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

Studies on Monotreme Proteins. VI.* Amino Acid Sequence of the β-Chain of Haemoglobin from the Platypus, *Ornithorhynchus anatinus*

R. G. Whittaker and E. O. P. Thompson

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

Abstract

The amino acid sequence of the 146 residues of the β-chain of the major haemoglobin from the platypus has been determined. The soluble peptides derived from the chain by tryptic digestion were isolated by paper ionophoresis and chromatography. The amino acid sequences were determined by the dansyl–Edman procedure or by further digestion with other enzymes. The tryptic peptides were aligned by homology with other β-globins. There were 14 changes in sequence compared with echidna β-chain.

The number of changes in sequence compared with human β-chain is 34 which is less than the 39 changes between human and platypus α-chains. Generally there are more changes between β-chains; there are only three other examples reported where there are more changes between α-chains than β-chains, these are of echidna, rabbit and dog globins.

By comparison with the 'contact sites' in horse haemoglobin there is one change in β-haem contacts, three changes in β_1–α_1 contacts and no changes in β_2–α_1 contacts.

The date of divergence of the monotremes from the other mammals was estimated at 132 ± 33 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-chain sequences and the significance of this discrepancy to the validity of the method is discussed.

Introduction

In the previous paper in this series (Whittaker and Thompson 1974) the fractionation of the haemoglobins and the globin chains from the major haemoglobin of the platypus was reported together with the amino acid sequence of the α-chain. In the present paper the complete amino acid sequence of the β-chain is given.

Since there are no significant fossil remains of the monotremes (Young 1962) no accurate palaeontological estimate of the date of divergence from the rest of the mammals is possible. In the previous paper the number of differences between α-globin chains, after allowing for hidden mutations, and palaeontological estimates of dates of divergence were used to derive an estimate of the time of divergence of the monotremes. In this paper these calculations are repeated, using β-chain sequences, to give a second estimate of the divergence point and to test the validity of the necessary assumption of a constant mutation rate.

Materials and Methods

The *S*-carboxymethylated β-chain of platypus haemoglobin was prepared as described previously (Whittaker and Thompson 1974).

* Part V, *Aust. J. Biol. Sci.*, 1974, **27**, 591–605.

The methods of high-voltage ionophoresis, peptide mapping, amino acid analysis, cyanogen bromide cleavage, citraconylation, sequence determination by the dansyl–Edman procedure and digestion with trypsin, chymotrypsin and thermolysin were the same as previously described (Air and Thompson 1969, 1971; Beard and Thompson 1971; Whittaker *et al.* 1972; Whittaker and Thompson 1974).

Isolation of Tryptic Peptides

No insoluble fraction was obtained after tryptic digestion of the β-globin at pH 8·7 and standing the digestion mixture at 4°C for 3 days. The digestion mixture was then freeze-dried and pyridine–acetate buffer (pH 6·4) added. A gelatinous precipitate formed slowly on standing at 4°C. The material soluble at pH 6·4 was peptide-mapped (Thompson *et al.* 1969).

The gelatinous pH 6·4 insoluble material was freeze-dried and 3·5% formic acid (pH 1·8) added. The soluble material was fractionated by ionophoresis at pH 1·8 and the small amount of insoluble material was taken for analysis.

Estimation of Dates of Divergence

The β-chain sequence of the platypus and the β-chain sequence of the echidna (Whittaker *et al.* 1972) were used to estimate a date of divergence of the monotremes from the other mammalian groups in the same manner as previously reported for the α-chain sequences (Whittaker and Thompson 1974).

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 suggested by Romero-Herrera *et al.* (1973), and the accepted point mutations of Dayhoff (1972) which statistically allow for hidden mutations.

The species compared were human, rabbit, dog, horse, B bovine, grey kangaroo, potoroo, echidna, platypus, chicken AII and frog. The β-chain sequences for a number of primates, sheep and red kangaroo are also available in a sufficiently completed form to use in the calculations but they were omitted to avoid undue emphasis of their particular lineages. The sequences are given by Dayhoff (1972) except for the potoroo and red kangaroo (Thompson and Air 1971), the echidna (Whittaker *et al.* 1972), and the chicken (Matsuda *et al.* 1973).

Results

As in previous papers tryptic peptides are numbered from the *N*-terminus (see Fig. 1) following the nomenclature for the human β-chain (Gerald and Ingram 1961).

Fractionation of Tryptic Peptides

The peptide map for the whole tryptic digest, obtained by paper ionophoresis–chromatography, is shown in Fig. 2.

All peptides except for βTp12B occurred in the peptide map. βTp15 was purified and separated from βTp10A by paper ionophoresis at pH 1·8. βTp5 was isolated in low yield, being largely insoluble at pH 6·4.

Isolation of βTp12B and βTp5

After freeze-drying of the pH 6·4 insoluble fraction, 3·5% formic acid (pH 1·8) was added. The soluble fraction mainly contained βTp5 which was further purified by ionophoresis at pH 1·8. The insoluble fraction at pH 1·8 contained βTp12B in a highly impure state.

A pure sample of βTp12B was obtained from a tryptic digest of citraconylated β-chain at pH 8·7. Initially there was no precipitate at pH 8·7. After freeze-drying, pyridine–acetate buffer (pH 6·4) was added and the material insoluble at pH 6·4 was decitraconylated in 5% formic acid, freeze-dried and redigested with trypsin

at pH 8·7. Peptide βTp12B, which was substantially pure, precipitated out. Presumably peptide(s) responsible for augmenting the solubility of βTp12B at pH 8·7 in previous digests, probably by hydrophobic interactions, had been removed in the pH 6·4 soluble fraction of the digested citraconylated β-chain.

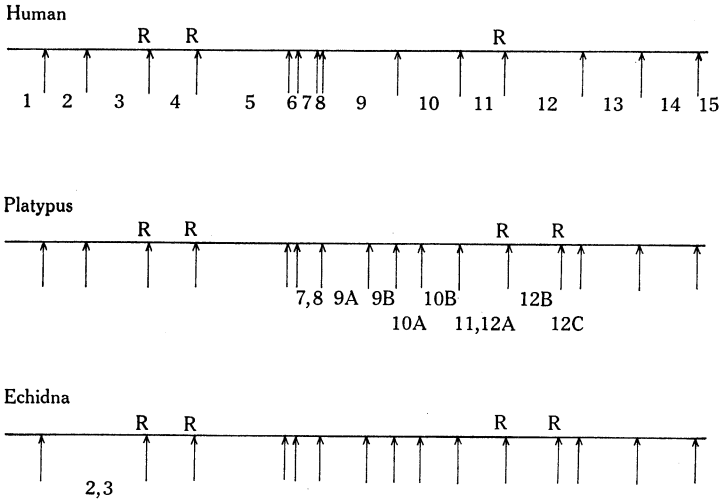

Fig. 1. Nomenclature for β-chain tryptic peptides of the major haemoglobin of the platypus compared with the β-chain tryptic peptides of human haemoglobin. The distribution of lysyl and arginyl residues in the echidna-I(B) β-chain is shown to indicate the strong similarities between both monotremes. The nomenclature of the monotreme β-chains follows that of the human β-chain. Where extra lysyl or arginyl residues have occured the letters A, B and C have been utilized. When a lysyl or arginyl residue has not occurred the peptide has been given all the numbers of the corresponding human peptides. The diagrams are scaled to show the relative number of residues in each peptide. ↑ Lysyl residue. ^{R}↑ Arginyl residue.

Amino Acid Composition of Tryptic Peptides

The amino acid composition of purified peptides is shown in Table 2. The total compositions are in good agreement with the total analysis of the β-chain (Table 1).

Amino Acid Sequences of Tryptic Peptides

All residues identified by the dansyl–Edman procedure are presented in *italic* fount. The alignment of the peptides has been based on homology with previously sequenced β-chains.

βTp1

The sequence of this peptide was

Val-His-*Leu-Ser-Gly-Gly-Glu-Lys.*
1 8

No dansyl derivative was found at position 2 and as all other residues were identified the histidinyl residue was assigned to this position. The peptide was basic on iono-

phoresis at pH 6·4 with the mobility [R_{Asp} = +0·21 (Offord 1966)] indicating a glutamic acid residue at position 7.

βTp2

The sequence of this peptide, which gave a positive Ehrlich reaction for tryptophan, was

Ser-Ala-Val-Thr-Asn-Leu-Trp-*Gly-Lys*.
9 17

No dansyl derivative was found at position 15 and as all other residues were identified the tryptophan residue was assigned to this position. The peptide was basic on ionophoresis at pH 6·4, indicating the presence of an asparaginyl residue at position 13.

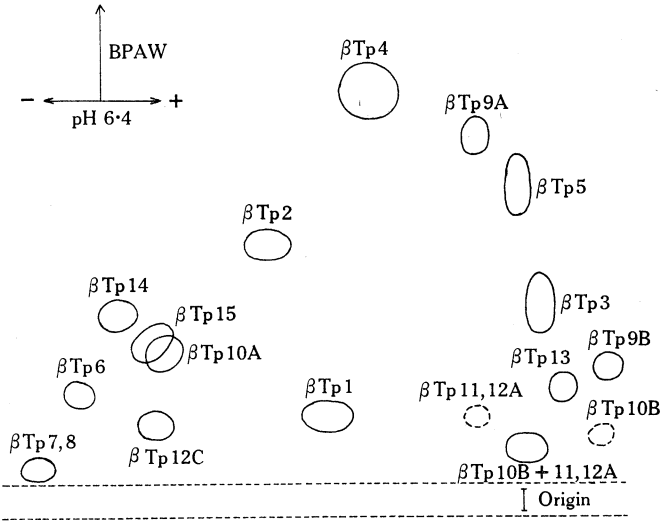

Fig. 2. Peptide map of the tryptic peptides of the β-chain of platypus globin. Ionophoresis at pH 6·4 was followed by chromatography with butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12 v/v) as indicated. The peptides are given the identification number corresponding to their position in the chain as shown in Fig. 1.

βTp3

This acidic peptide sequenced directly as

Val-Asn-Ile-Asn-Glu-Leu-Gly-Gly-Glu-Ala-Leu-Gly-Arg.
18 30

Thermolysin digestion gave four major peptides which were isolated by paper ionophoresis at pH 6·4 followed by paper chromatography. The compositions of the hydrolysed peptides were determined by ionophoresis at pH 1·8. The sequences of the peptides were: *Val-Asn*, neutral at pH 6·4; *Ile-Asn-Glu*, acidic at pH 6·4 with the mobility (R_{Asp} = −0·47) indicating one acidic residue—free glutamic acid was isolated by ionophoresis at pH 1·8 after removal of the first two residues by the dansyl–Edman procedure; *Leu-Gly-Gly-Glu-Ala*, acidic at pH 6·4; *Leu-Gly*-Arg, basic at pH 6·4.

βTp4

The sequence of this basic peptide, which gave a positive Ehrlich's reaction for tryptophan, was

Leu-Leu-Val-Val-Tyr-Pro-Trp-*Thr-Gln*-Arg.
31 40

Direct degradation gave

Leu-Leu-Val-Val-Tyr-Pro.

From a chymotryptic digestion the basic peptide *Thr-Gln*-Arg was isolated. Tryptophan was assigned to position 37 by difference.

Table 1. Amino acid composition of platypus β-globin

Samples were hydrolysed at 110°C for 24 h under vacuum with 6N HCl containing phenol at a concentration of 1 mg/ml. Values are given as moles per mole of protein and are uncorrected for losses on hydrolysis or incomplete hydrolysis. Also shown is the amino acid composition determined by sequence data

Amino acid	Platypus globin Hydrolysate	Sequence	Amino acid	Platypus globin Hydrolysate	Sequence
Lysine	12·5	12	Glycine	15·3	15
Histidine	7·6	8	Alanine	13·8	14
Arginine	4·0	4	Valine	12·0[B]	15
SCM-cysteine	0·8	1	Methionine	0·9	1
Aspartic acid	15·0	15	Isoleucine	1·3[B]	2
Threonine	3·7[A]	4	Leucine	17·0[B]	19
Serine	7·8[A]	9	Tyrosine	2·0	2
Glutamic acid	10·0	10	Phenylalanine	8·0	8
Proline	4·2	4	Tryptophan	2·9[C]	3

[A] Hydrolysis of a sample at 104°C for 24 h gave 3·9 threonine and 9·1 serine residues relative to 8 phenylalanine residues.
[B] Hydrolysis of a sample at 110°C for 72 h gave 14·6 valine, 1·9 isoleucine and 18·1 leucine residues relative to 8 phenylalanine residues. The slow release of these residues being due to three Val-Leu, two Val-Val and one Ile-Val peptide linkages in the platypus sequence.
[C] Determined by the spectrophotometric method of Beavan and Holiday (1952).

βTp5

The complete sequence of this acidic peptide was

Phe-Phe-Glu-Ala-Phe-Gly-Asp-Leu-Ser-Ser-Ala-Gly-Ala-Val-Met-Gly-Asn-Pro-Lys.
41 59

Direct degradation gave

Phe-Phe-Glx-Ala-Phe-Gly-Asx-Leu-Ser-Ser.

Cleavage at the methionine residue by digestion with cyanogen bromide gave rise to two peptides: CNBr1, acidic at pH 6·4, composition on ionophoresis at pH 1·8 consistent with a methionine residue at position 55. Mobility on ionophoresis at pH 6·4 indicated two acidic residues ($R_{Asp} = -0·33$); CNBr2, basic at pH 6·4, analysed and sequenced as *Gly-Asn-Pro-Lys.*

Thermolysin digestion was used to complete the sequence. The peptides identified were *Phe-Gly-Asp*, acidic at pH 6·4; *Leu-Ser-Ser*; *Ala-Gly-Ala*; *Val-Met-Gly-Asn-*

Table 2. Amino acid composition of platypus β-globin tryptic peptides

Soluble peptides were purified by paper ionophoresis at pH 6·4 followed by paper chromatography. Hydrolysates were made at 110°C for 24 h. Values are not corrected for losses on hydrolysis or incomplete hydrolysis and are given as moles per mole of peptide, with preferred values in parentheses. Detection by ninhydrin usually results in low recovery of the *N*-terminal amino acid residues

Amino acid	βTp1	βTp2[A]	βTp3	βTp4[A]	βTp5[B]	βTp6	βTp7,8	βTp9A	βTp9B	βTp10A	βTp10B +11,12A[C]	βTp12B	βTp12C	βTp13[A]	βTp14	βTp15	Total in protein
Lys	1·0 (1)	1·0 (1)			1·0 (1)	1·0 (1)	0·9 (1)	0·9 (1)	1·1 (1)	0·9 (1)	0·9 (1)	0·1	1·0 (1)	1·0 (1)	1·0 (1)		12
His	0·8 (1)						·) (1)				1·7 (2)	0·1	0·6 (1)		1·9 (2)	1·0 (1)	8
Arg			1·0 (1)	1·0 (1)							0·9 (1)	0·9 (1)					4
SCM-Cys											0·7 (1)						1
Asp		0·9 (1)	2·0 (2)		2·0 (2)			0·9 (1)	3·0 (3)		3·8 (4)	1·2 (1)		1·0 (1)			15
Thr		0·8 (1)		0·9 (1)				0·8 (1)		1·0 (1)							4
Ser	0·9 (1)	0·8 (1)			1·8 (2)			0·9 (1)			1·0 (1)	0·2	0·9 (1)	1·0 (1)	0·9 (1)		9
Glu	1·0 (1)		2·0 (2)	1·1 (1)	1·1 (1)						2·2 (2)	0·2		3·2 (3)			10
Pro				1·0 (1)	0·7 (1)						1·1 (1)			1·0 (1)			4
Gly	2·1 (2)	1·0 (1)	3·0 (3)		2·9 (3)		1·0 (1)	1·1 (1)		1·0 (1)		1·0 (1)			2·1 (2)		15
Ala		1·1 (1)	1·1 (1)		2·9 (3)		2·0 (2)	1·1 (1)		1·0 (1)		1·0 (1)		2·1 (2)	2·1 (2)		14
Val	0·8 (1)	1·1 (1)	1·0 (1)	2·0 (2)	1·2 (1)	1·0 (1)		0·9 (1)			1·0 (1)	2·2 (3)		1·0 (1)	2·0 (2)		15
Met					+ve (1)												1
Ile			0·9 (1)									0·5 (1)					2
Leu	1·0 (1)	1·0 (1)	1·9 (2)	1·4 (2)	1·1 (1)			1·7 (2)	2·0 (2)		2·6 (3)	2·6 (3)			1·9 (2)		19
Tyr				0·9 (1)												0·9 (1)	2
Phe					2·9 (3)			1·0 (1)		1·0 (1)	1·1 (1)	0·2	1·0 (1)	1·0 (1)			8
Trp		+ve (1)		+ve (1)										+ve (1)			3
Total	8	9	13	10	19	2	5	10	6	5	18	11	4	12	12	9	146

[A] Tryptophan detected by the Ehrlich reagent.

[B] Following ionophoresis at pH 6·4 and paper chromatography methionine is always detected as the sulphone.

[C] This overlapping peptide was the major product of tryptic digestion.

Pro-Lys, basic at pH 6·4. The *N*-terminal peptide was not isolated from the digest but the mobility of the intact peptide at pH 6·4 together with the mobility of the CNBr1 fragment indicated a glutamic acid residue at position 43.

βTp6

The sequence of this peptide was

Val-Lys.
60 61

βTp7,8

The sequence of this peptide was

Ala-His-*Gly-Ala*-Lys.
62 66

No dansyl derivative was found at position 63 and as all other residues were identified histidine was assigned to this position.

βTp9A

This was a neutral peptide at pH 6·4 indicating the presence of an aspartic acid residue. The dansyl–Edman procedure gave the sequence

Val-Leu-Thr-Ser-Phe-Gly-Asp-Ala-Leu-Lys.
67 76

βTp9B

This peptide, which stained yellow-brown with ninhydrin, was acidic at pH 6·4, indicating the presence of at least two aspartic acid residues. The PTH-derivative was isolated and identified at step 1 of the dansyl–Edman procedure to give the sequence

Asn-Leu-Asp-Asp-Leu-Lys.
77 82

βTp10A

This peptide sequenced directly as

Gly-Thr-Phe-Ala-Lys.
83 87

βTp10B+11,12A

As has been the case with other β-globins, this was the major peptide containing *βTp10B* and *βTp11,12A*. An aspartic acid residue preceding lysine 95 results in incomplete cleavage of the Lys-Leu bond.

95 96

Direct degradation gave

Leu-Ser-Glx-Leu-X-SCMCys-Asx-Lys-Leu.

Thermolysin digestion gave five major peptides which when analysed either by ionophoresis at pH 1·8 or with the amino acid analyser, sequenced as: *Leu-Ser-Glu*, acidic at pH 6·4 ($R_{Asp} = -0{\cdot}41$); *Leu-His-SCMCys-Asp-Lys*, neutral at pH 6·4;

Leu-His; *Val-Asp-Pro-Glu-Asn*, mobility at pH 6·4 indicating two acidic residues ($R_{Asp} = -0{\cdot}59$)—dansyl Asn was obtained at the last sequencing step (dansyl derivative not hydrolysed); *Phe-Asn*-Arg, basic at pH 6·4.

The complete sequence, together with the bonds hydrolysed by thermolysin, was

Leu-Ser-Glu-Leu-His-SCMCys-Asp-Lys-Leu-His-Val-Asp-Pro-Glu-Asn-Phe-Asn-Arg.
88 ↑ ↑ ↑ ↑ 105

βTp12B

This insoluble peptide sequenced directly as

Leu-Gly-Asn-Val-Leu-Ile-Val-Val-Leu-Ala-Arg.
106 116

Dansyl dipeptides were major products of hydrolysis at positions 109, 110, 111, 112 and 113. Isolation of the neutral peptide *Leu-Gly-Asn* from a thermolysin digestion of *βTp12B* indicated the presence of an asparaginyl residue at position 108.

βTp12C

The sequence of this peptide was

His-*Phe-Ser*-Lys,
117 120

the histidinyl residue being placed by difference.

βTp13

The sequence of this acidic, tryptophan positive peptide was

Asp-Phe-Ser-Pro-Glu-Val-Gln-Ala-Ala-Trp-*Gln*-Lys.
121 132

As all other residues were identified on direct degradation the tryptophanyl residue was assigned to position 130 by difference.

Thermolysin digestion was used to determine the amide distribution, the following peptides being isolated: *Asp-Phe-Ser-Pro-Glu*, acidic at pH 6·4 with mobility indicating two acidic residues ($R_{Asp} = -0{\cdot}57$); *Val-Gln*, neutral at pH 6·4; *Ala-Ala*-Trp-*Gln*-Lys, basic at pH 6·4—this peptide also gave a positive Ehrlich reaction for tryptophan.

βTp14

This peptide gave a clear sequence with the dansyl–Edman procedure except for the steps where histidine became *N*-terminal. These residues were placed by difference to give the sequence

Leu-Val-Ser-Gly-Val-Ala-His-*Ala-Leu-Gly*-His-Lys.
133 144

βTp15

The sequence of this peptide, which was the same as the *C*-terminal peptide for all known β-chains, was

Tyr-His.
145 146

Complete Sequence

The complete amino acid sequence of the platypus β-chain is shown in Table 3.

Estimation of the Monotreme Date of Divergence

The β-chain sequences compared are shown in Table 4. The chains were scanned using the computer method of Beard (1971) to give the number of amino acid differences between chains. These data were then used to determine PAM units (number of accepted point mutations per 100 links of two sequences) by the information given

Table 3. Complete amino acid sequence of the β-chain of platypus haemoglobin

Residues are numbered from the *N*-terminal residue and the positions of cleavage by trypsin are shown by arrows. The dashed arrow indicates a site of incomplete cleavage. The 14 residues that differ from those present in echidna Hb-1(B) β-chain are in capital letters

```
                5                  10                  15                  20
Val-His-Leu-Ser-Gly-GLY-Glu-Lys-SER-Ala-Val-Thr-Asn-Leu-Trp-Gly-LYS-Val-Asn-ILE-Asn-
              βTp1              ↑               βTp2              ↑

               25                  30                  35                  40
Glu-Leu-Gly-Gly-Glu-Ala-Leu-Gly-Arg-Leu-Leu-Val-Val-Tyr-Pro-Trp-Thr-Gln-Arg-Phe-Phe-Glu-
          βTp3                  ↑                  βTp4                  ↑

     45                  50                  55                  60
ALA-Phe-Gly-Asp-Leu-Ser-Ser-Ala-GLY-Ala-Val-Met-Gly-Asn-PRO-Lys-Val-Lys-Ala-His-Gly-
                          βTp5                                  ↑ βTp6 ↑          βTp7,8

 65                  70                  75                  80                  85
Ala-Lys-Val-Leu-Thr-Ser-Phe-Gly-Asp-Ala-Leu-Lys-Asn-Leu-Asp-ASP-Leu-Lys-Gly-Thr-Phe-Ala-
       ↑                βTp9A                ↑          βTp9B          ↑       βTp10A
            90                  95                 100                 105
Lys-Leu-Ser-Glu-Leu-His-Cys-Asp-Lys-Leu-His-Val-Asp-Pro-Glu-Asn-Phe-Asn-Arg-Leu-Gly-Asn-
  ↑          βTp10B          ⇡                βTp11,12A                  ↑

    110                 115                 120                 125
Val-Leu-ILE-Val-Val-Leu-Ala-Arg-His-Phe-Ser-Lys-ASP-Phe-SER-Pro-Glu-VAL-Gln-Ala-Ala-
      βTp12B                ↑  βTp12C    ↑                              βTp13

130                 135                 140                 145
Trp-Gln-Lys-Leu-Val-Ser-Gly-Val-ALA-His-Ala-Leu-GLY-His-Lys-Tyr-His
           ↑                          βTp14                        ↑ βTp15
```

by Dayhoff (1972). By use of these units shown in Table 4 (omitting the monotreme comparisons) and the dates of divergence estimates (Table 5) the average time for one accepted point mutation to occur per 100 residues of one chain was calculated to be $8 \cdot 75 \pm 1 \cdot 14$ million years (95% confidence interval of the mean, $n = 35$).

By application of the mean to the PAM units obtained by comparing the monotremes with all the other mammalian sequences a range of estimates for the date of divergence of the monotremes was obtained. The mean of this range was 132 million years with a 95% confidence interval of 99–165 million years ($n = 14$). This estimate differs widely from the previously reported α-chain comparison estimate (Whittaker and Thompson 1974) of 143–217 million years. The α-chain calculations have been repeated to allow for the inclusion of the opossum (Stenzel 1974) and the viper α-chains (Duguet *et al.* 1974). The divergence points of chicken and carp

have also been changed to 295 million years (previously 220) and 425 million years (350) respectively to suit the information given by Romer (1966).

The average time for one accepted point mutation to occur per 100 residues of α-chain then recalculates to be $9{\cdot}68 \pm 1{\cdot}06$ million years ($n = 44$) and using this mean the estimate of the monotreme divergence point from the rest of the mammals becomes 123–299 million years ($n = 21$). The wide statistical variation is due to the high mutational rate of the opossum α-chain.

Table 4. Matrices of differences between β-globins

The matrix of observed amino acid differences between β-chain sequences compared over 146 residues is shown in the upper triangle. PAM units (per 100 links) are shown in the lower triangle

	Human	Rabbit	Dog	Horse	Bovine	Grey kangaroo	Potoroo	Echidna	Platypus	Chicken AII	Frog
Human		14	15	26	25	38	42	31	34	45	67
Rabbit	11		21	26	30	37	38	31	33	48	64
Dog	11	16		31	29	36	36	31	36	42	65
Horse	20	20	25		32	44	43	37	36	48	69
Bovine	20	24	23	26		44	48	38	41	55	68
Grey kangaroo	33	31	31	39	39		15	39	40	47	73
Potoroo	37	33	31	38	44	11		35	37	45	71
Echidna	25	25	25	31	33	34	29		14	42	65
Platypus	28	27	31	31	36	35	31	11		42	69
Chicken AII	40	44	37	44	53	43	40	37	37		71
Frog	72	67	69	76	74	83	79	69	76	79	

Table 5. Palaeontological estimates of dates of divergence

Divergence point	Estimated time of divergence (million years)
Frog–birds, mammals	350[A]
Chicken–mammals	295[A]
Marsupials–eutherian mammals	104[B]
General radiation of eutherian mammals	68[B]
Horse–bovine	55[B]

[A] Romer (1966). [B] Romero-Herrera *et al.* (1973).

Note: No reliable estimate of the kangaroo–potoroo date of divergence is available. The comparison of these species was omitted from the calculations.

Using the modified α-chain calculations the average estimate for the divergence of the platypus from the echidna becomes 73 million years while using the β-chain calculation the estimate is 48 million years.

Discussion

The length of the β-globin chain of the platypus is 146 residues. The distribution of trypsin-sensitive basic amino acids is similar to the distribution in the echidna

β-chain with the exception of a lysyl residue replacing an histidinyl residue at position 17; lysine commonly occurs in this position. The unusual change which occurs in the G-helix of the echidna β-globin also occurs in the platypus. That is, an asparaginyl residue is present at residue 104(G6) which has been previously occupied by either lysyl or arginyl residues. The next position (G7), previously leucine in all β-chains, is occupied by an arginyl residue. The rest of the G-helix is homologous to other β-chains. The effects of these changes and a possible mutational mechanism for its occurrence has been presented previously (Whittaker *et al.* 1972).

Nine other amino acid changes have occurred that have not been previously listed in the 'Atlas of Protein Sequence and Structure' (Dayhoff 1972). Six of these also occur in the β-chain of the echidna and are: asparagine at position 13, a variable position often with a small polar, non-ionizable residue; asparagine at position 21, a variable position which is usually occupied by an acidic residue; leucine at position 23, usually occupied by valine; serine at position 119, usually glycine; serine at position 135, usually threonine; histidine at the variable position 139 usually occupied by an asparaginyl or seryl residue. All of these changes appear conservative and could arise through a single base mutation.

Of the three other amino acid substitutions peculiar to the platypus, two would appear conservative. Position 6, usually occupied by glutamic acid, proline or alanine, is glycine in the platypus and serine in the echidna. Position 142, which is usually alanine, is glycine in the platypus.

The remaining substitution, alanine at position 44, is of considerable interest because it is a contact point between the β-chain and the haem group. This position is usually occupied by serine or histidine, and in the horse, interaction is between the hydroxyl group of serine and the carboxyl group of the propionic acid side chain from ring III. This interaction is one of the only two polar interactions between the β-chain and the haem group. However, the substitution of an alanine at this position may be of little consequence as 'all polar contacts are in water so that their contribution to the binding energy should be small' (Perutz *et al.* 1968).

There are no differences from horse β-chain in the β–α_2 contacts. For the β–α_1 contacts there are variations in three positions, all of which have been reported in other β-chain sequences. These are alanine at position 51 (proline in horse), serine at position 119 (glycine) and serine at position 123 (threonine).

As the platypus is semi-aquatic and the echidna is terrestrial it would seem reasonable to anticipate some differences in haemoglobin characteristics such as oxygen dissociation curves, Bohr effects and the effect of 2,3-diphosphoglycerate. However, the sequences of the α- and β-chains from the major haemoglobins of both animals (Whittaker *et al.* 1972, 1973; Whittaker and Thompson 1974) show no changes from the horse sequences in the residues involved in the Bohr effect or in the binding site of 2,3-diphosphoglycerate. The α_1–β_2 contact seems mainly responsible for haem–haem interaction (Perutz *et al.* 1968) but there is only one change in the residues involved in this contact with both animals having serine at position 44 (CD2) of the α-chain sequence. This residue is only involved in the deoxyhaemoglobin contact (Bolton and Perutz 1970) and is normally occupied by a prolyl residue.

If differences in oxygen uptake characteristics between the haemoglobins of the two monotremes do occur then they must be due to changes in other areas such as α-haem, β-haem or α_1–β_1 (or all of these) contact sites.

The calculations of a date of divergence of the monotremes from the other mammalian groups using the platypus β-chain and the echidna β-chain previously reported (Whittaker *et al.* 1972) gives an estimate of 132 ± 33 million years, the approximate time of the divergence of the marsupials from the placental mammals.

This estimate differs widely from that given when α-chain sequences are considered. The α-chain calculations give a divergence date of 211 ± 88 million years, well before the divergence point of the marsupials.

This discrepancy between estimates is due to the slower mutational rate of the platypus and echidna β-chains. Usually the mutational rate for β-chains is faster than for α-chains (Air *et al.* 1971). This slower rate of mutation of the β-chain has previously been observed only in comparisons between the globin chains of the dog and rabbit. Either the α-chains or the β-chains or both of the platypus and echidna have mutated at an atypical rate to give the results obtained. Also many workers have demonstrated that different lineages mutate at varying rates (Goodman *et al.* 1971; Dayhoff 1972; Romero-Herrera *et al.* 1973; Goodman *et al.* 1975), suggesting that the assumption of a constant rate of mutation, necessary for divergence point estimations, may not be valid.

Calculations based on the minimum number of nucleotide replacements required to give the observed sequence differences do not add to the accuracy of the method (Beard 1971) and, on preliminary investigation, consideration of hypothetical ancestral sequences (Dayhoff 1972) does not seem promising.

Acknowledgments

This work was supported in part by the Australian Research Grants Committee. The authors are indebted to Mr T. R. Grant and Mr F. N. Carrick for samples of platypus blood and to Mr R. G. Mann for skilled technical assistance.

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 (Lond.)* **229**, 391.
Beard, J. M. (1971). Ph.D. Thesis, University of New South Wales.
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.
Bolton, W., and Perutz, M. F. (1970). *Nature (Lond.)* **228**, 551.
Dayhoff, M. O. (1972). 'Atlas of Protein Sequence and Structure.' (National Biomedical Research Foundation: Silver Spring.)
Duguet, M., Chauvet, J. P., and Acher, R. (1974). *FEBS Lett.* **47**, 333.
Gerald, P. S., and Ingram, V. M. (1961). *J. Biol. Chem.* **236**, 2155.
Goodman, M., Barnabas, J., Matsuda, G., and Moore, G. W. (1971). *Nature (Lond.)* **233**, 604.
Goodman, M., Moore, G. W., and Matsuda, G. (1975). *Nature (Lond.)* **253**, 603.
Matsuda, G., Maita, T., Mizuno, K., and Ota, H. (1973). *Nat. New Biol.* **244**, 244.
Offord, R E. (1966). *Nature (Lond.)* **211**, 591.
Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. G. (1968). *Nature (Lond.)* **219**, 131.
Romer, A. S. (1966). 'Vertebrate Palaeontology.' (University of Chicago Press.)
Romero-Herrera, A. E., Lehmann, H., Joysey, K. A., and Friday, A. E. (1973). *Nature (Lond.)* **246**, 389.
Stenzel, P. (1974). *Nature (Lond.)* **252**, 62.

Thompson, E. O. P., and Air, G. M. (1971). *Aust. J. Biol. Sci.* **24**, 1199.
Thompson, E. O. P., Hosken, R., and Air, G. M. (1969). *Aust. J. Biol. Sci.* **22**, 449.
Whittaker, R. G., and Thompson, E. O. P. (1974). *Aust. J. Biol. Sci.* **27**, 591.
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.
Young, J. Z. (1962). 'The Life of Vertebrates.' (Oxford University Press: London).

Manuscript received 10 February 1975