STUDIES ON MARSUPIAL PROTEINS

VI.* EVOLUTIONARY CHANGES IN β-GLOBINS OF THE MACROPODIDAE AND THE AMINO ACID SEQUENCE OF β-GLOBIN FROM *POTOROUS TRIDACTYLUS*

By E. O. P. THOMPSON† and G. M. AIR†

[Manuscript received May 27, 1971]

Abstract

Haemoglobins from 13 species of the subfamily Macropodinae and one species of the subfamily Potoroinae have been studied.

The red kangaroo *Megaleia rufa* contains a single haemoglobin which has been separated into the α- and β-chains. The soluble β-chain peptides were isolated and the amino acid compositions determined. The only difference between these peptides and those of the grey kangaroo *Macropus giganteus* was in peptide βTp5B. Amino acid sequence analysis of this peptide showed that the alanyl residue β56 in the grey kangaroo is replaced by a glycyl residue in the red kangaroo. Three insoluble tryptic peptides were also purified and analysed. For two of them the analysis showed some contamination with the third insoluble peptide and their compositions were calculated by difference. From these peptide compositions and the analyses of the β-chains it is concluded that there is probably only the one amino acid sequence difference between the β-globins of the two species.

The βTp5B peptides from 11 other macropodine species were isolated and the amino acid sequences determined. All species contained methionine at position β55, with alanine, glycine, serine, or threonine at position β56. The potoroo *Potorous tridactylus* had leucine at position β55.

The amino acid sequence of β-globin from *P. tridactylus* was determined. There were 15 differences in amino acid sequence from that of the grey kangaroo. Calculations based on the number of changes in the β-chain sequence suggest that the Potoroinae diverged from the Macropodinae about 50 million years ago, whereas the macropodine species have undergone a comparatively recent radiation.

I. INTRODUCTION

In a previous paper (Air and Thompson 1969) the amino acid sequence of the β-chain of haemoglobin from *Macropus giganteus* Shaw, 1790, was presented. This species has two allelic β-globin chains with glutamine or histidine as residue β2. This paper describes comparative studies on haemoglobins from 13 other species of the family Macropodidae.

The red kangaroo *Megaleia rufa* (Desmarest, 1822) is a closely related species to the grey kangaroo and the amino acid composition and peptide maps of the α- and β-globin chains, amino acid composition of β-chain tryptic peptides, and partial sequence of the β-globin have been determined. Only one difference in sequence was detected, located in tryptic peptide βTp5B which contains the methionine residue of the β-chain. This peptide from 11 other species of macropods has been isolated and sequenced.

* Part V, *Aust. J. biol. Sci.*, 1971, **24**, 765–86.

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

The β-globin of the potoroo *Potorous tridactylus* (Kerr 1792) does not contain methionine and amino acid analysis of the β-chain showed several changes from the macropodine species. The complete amino acid sequence obtained from the sequences of the tryptic peptides of this β-chain is presented.

Calculations based on the number of differences between the amino acid sequences of β-globin from the grey kangaroo and the potoroo give a probable date of their divergence as separate species.

II. Materials and Methods

(*a*) *Preparation of Haemoglobins and Globin Chains*

The haemoglobins and the α- and β-chains were prepared by carboxymethylcellulose chromatography as previously described (Thompson, Hosken, and Air 1969). In general a gradient of 0·01M Na^+ to 0·05M Na^+ was used to separate globin α- and β-chains. The separated chains were converted to the *S*-carboxymethylated derivatives.

For chromatographic separation of the haemoglobins of the potoroo the haemolysate was dialysed against 0·05M Tris–HCl at pH 8 and chromatographed on a column (16 cm by 2·8 cm diam.) of DEAE-Sephadex (A50) equilibrated against the same buffer. The haemoglobins were fractionated by a linear gradient of 0·05M Tris buffer from pH 8·0 to pH 7·0 (300 ml each chamber). After the first peak had been eluted a further 250 ml of pH 7·0 buffer completed the elution of the second component.

(*b*) *Enzyme Digestions and Purification of Peptides*

Digestions with TPCK-trypsin and chymotrypsin were as previously described (Thompson, Hosken, and Air 1969; Air and Thompson 1969). Digestions with thermolysin followed the method of Thompson, Sleigh, and Smith (1971).

Individual peptides were isolated from large-scale (100–150 mg) tryptic digests loaded directly onto a 1 by 120 cm column of Sephadex G25 or G50 and eluted with 0·01N ammonia or 1% ammonium bicarbonate, pH 9·1. The fractions were measured at 230 and 280 nm and the peak tubes combined and dried under vacuum. Each fraction was then dispersed in pH 6·4 pyridine–acetate buffer and any insoluble material isolated by centrifugation and washing with the buffer.

Soluble peptides were isolated from each fraction, or from the soluble portion of small-scale digests, by paper ionophoresis at either pH 6·4 or pH 1·75 followed by chromatography in n-butanol–acetic acid–water–pyridine (15 : 3 : 12 : 10 v/v) as used previously (Air and Thompson 1969, 1971). Peptides were eluted from the paper with 50% pyridine.

(*c*) *Amino Acid Analyses and Amino Acid Sequence Determination*

Amino acid analyses were carried out as previously described (Thompson, Hosken, and Air 1969). Alkaline hydrolysates for the analysis of tryptophan were done under nitrogen as described by Beard and Thompson (1970).

Amino acid sequences were determined by the "dansyl"–Edman method (Gray 1967) with the modifications of Air and Thompson (1969, 1971).

(*d*) *Cyanogen Bromide Cleavage*

This was carried out by the method previously used (Thompson, Hosken, and Air 1969) at a concentration of 20 mg globin/ml in 70% formic acid. Small samples were taken for end-group assay by the dansyl procedure.

In the case of *S*-carboxymethylated β-globin from the red kangaroo the reaction mixture was loaded on a column of Sephadex G50, equilibrated, and developed with 10% formic acid. Tubes were read at 280 nm and the pooled fractions freeze-dried.

III. Results

As in previous papers tryptic peptides are numbered from the *N*-terminus. Cyanogen bromide fragments are numbered from the *N*-terminus (CNBr1, CNBr2).

(*a*) *The Globins of the Red Kangaroo*

The red kangaroo (four animals studied) gave only a single haemoglobin, which chromatographed on polycarboxylic acid resins in the same position as the Hb-II of the grey kangaroo (Thompson, Hosken, and Air 1969). Richardson (1970) has chromatographed haemolysates from 88 different animals and observed only a single haemoglobin. The globin chains were readily separated using 8M urea–thiol buffers (Clegg, Naughton, and Weatherall 1965) with a gradient from 0·01M Na^+ to 0·05M Na^+.

Table 1

AMINO ACID COMPOSITION OF THE β-CHAIN OF HAEMOGLOBIN FROM THE RED KANGAROO, *MEGALEIA RUFA*

Values given are moles per mole of globin calculated assuming a β-chain length of 146 residues containing eight phenylalanine residues. Hydrolysis was done at 105°C

Amino Acid	24 hr	48 hr	72 hr	140 hr	Average*	Grey β-II†
Lysine	12·6	12·6	13·3	12·7	13	13
Histidine	7·9	7·9	7·8	7·9	8	8
Arginine	1·8	1·9	1·9	1·8	2	2
SCM-cysteine	1·8	1·8	1·7	1·7	2	2
Aspartic acid	14·3	14·6	14·6	14·4	15	15
Threonine	5·5	4·1	3·1	1·5	(7)	7
Serine	3·2	2·0	1·1	0·3	(4)	4
Glutamic acid‡	12·4	13·0	12·6	12·6	13	11
Proline	3·1	3·3	3·5	3·1	3	3
Glycine	12·1	12·3	12·2	12·3	12	11
Alanine	16·6	16·7	17·2	16·9	17	18
Valine	11·7	11·8	12·1	12·4	12	12
Methionine	—	0·7	0·8	0·5	1	1
Isoleucine	7·1	7·6	7·7	7·6	8	8
Leucine	17·5	17·8	18·0	17·7	18	18
Tyrosine	2·1	2·1	2·0	1·8	2	2
Phenylalanine	8·0	8·0	8·0	8·0	8	8
Tryptophan§					3	3

* Average or extrapolated value.

† The values for grey kangaroo β-II chain are from the sequence data corrected for the additional tryptophan residue.

‡ Our glutamic acid values are usually high (see text).

§ *N*-Bromosuccinimide titration data (see text).

The amino acid composition of the β-chain was calculated from results on hydrolysates obtained after 24, 48, 72, and 140 hr hydrolysis. The results are given in Table 1. After allowing for progressive destruction of serine and threonine the analysis of the β-chain agreed with the sequence analysis of grey kangaroo (Air and

Thompson 1969) except for extra residues of glycine, tryptophan, and glutamic acid and one residue less of alanine and serine. The values for serine and tryptophan in the grey kangaroo have been recently found to require revision since in place of a serine residue an additional tryptophan residue has been detected (see later). The glutamic acid value is probably high due to a loss of this amino acid in the standard mixture on which the analyses depend, resulting in consistently high values for this amino acid in our analyses. The difference between the values for glycine and alanine are real and have been confirmed by sequence analysis.

Peptide maps of the β-globin (cf. Air and Thompson 1969) and α-globin (cf. Beard and Thompson 1971) of red kangaroo were indistinguishable from those of grey kangaroo.

(*b*) *Inferred Sequence of β-Globin of Red Kangaroo*

Fractionation of tryptic digests of the *S*-carboxymethylated β-chain on a column of Sephadex G25 in 0·01N NH_3 and by peptide mapping gave results almost identical to those reported for β-globin from the grey kangaroo (Air and Thompson 1969). The only variation in the peptide maps was a large increase in the intensity of the βTp5B (sulphoxide) spot and a corresponding decrease in the βTp5B intensity, due to oxidation of the methionine residue. This oxidation appears to be caused by changes in the pyridine–acetate buffer in the pH 6·4 high-voltage electrophoresis tanks. The methionine is not oxidized in whole globin since cyanogen bromide cleavage takes place at the methionine β55 residue. The bond involving methionine sulphoxide is stable (Gross 1967).

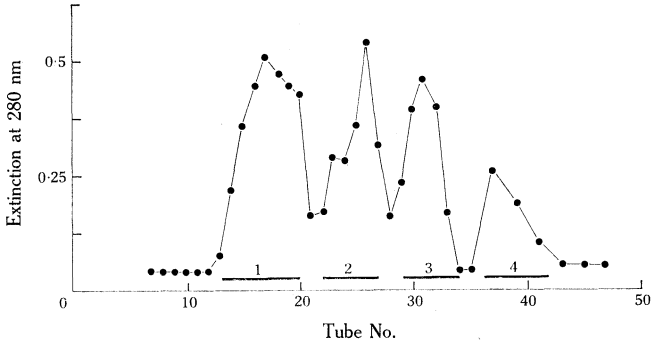


Fig. 1.—Gel filtration of cyanogen bromide fragments from red kangaroo β-globin on a column of Sephadex G50 (120 by 1 cm diam.) eluted with 10% formic acid. Fraction size 2·7 ml; flow rate 11 ml/hr; the fractions pooled for analyses are indicated by bars. Fraction 4 contained the reagent and no peptide material.

To determine the locus of the glycine–alanine change in the red kangaroo β-globin the chain was cleaved by cyanogen bromide and the products fractionated by gel filtration. Three main peptide fractions were obtained as shown in Figure 1. The larger fragment CNBr2 (91 residues) was eluted first followed by CNBr1 (55 residues) and a smaller fragment comprising residues 131–146. The smaller fragment corresponds to a cleavage following a tryptophan residue and could be due to the susceptibility of such bonds to bromine oxidation and acid cleavage (Ramachandran and Witkop 1967).

Samples of CNBr1 and CNBr2 were subjected to *N*-terminal group determination by the dansyl method. CNBr1 had *N*-terminal valine as in the grey kangaroo but CNBr2 had *N*-terminal glycine in contrast to the alanine known for the grey kangaroo. To confirm this difference tryptic digests of β-globin were fractionated and the βTp5B and βTp5B (sulphoxide) spots isolated, hydrolysed, and analysed. The peptide gave a residue of glycine replacing one of the two alanine residues of this peptide in the grey kangaroo. Four different animals of each species were examined and in each case there was glycine as residue β56 in the red kangaroos and alanine in the grey kangaroos.

The compositions (Table 2) of all other soluble tryptic peptides from red and grey kangaroos were identical (cf. Air and Thompson 1969). Time did not allow sequencing of each peptide but it is presumed that the sequences are probably the same. In the case of βTp13 a noticeable difference was the complete absence of serine from the hydrolysate of the red kangaroo whereas 0·4 residues had been found in the grey βTp13 and assigned to residue 126. Re-examination of the βTp13 peptide from grey kangaroo and parallel experiments with the peptide from red kangaroo confirmed that residue 126 should be threonine rather than serine and residue 130 tryptophan instead of threonine as previously reported. The tryptophan was not detected in acid hydrolysates, but alkaline hydrolysis of βTp13 from grey kangaroo gave one residue of tryptophan. Further data on βTp13 is given in the data on the sequences from the potoroo β-chain.

There are three insoluble β-chain tryptic peptides. Two of these, βTp4 and βTp5A, are in CNBr1 and one, βTp12, in CNBr2. When β-globin is digested with trypsin and the products fractionated on Sephadex G25 in 0·01N NH_3 the βTp4 and βTp12 are present in the first peak fraction while βTp5A is in the second peak fraction (Air and Thompson 1969). After freeze-drying and dissolving the soluble peptides in pH 6·4 buffer the insoluble residues can be purified on Sephadex G50 in formic acid solutions. In this way βTp12 was isolated from both CNBr2 and β-globin tryptic digests and its amino acid composition was identical with that from the grey kangaroo. The peptides βTp4 and βTp5A were obtained mixed with βTp12 from a globin digest, as evidenced by the presence of *S*-carboxymethylcysteine. By subtracting the contribution of βTp12 to the amino acid composition of the mixtures it was apparent that both βTp4 and βTp5A had compositions identical with those of the peptides from the grey kangaroo.

The only difference found between the β-globins of the grey and red kangaroo is a change of β56 from alanine in the grey to glycine in the red kangaroo. The peptide maps of tryptic peptides of grey and red kangaroo β-globin are identical which suggests that the number of amide groups in the peptides are the same.

(*c*) *The β-Globins of the Wallaroo,* Macropus robustus

The wallaroo was found to have two allelic haemoglobins similar to those of the grey kangaroo on starch-gel electrophoresis and ion-exchange chromatography. Separation of the globin chains on CM-cellulose in urea–thiol buffers gave β-I, β-II, and α-globins in the same elution pattern as that obtained from grey kangaroo globin (Thompson, Hosken, and Air 1969). In an extended survey Richardson (1970) has examined several species and subspecies of wallaroos. Of those studied here, *Macropus robustus robustus* (Gould, 1841), inhabits the coastal range areas of New South Wales

TABLE 2

AMINO ACID COMPOSITIONS OF RED KANGAROO β-GLOBIN TRYPTIC PEPTIDES

Peptides were purified by gel filtration and peptide mapping except βTp12, which was obtained from the insoluble portion of a tryptic digest after gel filtration. Values given are moles per mole of protein with preferred values in parentheses

Amino Acid	βTp1	βTp2	βTp3	βTp4†	βTp5A‡	βTp5B	βTp6,7,8	βTp9A	βTp9B	βTp10A	βTp10B +11	βTp12	βTp13	βTp14	βTp15
Lysine	1·0 (1)	0·9 (1)			1·0 (1)	1·0 (1)	1·1 (1)	1·0 (1)	1·1 (1)	1·4 (1)	2·0 (2)	1·0 (1)	0·7 (1)	1·0 (1)	
Histidine	1·1 (1)				0·9 (1)		1·0 (1)				1·8 (2)	0·9 (1)		1·0 (1)	1·0 (1)
Arginine			1·0 (1)	1·0 (1)											
SCM-cysteine											+ (1)	0·7 (1)			
Aspartic acid		0·9 (1)		0·2	3·1 (3)	1·0 (1)		0·4 (1)	2·8 (3)		2·6 (3)	1·1 (1)	1·2 (1)	0·9 (1)	
Threonine	0·8 (1)	0·9 (1)	1·0 (1)	1·3 (1)	0·2					0·8 (1)		0·3	1·7 (2)		
Serine		1·6 (1)		1·6 (1)	1·0 (1)	0·4				0·6	1·1 (1)				
Glutamic acid	2·2 (2)		3·1 (3)	0·3	0·2	0·3					2·1 (2)	1·3 (1)	3·1 (3)		
Proline				1·4 (1)	0·1	0·6 (1)					0·6 (1)				
Glycine		1·4 (1)	3·0 (3)		1·1 (1)	1·1 (1)	1·1 (1)	1·1 (1)		1·0 (1)		1·5 (2)		1·2 (1)	
Alanine	1·0 (1)	1·2 (1)	2·1 (2)	0·3	1·2 (1)	0·9 (1)	2·1 (2)	2·1 (2)		1·2 (1)		1·2 (1)	1·2 (1)	3·9 (4)	
Valine	0·9 (1)		0·9 (1)	1·1 (1)	0·2	1·0 (1)	1·0 (1)	2·0 (2)			1·0 (1)	1·1 (1)	1·1 (1)	1·9 (2)	
Methionine						+ (1)									
Isoleucine		0·9 (1)	1·0 (1)	0·5 (1)				1·0 (1)				2·4 (3)	1·0 (1)		
Leucine	1·0 (1)	0·9 (1)	1·0 (1)	1·8 (2)	1·2 (1)		1·0 (1)	1·0 (1)	2·0 (2)		2·5 (3)	2·6 (3)		1·5 (2)	
Tyrosine				1·3 (1)											1·0 (1)
Phenylalanine					3·0 (3)			0·8 (1)		1·0 (1)	1·0 (1)	1·0 (1)	1·1 (1)		
Tryptophan*		+ (1)		0·3 (1)									+ (1)		
Total in peptide	8	9	13	10	12	7	7	10	6	5	17	16	12	12	2

* Detected by Ehrlich reagent.

† Purified by gel filtration on Sephadex G50 in 10% formic acid. Allowance made for contamination with βTp12, most of which eluted in earlier fractions.

‡ From the insoluble portion of a tryptic digest of globin. Not pure (see text).

and Queensland and was found by Richardson (1970) to have only Hb-II in 27 animals studied. Both Hb-I and Hb-II occurred in the subspecies *Macropus robustus erubescens* (Sclater, 1870), which inhabits the dry plains and ranges of western New South Wales and inland Australia, with a gene frequency for Hb-II varying from 0·1 to 0·4 in different areas (97 animals studied).

From four animals, one homozygous for Hb-I (*M.r.erubescens*) and three homozygous for Hb-II (*M.r.robustus*) the β-globin tryptic peptide maps were found to be identical to those from grey kangaroo β-I and β-II globins. The composition and sequence of the βTp1 peptides were identical to those of grey kangaroo with glutamine at position β2 in the β-I chain substituted by histidyl residues in the β-II chains.

The composition of βTp5B and its sulphoxide peptides included a glycyl residue in all four animals as in red kangaroo but different from that of grey kangaroos. Cyanogen bromide cleavage of the methionyl (β55) bond gave dansyl-glycine in addition to dansyl-valine when the cleaved material was subjected to end group analysis, hence the glycine is at position β56 as in the red kangaroo.

(*d*) *The β-Globins of Other Macropods*

The species included in this survey are listed in Table 3, with the amino acid sequences of the tryptic peptide βTp5B for each species. This peptide shows con-

TABLE 3

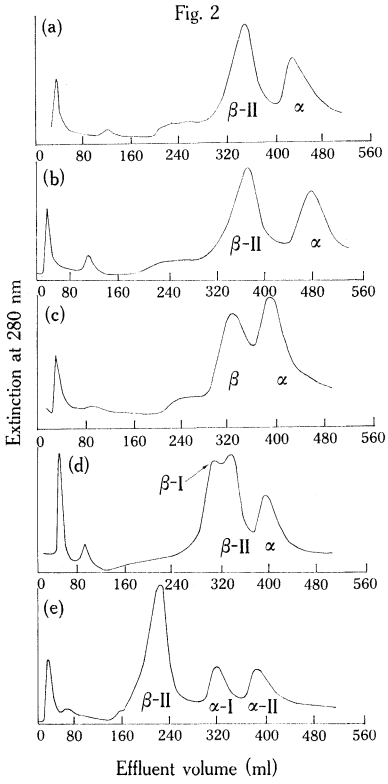
SEQUENCE OF TRYPTIC PEPTIDE βTp5B FROM DIFFERENT MACROPODS

Species	Common Name	No. Examined	Sequence
Subfamily Macropodinae			
Macropus giganteus	Eastern grey kangaroo	4	Ala-Val-Met-Ala-Asn-Pro-Lys
M. fuliginosus	Western grey kangaroo	1	
Megaleia rufa	Red kangaroo	4	Ala-Val-Met-Gly-Asn-Pro-Lys
M. robustus	Wallaroo or euro	4	
M. parryi	Pretty-face or whiptail wallaby	2	
M. rufogrisea	Bennets or red-necked wallaby	2	
Wallabia bicolor	Swamp wallaby	2	
M. antilopinus	Antelopine kangaroo	1	
Thylogale stigmatica wilcoxi	Red-legged pademelon	1	Ala-Val-Met-Ala-Asn-Pro-Lys
Thylogale billardieri	Red-bellied pademelon	1	Ala-Val-Met-Ser-Asn-Pro-Lys
M. eugenii	Tammar	4	
Petrogale xanthopus	Yellow-footed rock wallaby	1	Ala-Val-Met-Gly-Asn-Ala-Lys
M. agilis	Agile wallaby	2	Ala-Val-Met-Thr-Asn-Ala-Lys
		1	Mixture*
Subfamily Potoroinae			
Potorous tridactylus	Potoroo		Ala-Val-Leu-Gly-Asn-Ala-Lys

* Ala-Val-Met-Thr-Asn-Ala-Lys plus Ala-Val-Met-Ala-Asn-Pro-Lys.

siderable variation, particularly at residue β56 where alanine, glycine, serine, or threonine is present. Fractionation of the globins gave elution curves as shown in

Figure 2. All animals other than heterozygote grey kangaroos and wallaroos gave one β-chain peak followed by one α-globin except two of the three samples of *Macropus agilis* where two β-globins were present. The tryptic peptide βTp5B from another



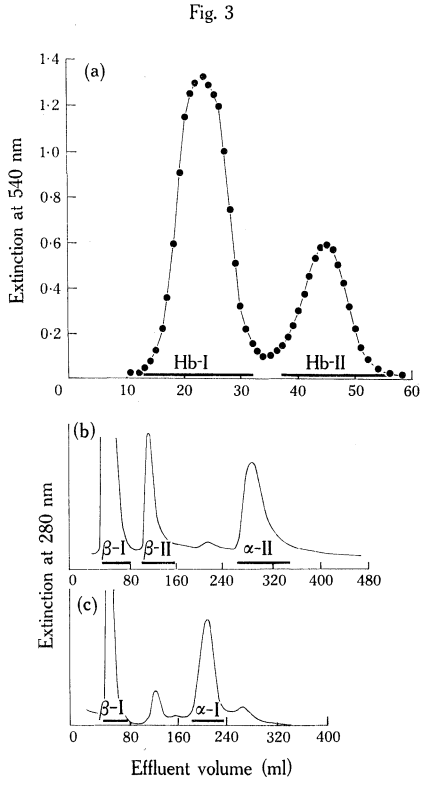


Fig. 2.—Continuous recording of the extinction at 280 nm of the effluent from a linear gradient elution of macropod and potoroo globins from CM-cellulose equilibrated with 8M urea–thiol buffer. Initial buffer was 8M urea–0·05M mercaptoethanol–0·01M Na^+ (phosphate)–0·001M EDTA at pH 6·7. Flow rate 96 ml/hr; column size 1·8 cm diam. by 12 cm. Limit buffer was 250 ml and contained 8M urea–0·05M mercaptoethanol–0·05M Na^+ (phosphate)–0·001M EDTA. Approximately 100 mg of globin was loaded in each case. (*a*) Grey kangaroo Hb-II homozygote; (*b*) wallaroo Hb-II homozygote; (*c*) tammar; (*d*) agile wallaby (two β-chains); (*e*) potoroo (two α-chains).

Fig. 3.—(*a*) Separation of Hb-I and Hb-II of potoroo red cell haemolysate on a column of DEAE-Sephadex A50 (2·8 cm diam. by 16 cm) by linear gradient elution at room temperature with 0·05M Tris–HCl from pH 8·0 to pH 7·0 (300 ml each chamber). After completion of elution of Hb-I (tube 33) the elution was continued with the buffer of pH 7·0. Fraction size 8 ml; flow rate approximately 100 ml/hr. (*b*) Continuous recording as in Figure 2 of the elution of the globin from Hb-I of the potoroo from CM-cellulose using 8M urea–thiol buffers at pH 6·7. (*c*) Continuous recording as in (*b*) of the globin of Hb-II of the potoroo.

sample of *M. agilis* gave two sequences as shown in Table 3, the sequences being inferred from the relative intensities of the two dansyl amino acids obtained after steps 3 and 5 of the Edman degradation. Since this sample gave only one β-globin

peak this is not the reason for the double peak shown in Figure 2, where the separation on CM-cellulose would suggest there is a charge difference between the two β-globins. Note that polymorphism could be present in many of these species, particularly where only one or two samples were available. The methods used in this study would not detect variations among non-charged residues unless there was a marked difference in peptide chromatographic rates. Despite the apparent similarity a minimum requirement to give confidence that two peptides from different globins were identical would be amino acid analysis. The fingerprint of the β-globin tryptic peptides of the potoroo showed only two changes in position on the peptide map, although there are changes in the amino acid sequence of six of the 12 soluble peptides compared with those of β-globin from the grey kangaroo.

(*e*) *The Globin Chains of the Potoroo*

When globin from the potoroo was fractionated on CMC-cellulose the elution curve shown in Figure 2(*e*) was obtained. There are two α-chains (α-I and α-II) and one β-chain (β-II). This pattern was obtained from the globin of four animals whose haemoglobin on starch-gel electrophoresis gave two bands (P. G. Johnstone, personal communication). Blood from an animal giving only a single band on starch-gel electrophoresis gave only α-II and β-II globin peaks on CM-cellulose chromatography. Blood from a fifth animal gave two bands on starch-gel electrophoresis but four peaks when the globin was chromatographed, β-I, β-II, α-I, and α-II, named in order of emergence from the column.

The haemoglobin from this animal when fractionated on DEAE-Sephadex gave a major peak, Hb-I and a minor component, Hb-II (Fig. 3). The globins from each of these peaks were separately examined by CM-cellulose chromatography. Hb-I globin gave β-I, β-II, and α-II chains while Hb-II globin gave β-I and α-I chains (Fig. 3). The Hb-I peak from this animal obviously contains two haemoglobins, i.e. $\alpha_2^{II}\beta_2^{I}$ and $\alpha_2^{II}\beta_2^{II}$.

The amino acid compositions of each of these globins were determined on a 24-hr hydrolysate. These analyses are shown in Table 4. No significant difference was apparent between the β-I and β-II globin results, suggesting the polymorphism may be due to an amide difference.

The results for the α-chains show a probable difference in the contents of lysine and glutamic acid residues which could be responsible for the chromatographic differences seen.

(*f*) *The Amino Acid Sequence of the β-II Globin of the Potoroo*

Since β-II was the most prevalent β-chain in the animals studied tryptic peptides of this chain were isolated for sequence studies. The fractionation was by gel-filtration on Sephadex G50 (Fig. 4) followed by separation of the bulked peak tubes into fractions soluble and insoluble at pH 6·4. On the two-dimensional peptide map the soluble peptides occupied similar positions to the peptides from grey kangaroo (Air and Thompson 1969) except for βTp3 (more acidic) and βTp2 (faster during chromatography).

The use of ammonium bicarbonate rather than 0·01N ammonia for the gel filtration gave a better resolution according to molecular weight than previously obtained. This was evident in that fraction 1 contained predominantly the double peptide βTp10B+11 (17 residues) with some βTp3 (13 residues), fraction 2 contained the bulk of peptides containing 7–12 residues, fraction 3 contained four main peptides of 5–9 residues, the larger one βTp2 being retarded by its aromatic tryptophanyl residue, while fraction 4 contained only the dipeptide Tyr-His (βTp15). The insoluble

TABLE 4

AMINO ACID COMPOSITION OF α-I, α-II, β-I, and β-II CHAINS FROM POTOROO GLOBIN

Values given are moles per mole globin calculated from the μmoles recovered from the column and converted to moles per mole of 141 residues for α-chain and 146 residues for β-chain by reference to the amino acids stable to hydrolysis. Hydrolysis was for 24 hr at 120–125°C in 6N HCl containing 1 mg phenol per millilitre

Amino Acid	Globin*					β-II Sequence†
	α-I	α-II	β-I(1)	β-I(2)	β-II	
Lysine	10·4	11·5	13·1	13·2	13·0	13
Histidine	10·6	10·6	7·9	7·9	7·6	8
Arginine	2·9	3·0	2·1	2·2	2·0	2
SCM-cysteine	1·0	1·0	1·8	1·9	1·8	2
Aspartic acid	11·6	11·7	15·4	15·4	15·3	15
Threonine‡	5·7	5·9	3·6	3·4	3·3	5
Serine‡	4·7	4·7	3·6	3·1	3·6	8
Glutamic acid	8·6	7·6	10·5	10·6	10·7	10
Proline	4·9	4·5	2·4	2·1	1·8	2
Glycine	9·3	10·3	13·2	13·1	13·2	13
Alanine	18·2	18·1	15·7	15·0	15·1	15
Valine	11·2	11·2	13·1	12·0	11·9	12
Methionine	0	0	0	0	0	0
Isoleucine	3·1	3·2	6·8	6·7	6·5	7
Leucine	17·2	17·1	20·9	20·8	20·7	21
Tyrosine	3·0	3·0	2·3	2·2	2·3	2
Phenylalanine	8·0	8·0	8·0	8·0	8·0	8
(Tryptophan)						(3)
Total						146

* α-I and β-I(1) globin were isolated from Hb-II; α-II, β-I(2), and β-II globin from Hb-I (Fig. 3). Results for the two β-I samples are given.

† From Table 5.

‡ Uncorrected for decomposition, which is considerable at this temperature where liberation of valine and isoleucine is more complete.

fractions also showed better resolution by molecular weight. That from fraction 1 was βTp12 (16 residues), from fraction 2 βTp5A (12 residues), while that from fraction 3 was a mixture of βTp5A and βTp4 (12 residues)—the latter containing tryptophan and tyrosine which contribute to its retardation.

The tryptic peptides are considered in turn with the evidence on which their sequence is based. Residues identified as dansyl derivatives are given in *italic* font.

(i) *βTp1*

This peptide lacked threonine and alanine compared with the peptide from grey kangaroo but two serine residues were present. The peptide was neutral showing that no amide residues were present. The amino acid sequence is

Val-His-Leu-Ser-Ser-Glu-Glu-Lys

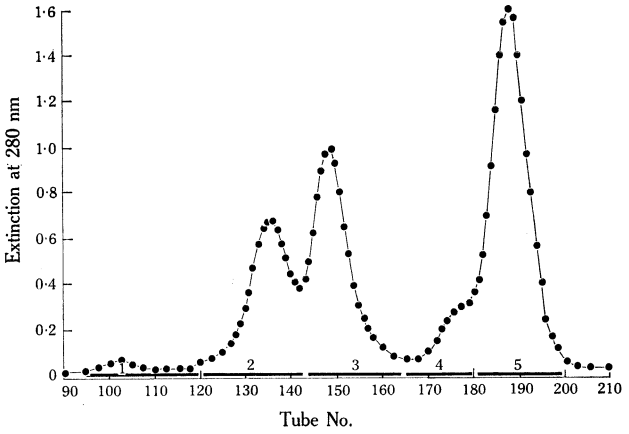


Fig. 4.—Gel filtration of a tryptic digest of the β-chain of potoroo globin (145 mg) on a column (110 cm by 2·8 cm diam.) of Sephadex G50 in 1% ammonium bicarbonate, pH 9·1. Flow rate 36 ml/hr; fraction size 3 ml. Tubes containing fractions combined for peptide recovery are shown as bars. The peptides present in each fraction are as follows:

Fraction 1	Fraction 2	Fraction 3	Fraction 4	Fraction 5
Peptides soluble at pH 6·4				
βTp3	βTp1	βTp2	βTp15	Phenol
βTp10B*	βTp3*	βTp6,7,8*		
βTp10B+11	βTp5B	βTp9B*		
	βTp6,7,8	βTp10A		
	βTp9A			
	βTp9B			
	βTp10B			
	βTp13			
	βTp14			
Peptides insoluble at pH 6·4				
βTp12	βTp5A	βTp4+βTp5A*		

* Indicates major proportion of peptide is in another fraction.

(ii) *βTp2*

The peptide showed two differences in amino acid composition from that of grey kangaroo with leucine and glycine replacing alanine and aspartic acid. This accounts for its faster chromatographic rate on peptide maps. The dansyl-Edman degradation showed the sequence

Gly-Leu-Ile-Thr-*Ser-Leu*-Trp-*Gly-Lys*

Digestion with thermolysin gave peptides consistent with this sequence and splitting at the bonds shown

Gly-Leu-*Ile-Thr-Ser*-Leu-Trp-Gly-Lys

(iii) *βTp3*

Analysis showed two differences from the peptide from grey kangaroo with isoleucine and aspartic acid replacing alanine and valine. The sequence is

Ile-Asp-Ile-Glu-Gln-Thr-Gly-Gly-Glu-Ala-Leu-Gly-Arg

This peptide was more acidic than that from grey kangaroo. By identification of the phenylthiohydantoins (PTH) residue 2 was shown to be aspartic acid, residue 4 was glutamic acid, and 5 was glutamine; thus the allocation of the amide residues is the same as in grey kangaroo.

(iv) *βTp4*

This has proved the most difficult of the insoluble peptides to obtain pure, and it could not be completely separated from βTp5A. Analysis of the mixture and subtraction of the contribution expected from the βTp5A showed that the composition of βTp4 was very similar to that from grey kangaroo. A solution made by heating the mixture of insoluble peptides for a few minutes in 50% pyridine and standing it in the cold contained mainly βTp4, the βTp5A having precipitated. The dansyl–Edman degradation gave the sequence

Leu-Leu-Ile-Val-Tyr-Pro-Trp-Thr-Ser-Arg

The evidence for the Thr-Ser- sequence was weak by this method, so a sample was digested with chymotrypsin and the mixture subjected to peptide mapping. The strongest basic peptide contained threonine, serine, and arginine. The sequence data was consistent with the sequence shown and is the same as in grey kangaroo.

(v) *βTp5A*

This was obtained almost pure in the insoluble portion of fraction 2 from Sephadex. The amino acid analysis was similar to that for the peptide from grey kangaroo. Minor spots during the dansyl–Edman degradation showed that some βTp13 was present. Subtracting the contribution that this would make to individual amino acid contents gave the composition

$Lys_{1.0}$ $His_{1.0}$ $Asp_{2.0}$ $Ser_{1.4}$ $Gly_{1.0}$ $Ala_{0.8}$ $Leu_{1.0}$ $Phe_{2.7}$

The peptide was neutral on ionophoresis at pH 6·4 showing two aspartyl residues must be present. The sequence is

Phe-Phe-Asp-His-Phe-Gly-Asp-Leu-Ser-Ser-Ala-Lys

(vi) *βTp5B*

This peptide occurred in the same position as βTp5B in the grey kangaroo tryptic peptide map but gave no spot in the sulphoxide position since it contained no methionine.

The amino acid analysis and sequence data showed it to be

Ala-Val-Leu-Gly-Asn-Ala-Lys

(vii) *βTp6,7,8*

The analysis and sequence of this peptide was identical to that found for the grey kangaroo

Val-Leu-Ala-His-Gly-Ala-Lys

(viii) *βTp9A*

The composition of this peptide showed one residue of alanine less and one residue of serine present in comparison with the peptide from grey kangaroo. The sequence data showed this replacement was at residue 4:

Val-Leu-Val-Ser-Phe-Gly-Asp-Ala-Ile-Lys

(ix) *βTp9B*

Identical in composition and sequence to the peptide from grey kangaroo. The amide allocations were confirmed by identification of the PTH derivatives for residues 1, 3, and 4:

Asn-Leu-Asp-Asn-Leu-Lys

(x) *βTp10A*

Identical in composition and sequence to the peptide from grey kangaroo

Gly-Thr-Phe-Ala-Lys

(xi) *βTp10B+11*

This double peptide is the major product of tryptic digestion due to a poorly hydrolysed Asp-Lys-Leu bond. The composition was identical to that of the peptide from grey kangaroo. The small amount of βTp10B as well as βTp10B+11 gave a clear sequence for the βTp10B section by the dansyl–Edman technique.

Digestion of βTp10B+11 with thermolysin and fractionation by ionophoresis at pH 6·4 and chromatography gave peptides which by ionophoretic rate, composition, and sequence confirmed that the sequence was identical with that of the grey kangaroo. An acidic peptide *Leu-Ser-Glu* and a neutral peptide *Leu-His-SCMCys-Asp*-Lys were derived from the 10B section. Two basic peptides *Leu*-His and *Phe*-Lys were obtained. An acidic peptide *Val-Asp-Pro-Glu-Asn* gave a clear sequence with identification of the PTH derivatives for residues 2, 4, and 5 giving the amide distribution. The sequence, showing the bonds cleaved by thermolysin, is

Leu-Ser-Glu-Leu-His-SCMCys-Asp-Lys-Leu-His-Val-Asp-Pro-Glu-Asn-Phe-Lys
↑ ↑ ↑ ↑

(xii) *βTp12*

The composition of this peptide was obtained from the insoluble residue of fraction 1 from the Sephadex gel filtration as

Lys_1 $His_{0.8}$ $SCMCys_{0.7}$ $Asp_{1.2}$ $Glu_{1.3}$ $Gly_{1.8}$ $Ala_{1.2}$ $Val_{1.9}$ $Leu_{3.4}$ $Phe_{1.0}$

The dansyl–Edman method gave 15 residues of the sequence, the degradation not stopping at residue 4 as with the peptide from the grey kangaroo.

Digestion with pepsin and peptide mapping gave several peptides that confirmed the sequence. Thus basic peptides

Leu-Ala-Glu-His-Phe-Gly-Lys and *Ala-Glu-His-Phe-Gly*-Lys

from the *C*-terminal sequence were obtained. A neutral peptide Leu-Leu-Gly-Asn-Val represented the *N*-terminal sequence while an acidic peptide *Val-Ile-SCMCys* confirmed the sequence near the cysteine residue. The sequence, showing the bonds cleaved by pepsin, is

Leu-Leu-Gly-Asn-Val-Leu-Val-Ile-SCMCys-Leu-Ala-Glu-His-Phe-Gly-Lys

(xiii) *βTp13*

The amino acid composition of this peptide showed some additional amounts of serine, glycine, and alanine compared with the peptide from grey kangaroo. After purification by a longer ionophoresis (50 V/cm, 2 hr) at pH 6·4 the peptide gave a clear dansyl–Edman sequence for 9 residues

Asx-Phe-Thr-Ile-Asx-Ala-Glx-Val-Ala

No dansyl derivative was obtained for residue 10 and only a weak dansyl-Glx for residue 11.

The peptide was digested with thermolysin and the mixture fractionated by ionophoresis at pH 6·4 and chromatography. Acidic peptides *Asp-Phe-Thr* and *Ile-Asp-Ala-Gln* were obtained, the PTH derivatives confirming that Asp and Gln residues were present in the terminal sequence shown above. The basic area contained a peptide Gln-Lys and other faint spots giving valine, alanine, glutamine, and lysine on acid hydrolysis. The neutral peptides included *Val-Ala* and a faint ninhydrin-positive zone, running faster during chromatography, which gave only valine and alanine on acid hydrolysis. This peptide obviously contained an additional residue which was identified as tryptophan by the Ehrlich colour reaction. Hence residue 10 is tryptophan and the sequence of βTp13, with the bonds cleaved by thermolysin, is

Asp-Phe-Thr-Ile-Asp-Ala-Gln-Val-Ala-Trp-*Gln*-Lys

The identification of tryptophan as residue 10 raised the possibility that it was also present in Tp13 from the red kangaroo and the grey kangaroo. Thermolysin digestion of the peptide from red kangaroo gave a similar pattern of peptides to the potoroo although the sequence of the acidic peptides were slightly different: Glu-Phe-Thr and Ile-Asp-Thr-Gln.

For the grey kangaroo the peptide βTp13 was purified and a sample hydrolysed with NaOH for analysis. Tryptophan (1 residue) was obtained in addition to the other amino acids expected from the composition determined from acid hydrolysates (Air and Thompson 1969) except that threonine had been converted to glycine and the isoleucine and aspartic acid were not completely liberated. The previous error in assigning a threonine residue to position β130 in grey kangaroo was due to a wrong

identification of dansyl-serine as residue 126 rather than dansyl-threonine. The pepsin peptide identified from its amino acid composition as

Gln-Val-Ala-Thr-Gln-Lys

should have been

Thr-Gln-Val-Ala-Trp-Gln-Lys

with the tryptophan residue being destroyed during acid hydrolysis. The small peptide thought to be Gln-Lys was probably Trp-Gln-Lys based on its ionophoretic mobility and calculations according to Offord (1966), the tryptophan not surviving acid hydrolysis.

That three residues of tryptophan occurred in β-globin of grey kangaroo was not supported by *N*-bromosuccinimide titration results which gave 2·1 residues (Beard and Thompson 1971). Since this tryptophan at β130 is in a very hydrophobic environment (Perutz, personal communication) these titrations were repeated on samples that had stood for at least 24 hr in the 8M urea solution at pH 3·5 to allow unfolding to occur before titration with *N*-bromosuccinimide. Under these conditions (Beard and Thompson, unpublished data) mean values of 2·6, 2·7, and 2·9 residues per mole were obtained for grey kangaroo, red kangaroo, and potoroo β-globin respectively. Alkaline hydrolysis of two samples of grey kangaroo β-globin gave values of 2·6 and 3·0 residues confirming that three residues of tryptophan are present.

(xiv) *βTp14*

This peptide has the same amino acid composition and sequence as that from the grey kangaroo namely

Leu-Val-Ala-Gly-Val-Ala-Asn-Ala-Leu-Ala-His-Lys

(xv) *βTp15*

This peptide is the same as in grey kangaroo and all known β-globins

Tyr-His

(xvi) *Complete Sequence*

The complete amino acid sequence of the potoroo β-chain is shown in Table 5. The peptides have been aligned from their strong homology with kangaroo β-globin.

IV. Discussion

In comparative studies of amino acid sequences of proteins from different species limitations of time and facilities make it difficult to completely sequence all peptides. The amino acid compositions of the tryptic peptides of the β-globin of red kangaroo, however, suggest that a change from an alanyl to a glycyl residue at β56 is the sole difference between the β-globins of grey and red kangaroos respectively. The polymorphism at position β2 found in grey kangaroos was not present in the red kangaroos examined, but in the case of the wallaroo, with glycine at β56 as in red kangaroo, there is an allelic glutaminyl or histidyl residue at β2 as in the grey

TABLE 5

AMINO ACID SEQUENCE OF POTOROO β-GLOBIN

Residues are numbered from the *N*-terminal residue and the positions of cleavage by trypsin are indicated by arrows. The 15 residues which differ from those present in grey kangaroo β-chain are shown in capital letters

```
                5                  10                  15                  20                  25                  30
Val-His-Leu-SER-SER-Glu-Glu-Lys-GLY-LEU-Ile-Thr-Ser-Leu-Trp-Gly-Lys-ILE-ASP-Ile-Glu-Gln-Thr-Gly-Gly-Glu-Ala-Leu-Gly-Arg-
          βTp1               ↑          βTp2               ↑                  βTp3                              ↑

           35                  40                  45                  50                  55
Leu-Leu-Ile-Val-Tyr-Pro-Trp-Thr-Ser-Arg-Phe-Phe-Asp-His-Phe-Gly-Asp-Leu-Ser-SER-Ala-Lys-Ala-Val-LEU-GLY-Asn-ALA-
                    βTp4               ↑                  βTp5A              ↑          βTp5B

 60                  65                  70                  75                  80                  85
Lys-Val-Leu-Ala-His-Gly-Ala-Lys-Val-Leu-Val-SER-Phe-Gly-Asp-Ala-Ile-Lys-Asn-Leu-Asp-Asn-Leu-Lys-Gly-Thr-Phe-Ala-Lys-Leu-
↑          βTp6,7,8          ↑                  βTp9A              ↑          βTp9B          ↑          βTp10A          ↑

 90                  95                 100                 105                 110                 115
Ser-Glu-Leu-His-Cys-Asp-Lys-Leu-His-Val-Asp-Pro-Glu-Asn-Phe-Lys-Leu-Leu-Gly-Asn-VAL-LEU-Val-Ile-Cys-Leu-Ala-Glu-His-Phe-
          βTp10B          ↑          βTp11              ↑                  βTp12

120                 125                 130                 135                 140                 145
Gly-Lys-ASP-Phe-Thr-Ile-Asp-ALA-Gln-Val-Ala-Trp-Gln-Lys-Leu-Val-Ala-Gly-Val-Ala-Asn-Ala-Leu-Ala-His-Lys-Tyr-His
     ↑                  βTp13              ↑                  βTp14                              ↑ βTp15
```

kangaroo. The presence of the same polymorphism at β2 in the grey kangaroo and wallaroo but with a consistent species difference of a glycine–alanine change at β56 is difficult to explain in terms of evolution. Either the grey kangaroo–wallaroo common ancester had the polymorphism at β2 and the glycine–alanine change became incorporated into both alleles by independent mutations or by crossing-over with selection of only one of the products, or the β2 polymorphism has evolved independently in both species.

For a total of 14 species of macropods the sequences of the peptide βTp5B, the locus of the difference between red and grey kangaroos, have been determined. The residues in position β56 vary between glycine, alanine, serine, and threonine residues while position β58 is either alanine or proline. These could all occur by one base mutation in the nucleotide sequence of DNA.

TABLE 6

TIME FOR ONE CHANGE IN β-GLOBIN AMINO ACID SEQUENCE PER HUNDRED RESIDUES

β-sequences Compared	No. of Differences	Differences per 100 Residues* ÷2	Time since Divergence (millions of years)	Rate of Change (millions of years)
Human–rhesus monkey	8	2·7	30	11·1
Human–horse	26	8·9	80	9·0
Human–sheep B	26	8·9	80	9·0
Human–rabbit	14	4·8	80	16·7
Human–grey kangaroo	38	13·0	120	9·2
Rhesus monkey–horse	28	9·6	80	8·3
Rhesus monkey–sheep B	27	9·2	80	8·7
Rhesus monkey–rabbit	16	5·5	80	14·5
Rhesus monkey–grey kangaroo	36	12·3	120	9·8
Horse–sheep B	33	11·3	60	5·3
Horse–rabbit	25	8·6	80	9·3
Horse–grey kangaroo	46	15·8	120	7·6
Sheep B–rabbit	29	9·9	80	8·1
Sheep B–grey kangaroo	42	14·7	120	8·2
Rabbit–grey kangaroo	37	12·7	120	9·4
Mean rate of change				9·6±1·5†

* Residues are 146 in number in each sequence compared.

† 95% confidence limits.

This variation in amino acid sequence of peptide βTp5B suggests that this peptide could be of use in taxonomic work in the Macropodidae although the information is very limited compared with a complete sequence determination.

The potoroo is a member of the subfamily Potoroinae, which is considered to have diverged well before the radiation of the main macropodine species. Some differences were seen in the fingerprint of tryptic peptides from potoroo β-chain and the complete amino acid sequence has been determined.

The relationship to the macropods is clearly revealed, e.g. by the identical positioning of lysyl and arginyl residues in the chain which show quite marked differences from other mammals (Dayhoff 1969), but there is a total of 15 differences between the amino acid sequences of the β-globin from the potoroo and the grey kangaroo.

In a previous paper (Air *et al.* 1971) a method of calculating the number of years since divergence of two living species, t, was given based on the number of differences in amino acid sequence between protein chains common to both species:

$$t = 100dY/2n,$$

where d is the number of amino acid differences in the polypeptide chains containing n amino acid residues and Y is the rate of change, expressed as years per amino acid change per 100 residues. The value of Y for β-globin chains was calculated from a comparison of all β-globins for which complete amino acid sequences were available and was $9{\cdot}6 \pm 1{\cdot}5$ (95% confidence limits) million years for the 15 comparisons involved, including the sequence of grey kangaroo β-globin. Using this value the time since divergence of the potoroine and macropodine lines would be about 50 million years ($41-57 \times 10^6$, 95% confidence interval).

This "biochemical clock" estimate is based on the original predictions of Zuckerkandl and Pauling (1962) and involves a knowledge of palaeontological estimates of dates of divergence. The dates used were based on information given by Romer (1966), and the Y index for β-globin chains is given in Table 6. That this is a valid estimate of times since divergence is shown by the close agreement of the times calculated for marsupial–eutherian mammal divergence using the α-, β-, and myoglobin chain data (Air *et al.* 1971). The dates of divergence estimated from palaeontological records are subject to considerable variation between different authorities and those used in Table 6 may be questioned. With increasing data of the type presented by Air *et al.* (1971) a revision of some of these palaeontological estimates, e.g. the date of divergence of rabbits from other species, may be possible.

V. Acknowledgments

The authors are indebted to Professor G. B. Sharman, Dr. W. E. Poole, Dr. B. J. Richardson, and Mr. P. G. Johnston for samples of blood and helpful discussions. We also thank Miss W. K. Fisher, Miss J. M. Salter, and Mr. R. G. Whittaker for skilled technical assistance.

This work was supported in part by the Australian Research Grants Committee and the Australian Wool Board.

VI. 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., and Thompson, E. O. P. (1970).—*Aust. J. biol. Sci.* **23**, 185.

Beard, J. M., and Thompson, E. O. P. (1971).—*Aust. J. biol. Sci.* **24**, 765.

Clegg, J., Naughton, M. A., and Weatherall, D. J. (1965).—*Nature, Lond.* **207**, 945.

DAYHOFF, M. O. (1969).—"Atlas of Protein Sequence and Structure." (National Biomedical Research Foundation: Silver Spring.)

GRAY, W. R. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 469. (Academic Press: New York.)

GROSS, E. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 238. (Academic Press: New York.)

OFFORD, R. E. (1966).—*Nature, Lond.* **211**, 591.

RAMACHANDRAN, L. K., and WITKOP, B. (1967).—In "Methods in Enzymology". (Ed. C. H. W. Hirs.) Vol. XI. p. 283. (Academic Press: New York.)

RICHARDSON, B. J. (1970).—Ph.D. Thesis, University of New South Wales.

ROMER, A. S. (1966).—"Vertebrate Paleontology." 3rd Edn. (Chicago Univ. Press.)

THOMPSON, E. O. P., HOSKEN, R., and AIR, G. M. (1969).—*Aust. J. biol. Sci.* **22**, 449.

THOMPSON, E. O. P., SLEIGH, R. W., and SMITH, M. B. (1971).—*Aust. J. biol. Sci.* **24**, 525.

ZUCKERKANDL, E., and PAULING, L. (1962).—In "Horizons in Biochemistry". (Eds. M. Kasha and B. Pullman.) p. 189. (Academic Press: New York.)