ox	Porpoise	Dolphin	Sperm whale	Harbour seal	Kangaroo	Platypus	Chicken
Man		1	1	6	7	13	14	17	16	17	22	23	17	18	25	29	20	21	24	24	21	25	35
Chimpanzee	1		2	7	8	14	15	18	17	18	21	24	18	17	24	28	19	20	23	23	22	26	35
Gibbon	1	1		5	6	14	13	18	15	16	23	24	18	19	26	30	21	22	25	25	22	26	36
Baboon	4	5	3		1	12	11	16	13	12	21	22	16	17	25	29	22	22	25	23	21	23	37
Macaque	5	5	4	1		12	10	15	12	12	22	21	16	17	25	29	22	22	24	23	21	23	37
Tree shrew	9	10	10	8	8		17	14	17	18	15	16	14	15	21	24	19	18	22	19	21	21	36
Marmoset	10	11	9	8	7	12		19	4	4	18	23	15	16	24	28	24	24	26	23	21	25	38
Potto	12	13	13	11	11	10	14		21	23	20	13	19	15	25	27	21	19	21	23	28	25	38
Woolly monkey	11	12	11	9	8	12	3	15		4	20	21	18	18	25	29	26	25	28	25	21	25	38
Squirrel monkey	12	13	11	8	8	13	3	17	3		19	23	18	18	26	30	27	26	30	26	22	25	39
Sport lemur	16	15	17	15	16	11	13	14	14	14		22	19	14	21	24	19	17	25	18	25	23	39
Galago	17	18	18	16	15	11	17	9	15	17	16		24	21	27	30	25	25	26	28	28	24	38
Badger	12	13	13	11	11	10	11	14	13	13	14	18		17	24	25	21	22	27	19	22	28	39
Horse	13	12	14	12	12	11	11	11	13	13	10	15	12		17	19	16	13	19	20	26	25	45
Sheep	18	18	19	18	18	15	18	18	18	19	15	20	18	12		6	23	18	28	27	32	32	43
Ox	22	21	22	22	22	18	21	20	22	22	18	22	18	14	4		23	17	31	28	36	36	45
Porpoise	14	14	15	16	16	14	18	15	19	20	14	18	15	11	17	17		10	15	17	29	27	39
Dolphin	15	14	16	16	16	13	18	14	18	19	12	18	16	9	13	12	7		18	19	30	28	38
Sperm whale	18	17	18	18	18	16	19	15	21	22	18	19	20	14	21	22	11	13		26	31	30	42
Harbour seal	18	17	18	17	17	14	17	17	18	19	13	21	14	14	20	21	12	14	19		27	25	38
Kangaroo	15	16	16	15	15	15	15	21	15	16	18	21	16	16	25	25	22	19	23	20		24	35
Platypus	18	19	19	17	17	15	18	18	18	18	17	18	21	18	25	28	20	21	22	18	18		37
Chicken	27	27	28	29	29	28	31	31	31	31	31	31	31	27	36	38	31	31	35	31	26	29	

Estimation of the Platypus Date of Divergence

The total numbers of amino acid differences between all myoglobins of known amino acid sequence are shown in Table 8. It is seen that platypus myoglobin shows very similar differences to kangaroo myoglobin in many cases, which is a reflection of the highly conserved structure and makes this protein less suitable than the α- and β-globin chains for evolutionary rate calculations. The amino acid differences were converted to PAM units (number of accepted point mutations per 100 links of two sequences) by the table given in Dayhoff (1972).

By use of the PAM units shown in Table 8 (omitting the platypus comparisons) and the dates of divergence estimates that come from the palaeontological record, the average time for one accepted point mutation to occur per 100 residues of one

* *Note added in proof.* A preparation of myoglobin from another red kangaroo showed residue 21 to be valine and residue 26 glutamine and this is probably the most common sequence.

chain was calculated to be $10{\cdot}58 \pm 0{\cdot}36$ million years (95% confidence interval of the mean, $n = 231$).

By application of the mean to the PAM units obtained by comparing platypus myoglobin with all the other mammalian sequences, a range of estimates for the date of divergence of the platypus was obtained. The mean of this range was 102 million years with a 95% confidence interval of 71–131 million years ($n = 21$).

Discussion

The method of Hapner *et al.* (1968) affords a convenient method for the preparation of myoglobin. There was predominantly a single myoglobin in the extracts of platypus skeletal muscle and a unique amino acid sequence was found. The sequence work on myoglobins is aided by this purity when compared with work on α- and β-globin preparations that can cross contaminate each other.

The amino acid sequence of platypus myoglobin shows 25 changes from the sequence of human myoglobin and 24 changes from that of kangaroo myoglobin—considerably less than was the case with platypus α-chain compared with human and kangaroo α-chains (39 and 46 changes respectively), or platypus β-chain compared with human and kangaroo β-chains (34 and 30 changes respectively). This is in keeping with previous reports of conservation of residues in myoglobin, particularly those residues that help maintain myoglobin as a monomeric molecule (cf. Romero-Herrera and Lehmann 1974*c*).

The three-dimensional structure of sperm whale myoglobin has been determined (Kendrew *et al.* 1960; Perutz *et al.* 1965; Watson 1969) as well as its amino acid sequence (Edmundson 1965). The amino acid residues which appear to be involved in the folding of the chain and interaction with the haem are either identical in platypus myoglobin or replaced by a very similar residue. Thus, of the residues listed by Perutz (1969) as interacting with the haem group, myoglobins from sperm whale, human and platypus are identical except for residue CD3(45) which is arginine in sperm whale but lysine in the other two, and residue G4(103) which is phenylalanine in platypus but tyrosine in the other two.

Internal hydrophobic side chains are identical except for residue B9(28) which is isoleucine in sperm whale but valine in human and platypus, residue F1(86) which is leucine in sperm whale and platypus but isoleucine in human, residue H19(142) which is isoleucine in sperm whale but methionine in the other two, and residue G11(110) which is alanine in sperm whale and platypus but lysine in human myoglobin. These substitutions maintain essential contact sites.

Of the numerous conserved hydrogen-bonding residues, B16(35) is an exception with glycine in human but serine and threonine in sperm whale and platypus respectively, and CD6(48) is glycine in platypus but histidine in sperm whale and human myoglobin resulting in a loss of bonding of this residue in platypus myoglobin to the side chain of lysine CD3(45).

Of the polar interactions only residue 59(E2) in platypus has lost its negative charge (Glu→Ala) but Watson (1969) showed only a weak interaction for this site. It seems likely that the three-dimensional structure of platypus myoglobin will closely follow that for sperm whale myoglobin.

When the amino acid sequence of platypus myoglobin is compared with the sequence of kangaroo myoglobin and other myoglobins of known sequence the

glutamyl residue at G17(116) is common to the monotreme, marsupial and all primate myoglobins (except chimpanzee which has the histidine which is common to myoglobins of most other placental mammals). Of the other substitutions there are only two positions that could indicate a similar line of descent: E17(74) asparagine and G4(103) phenylalanine which are identical in marsupial and monotreme myoglobin but different in human and sperm whale myoglobin.

There is only one change in platypus myoglobin at a residue position previously constant in known myoglobins. At position E2(59) an alanyl residue replaces the glutamyl residue for which the side chain is listed as being salt-linked to lysine E5(62).

From a consideration of the number of changes in amino acid sequences of globin chains when comparing monotreme globins with those from different mammalian species, an estimate of the average rate of mutation can be obtained. This type of calculation was introduced by Zuckerkandl and Pauling (1962) and it was hoped that application of this method to the amino acid sequences of marsupial and monotreme proteins, for which there is no adequate fossil record, would give a reliable estimate of the date of divergence from eutherian mammals. The earlier data with marsupial globin chains (Air *et al.* 1971) appeared promising and appeared to support the idea of a statistically constant rate of change, even though many reasons why this would seem unlikely can be advanced.

The estimates of the dates of divergence of monotremes from eutherian mammals for the α-, β- and myoglobin chains are:

	Million years	
α-chain	211 ± 88	
β-chain	132 ± 33	
Myoglobin	102 ± 31	(platypus only)

Although the ranges of these estimates overlap, the differences are most striking. The discrepancy between the estimates is probably due to varying rates of mutation of the different chains in particular lineages, a conclusion reached by other workers (e.g. Dayhoff 1972; Romero-Herrera *et al.* 1973; Goodman *et al.* 1975), suggesting that it is not valid to assume a constant rate of mutation in making divergence point estimations. This is a disappointing conclusion, which is not altered by using the minimum number of nucleotide replacements (Whittaker 1975) and allowing for multiple mutations at particular sites.

Acknowledgments

This work was supported in part by the Australian Research Grants Committee. We are indebted to Mr T. R. Grant and Mr F. N. Carrick for the platypus muscle, to Mr R. G. Mann for skilled technical assistance and to Dr Janet K. Allen for assistance with computing.

References

Air, G. M., and Thompson, E. O. P. (1969). *Aust. J. Biol. Sci.* **22**, 1437.
Air, G. M., and Thompson, E. O. P. (1971). *Aust. J. Biol. Sci.* **24**, 75.
Air, G. M., Thompson, E. O. P., Richardson, B. J., and Sharman, G. B. (1971). *Nature (London)* **229**, 391.
Ambler, R. P. (1963). *Biochem. J.* **89**, 349.
Anson, M. L., and Mirsky, A. E. (1930). *J. Gen. Physiol.* **13**, 469.

Beard, J. M., and Thompson, E. O. P. (1971). *Aust. J. Biol. Sci.* **24**, 765.
Beaven, G. H., and Holiday, E. R. (1952). *Adv. Protein Chem.* **7**, 319.
Berggard, I. (1961). *Ark. Kem.* **18**, 291.
Bradshaw, R. A., and Gurd, F. R. N. (1969). *J. Biol. Chem.* **244**, 2167.
Dautrevaux, M., Boulanger, Y., Han, K., and Biserte, G. (1969). *Eur. J. Biochem.* **11**, 267.
Dayhoff, M. O. (1972). 'Atlas of Protein Sequence and Structure'. (National Biochemical Research Foundation: Silver Spring.)
Deconinck, M., Peiffer, S., Depreter, J., Paul, C., Schnek, A. G., and Leonis, J., (1975). *Biochim. Biophys. Acta* **386**, 567.
Dreyer, W. J., and Bynum, E. (1967). 'Methods in Enzymology'. (Ed. C. H. W. Hirs.) Vol. 11. p. 32. (Academic Press: New York.)
Edmundson, A. B. (1965). *Nature (London)* **205**, 883.
Ellman, G. L. (1959). *Arch. Biochem. Biophys.* **82**, 70.
Gerald, P. S., and Ingram, V. M. (1961). *J. Biol. Chem.* **236**, 2155.
Goodman, M., Moore, G. W., and Matsuda, G. (1975). *Nature (London)* **253**, 603.
Griffiths, M. (1968). 'Echidnas'. (Pergamon Press: Oxford.)
Gross, E., and Witkop, B. (1962). *J. Biol. Chem.* **237**, 1856.
Han, K., Dautrevaux, M., Chaila, X., and Biserte, G. (1970). *Eur. J. Biochem.* **16**, 465.
Han, K., Tetaert, D., Moschetto, Y., Dautrevaux, M., and Kopeyan, C. (1972). *Eur. J. Biochem.* **27**, 585.
Hapner, K. D., Bradshaw, R. A., Hartzell, C. R., and Gurd, F. R. N. (1968). *J. Biol. Chem.* **243**, 683.
Kendrew, J. C. *et al.* (1960). *Nature (London)* **185**, 422.
Kluh, I., and Bakaradjieva, A. (1971). *FEBS Lett.* **17**, 31.
Nash, A. R., and Thompson, E. O. P. (1974). *Aust. J. Biol. Sci.* **27**, 607.
Offord, R. E. (1966). *Nature (London)* **211**, 591.
Perutz, M. F. (1969). *Proc. R. Soc. B.* **173**, 113.
Perutz, M. F., Kendrew, J. C., and Watson, H. C., (1965). *J. Mol. Biol.* **13**, 669.
Romero-Herrera, A. E., and Lehmann, H. (1971). *Biochem. Biophys. Acta* **251**, 482.
Romero-Herrera, A. E., and Lehmann, H. (1972*a*). *Biochem. Biophys. Acta* **278**, 62.
Romero-Herrera, A. E., and Lehmann, H. (1972*b*). *Biochim. Biophys. Acta* **278**, 465.
Romero-Herrera, A. E., and Lehmann, H. (1973*a*). *Biochim. Biophys. Acta* **317**, 65.
Romero-Herrera, A. E., and Lehmann, H. (1973*b*). *Biochim. Biophys. Acta* **322**, 10.
Romero-Herrera, A. E., and Lehmann, H. (1974*a*). *Biochim. Biophys. Acta* **336**, 318.
Romero-Herrera, A. E., and Lehmann, H. (1974*b*). *Biochim. Biophys. Acta* **359**, 236.
Romero-Herrera, A. E., and Lehmann, H. (1974*c*). *Proc. R. Soc. B.* **186**, 249.
Romero-Herrera, A. E., and Lehmann, H. (1975). *Biochim. Biophys. Acta* **393**, 205.
Romero-Herrera, A. E., Lehmann, H., Joysey, K. A., and Friday, A. E. (1973). *Nature (London)* **246**, 389.
Sanger, F., Thompson, E. O. P., and Kitai, R. (1955). *Biochem. J.* **59**, 509.
Tetaert, D. *et al.* (1974). *Biochim. Biophys. Acta* **351**, 317.
Thompson, E. O. P., and Air, G. M. (1971). *Aust. J. Biol. Sci.* **24**, 1199.
Watson, H. C. (1969). In 'Progress in Stereochemistry'. (Eds. B. J. Aylett and M. M. Harris.) Vol. 4. p. 299. (Butterworths Scientific Publications: London.)
Whittaker, R. G. (1975). Ph.D. Thesis, University of New South Wales.
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.
Whittaker, R. G., and Thompson, E. O. P. (1974). *Aust. J. Biol. Sci.* **27**, 591.
Whittaker, R. G., and Thompson, E. O. P. (1975). *Aust. J. Biol Sci.* **28**, 353.
Zuckerkandl, E., and Pauling, L. (1962). In 'Horizons in Biochemistry'. (Eds M. Kasha and B. Pullman.) p. 189. (Academic Press: New York.)

Manuscript received 15 September 1975