Aust. J. Biol. Sci., 1980, **33**, 153–67

Myoglobins of Cartilaginous Fishes. II.* Isolation and Amino Acid Sequence of Myoglobin of the Shark *Mustelus antarcticus*

W. K. Fisher, D. D. Koureas 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 red muscle of the gummy shark *M. antarcticus* was purified by gel filtration and ion-exchange chromatography on carboxymethyl cellulose in 8 M urea–thiol buffer. Amino acid analysis and sequence determination showed 148 amino acid residues. The amino terminal residue is acetylated as shown by nuclear magnetic resonance and mass spectrographic analysis of an *N*-terminal peptide. There is a deletion of four residues at the amino terminal end as well as one residue in the CD interhelical area relative to other myoglobins. These overall differences were also found previously in myoglobin of *Heterodontus portusjacksoni*.

The complete amino acid sequence has been determined following digestion with trypsin, chymotrypsin, thermolysin, staphylococcal protease and cyanogen bromide. Sequences of purified peptides were determined by the dansyl–Edman procedure. The amino acid sequence showed approximately 88 differences from mammalian, monotreme, bird and tuna myoglobins, slightly more than previously reported for *H. portusjacksoni* usually considered a more primitive animal. There were 24 residues common to both shark myoglobins that were different from those present in other myoglobins. The sequence has been compared to the myoglobin of yellowfin tuna and other myoglobins.

Introduction

In the first paper of this series (Fisher and Thompson 1979), the amino acid sequence of myoglobin from the shark *Heterodontus portusjacksoni* was presented. The myoglobin sequence as well as the amino acid sequence of the α- and β-globin chains of the haemoglobin (Nash *et al.* 1976; Fisher *et al.* 1977) supported the idea that this elasmobranch, which is generally considered to represent a primitive class of animals, diverged approximately 450 million years ago.

The origins and relationships between groups of living elasmobranchs are by no means clear, with an inadequate fossil record to indicate that particular groups have evolved from others. Most living groups, including some that on structural features would seem to be primitive, go back to the Upper Jurassic, and seem to have been separate lineages since that time (J. A. F. Garrick, personal communication). It was of interest therefore to examine other cartilaginous fishes to compare the amino acid sequence of a common protein and to see if any marked differences indicative of a different date of divergence could be detected.

The gummy shark *Mustelus antarcticus* Günther, 1870, is a small shark with rhombic, pavement-like teeth adapted for feeding on crustaceans, molluscs and small

* Part I, *Aust. J. Biol. Sci.*, 1979, **32**, 277–94.

fish. It is representative of 'modern' sharks (Romer 1966) and is sometimes included in a separate family of smooth dog fishes, the Triakidae (Grant 1978), first known in fossil records from the early Tertiary (Romer 1966). In this paper, the isolation of myoglobin from the red muscle of *M. antarcticus* is described, as well as the determination of its complete amino acid sequence. Peptides isolated from a variety of enzyme digests have been purified and sequenced by the dansyl–Edman procedure. Several large insoluble tryptic peptides, as well as 68 differences in amino acid sequence from the myoglobin of *H. portusjacksoni*, necessitated the use of a range of proteolytic enzymes to obtain sufficient overlapping peptides to confirm the sequence deduced. The sequence has been compared with that previously determined for myoglobin from *H. portusjacksoni* and other animals.

Materials and Methods

The methods of peptide mapping, amino acid analysis, sulfhydryl estimation, cyanogen bromide cleavage, sequence determination by the dansyl–Edman procedure, and digestion with trypsin, chymotrypsin, thermolysin, pepsin and staphylococcal protease were substantially the same as previously described (Air and Thompson 1969, 1971; Beard and Thompson 1971; Fisher and Thompson 1979).

Isolation of Gummy Shark Myoglobin

The gummy shark red muscle was not extensive, and was mostly confined to a thin layer under the skin between the first and second dorsal fins above the lateral line. Samples from many small sharks were pooled to obtain sufficient material for the sequence work.

The sections of muscle tissue scraped from the skin were homogenized for 2 min in a blender with 2 ml of 1·5 mM KCN per gram. The homogenate was centrifuged for 10 min at 30 000 *g* and 4°C. To the supernatant was added K_3FeCN_6 (10 mg/100 ml) to convert the myoglobin to metmyoglobin before dialysing overnight versus tap water. The dialysed solution was freeze-dried, re-extracted in a small volume of water, centrifuged and loaded on a Sephadex G75 column (2·7 by 150 cm), equilibrated and eluted with 0·05 M Tris–0·05 M potassium acetate –0·001 M ethylenediaminetetraacetate–0·05 M mercaptoethanol buffer, pH 7·0. The lower molecular weight haem protein fraction, measured at 540 nm, was collected and converted to apomyoglobin as previously described (Fisher and Thompson 1979).

Fractionation of Apomyoglobin

To remove contaminating proteins present in the crude apomyoglobin fraction from gel filtration, it was fractionated in urea–thiol buffers on carboxymethyl cellulose columns (Clegg *et al.* 1966). A column of CM-cellulose (2·1 by 20 cm) was equilibrated with buffer containing 8 M urea–0·0025 M Na_2HPO_4–0·001 M EDTA (acid form)–0·05 M mercaptoethanol, pH 6·7, and loaded with 150 mg globin dissolved in the starting buffer. A linear gradient from 0·005 M Na^+ to 0·055 M Na^+, 250 ml each chamber, was used and equal-sized fractions collected. The fractions were examined by cellulose acetate electrophoresis in 8 M urea to which was added 10% (v/v) of 90% formic acid, pH 2·1.

The fractions containing apomyoglobin were pooled and carboxymethylated as previously described (Thompson *et al.* 1969).

Fractionation of Enzyme Digests

All enzyme digests of apomyoglobin were fractionated by gel filtration on Sephadex G25 or G50 columns (2·8 by 140 cm) in 0·5% (w/v) NH_4OH, with the exception of peptic digests where 5% formic acid was used.

The fractions forming a peak at 220 nm (280 nm for formic acid columns) were pooled and freeze-dried. The freeze-dried material was suspended or dissolved in pH 6·4 buffer, centrifuged at 38 000 *g* to remove insoluble material, and peptide-mapped by paper ionophoresis at pH 6·4 followed by chromatography in the second dimension with butanol–pyridine–acetic acid–water (15 : 10 : 3 : 12 v/v).

Peptides were detected by 0·02% (w/v) ninhydrin and, after washing in acetone, eluted and stored in 60% (v/v) pyridine.

Determination of the Blocking Group on the Amino Terminal of the Myoglobin Chain

The purification of amino terminal chymotryptic fragments on sulfonated polystyrene columns was done as previously described (Fisher and Thompson 1979). The chymotryptic fragment was further digested with carboxypeptidase A to remove *C*-terminal tryptophan and repurified on the sulfonated polystyrene column.

After freeze-drying, the eluate was examined by amino acid analysis of an hydrolysate, by gas chromatography–mass spectrometry as previously described (Fisher and Thompson 1979), and by proton n.m.r. spectrometry on a JEOL 99·6 MH_3 JNM-FX100 spectrometer in Fourier Transform mode using tubes (1·7 nm o.d.) in a JEOL dual $^1H/^{13}C$ mini probe; pulse width was 30 μs (90°) with the pulse repetition of 5 s. Chemical shifts are reported downfield of the sodium salt of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS).

Results

Tryptic (T), chymotryptic (C), thermolysin (Th) and staphylococcal protease (E) peptides are numbered sequentially from the *N*-terminal end.

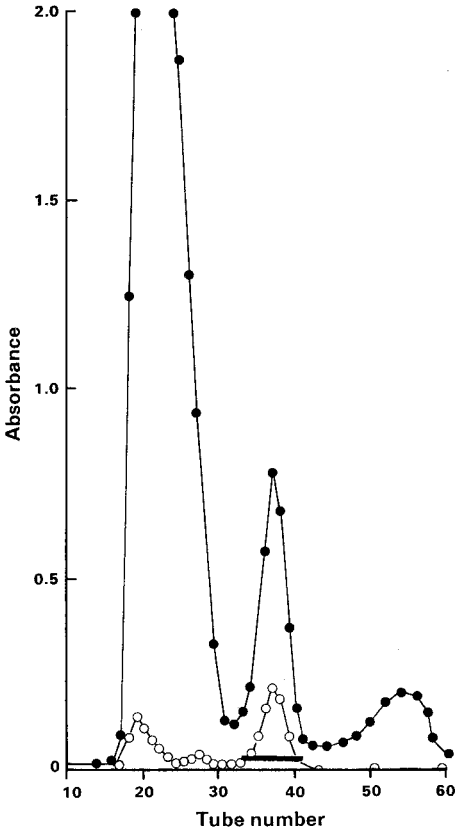


Fig. 1. Gel filtration of a red muscle extract from gummy shark in a column (150 by 2·7 cm) of Sephadex G75 in buffer containing 0·05 M Tris–0·05 M potassium acetate–0·001 M EDTA–0·05 M mercaptoethanol, pH 7. Fraction size 22 ml. Absorbance at 280 nm (●—●) and at 540 nm (○—○) measured with 2-mm light path. Bulked fractions containing myoglobin are shown by a bar.

Isolation of Shark Myoglobin

Due to the limited amounts of red muscle and the large amount of associated white muscle tissue, attempts to fractionate the extracts of red muscle by ion-exchange

chromatography on sulfopropyl-Sephadex as in previous work on myoglobins from platypus (Fisher and Thompson 1976) or *H. portusjacksoni* (Fisher and Thompson 1979) were not reproducible. Fractionation on diethylaminoethyl cellulose columns (Brown 1961) also was unreproducible, with limited binding of the myoglobin and consequent lack of fractionation from the larger amounts of protein extracted with the myoglobin from the mixture of red and white muscle tissue.

Gel filtration of the crude extract on Sephadex G75 gave a clearly separated, low molecular weight protein fraction containing myoglobin (Fig. 1). Cellogel iono-phoresis indicated some contaminant proteins that were readily removed by fraction-ation on CM-cellulose in urea–thiol buffers (Fig. 2). The major peak contained the apomyoglobin and, as seen in Fig. 2, two peaks were sometimes obtained. Peptide maps of tryptic digests of these two peak fractions indicated almost identical patterns, and it was concluded that the separation was due to partial deamidation of the myoglobin. Amino acid sequence work was mostly done on the pooled myoglobin, except when allocation of amide groups necessitated use of the more strongly bound component.

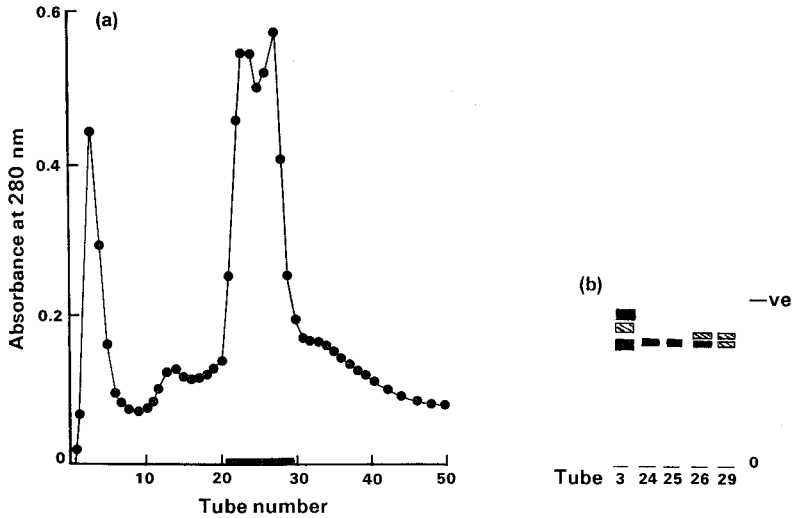


Fig. 2. (*a*) Fractionation of the gummy shark apomyoglobin fraction from Fig. 1 on a carboxymethyl cellulose column (15 by 1·8 cm) equilibrated with 8 M urea–0·05 M mercaptoethanol–0·005 M Na^+ (phosphate), pH 6·7. A linear gradient to buffer containing 0·055 M Na^+ with 250 ml of buffer in each chamber was applied. Fraction size 8·6 ml; flow rate 130 ml/h. Absorbance measured with 5-mm light path. Bulked fractions containing apomyoglobin used for sequence work are shown by a bar. Tube contents were examined by Cellogel electrophoresis in 8 M urea–formic acid, pH 2·1. (*b*) Electrophoretic patterns of tube 3 from the initial peak and tubes 24, 25, 26 and 29 of the split apomyoglobin peak respectively. Cross-hatched zones represent lesser amounts. The split apomyoglobin peak is due to partial deamidation.

Amino Acid Composition

As in other fish myoglobins, cysteine was present. The method of Ellman (1959) gave 0·9 mol/mol apomyoglobin.

The amino acid composition determined on 6 M HCl hydrolysates after hydrolysis for 24–96 h, with correction for destruction of serine and threonine, and the longest

time values to allow for incomplete liberation of valine and isoleucine residues (both Ile-Val and Ile-Ile sequences are present), agreed satisfactorily with the composition expected from the sequence. The analysis gave one less serine residue, and one more residue of lysine and glutamic acid, than the final sequence indicated. In our laboratory these residues are often subject to greater variation between analyses.

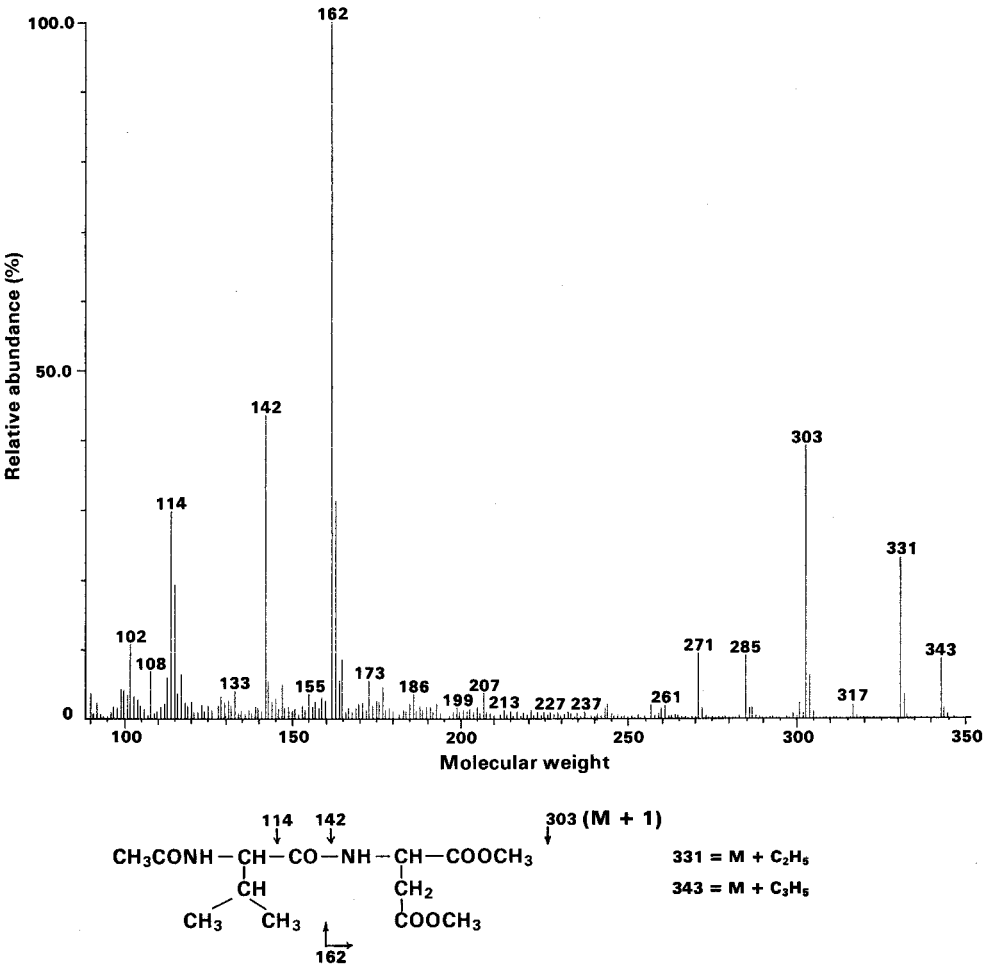


Fig. 3. Mass spectrograph trace of blocked amino terminal peptide from a chymotrypsin- and carboxypeptidase A-digested sample of gummy shark apomyoglobin. Arrows are drawn to refer to the diagrammatic representation of how the peptide Acetyl-Val-Asp, methylated with CH_2N_2, has been split in the mass spectrograph.

Determination of the Blocking Group on the Amino Terminus of Myoglobin

The chymotryptic acidic *N*-terminal peptide that was not adsorbed on sulfonated polystyrene had the composition $Asp_{1.0}\,Val_{1.0}\,Trp_{(1)}$ (Ehrlich positive) but after digestion with carboxypeptidase A and rechromatography as before the tryptophan residue was removed from the acidic peptide fraction.

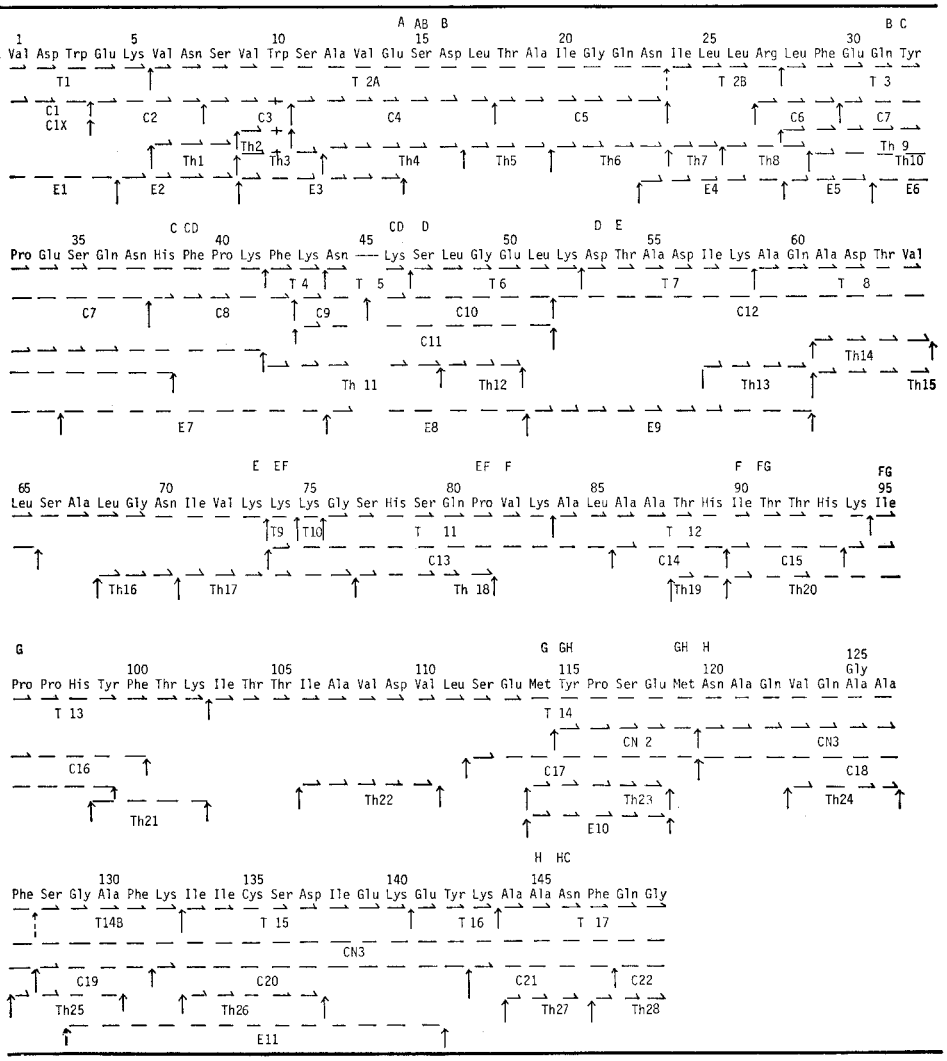
Proton n.m.r. spectroscopy gave a peak at 2·05 ppm, indicative of an acetyl blocking group. This was confirmed by the gas chromatograph–mass spectrographic analyses (Fig. 3) which showed mass ions, confirming the sequence Acetyl-Val-Asp.

Amino Acid Composition and Sequence of Peptides

Amino acid compositions of tryptic peptides, including *N*- and *C*-terminal portions of T2 split by chymotryptic-like activity, were determined. Satisfactory analyses were

Table 1. Amino acid sequence of the myoglobin of the gummy shark

Residues are numbered from the *N*-terminal residue and numbered sequentially except for one numbered gap corresponding to residue 45. The positions of cleavage by proteolytic enzymes are shown by vertical arrows. The dashed vertical arrows indicate an incomplete cleavage. Also shown are the cyanogen bromide fragments CN2–3. Residues identified by the dansyl–Edman procedure are underlined with an arrow. If no horizontal arrow is shown the residue did not give a clear result, or was not investigated



obtained for most tryptic peptides, with the larger insoluble and difficult to purify T2A, T8 and T14 peptides showing some contamination for the composition expected for the sequence reported. The sequence of the tryptic peptides is shown in Table 1.

Also shown in Table 1 is the designation and sequence of peptides obtained by digestion with other proteolytic enzymes. The quantitative amino acid composition of some of the key peptides used to confirm amino acid sequences, to confirm peptide identification after ionophoresis for amide group allocation, or to identify overlapping peptides which prove alignment of tryptic peptides, were also determined. In general, hydrolysates of peptides were examined by ionophoresis at pH 1·8 (Dreyer and Bynum 1967) and a portion of the hydrolysate was retained for later quantitative analysis if this was necessary.

In Table 1 each residue identified by the dansyl–Edman procedure is shown. The residues are numbered from the *N*-terminal residue as 1, but homology indicates that this residue is equivalent to residue 5 in mammalian myoglobins, as was the case with *H. portusjacksoni* myoglobin. Also like the other shark myoglobin, there is a missing residue in the interhelical CD region so that there is a total of 148 residues. The gap has been numbered 45 to preserve the homology with other myoglobins. This results in residues 46–149 having a number one more than the linear arrangement in the 148 residues of the shark myoglobin. The helical and interhelical regions originally allocated on the basis of X-ray diffraction (Perutz *et al.* 1965) are shown in Table 1.

Because of the relatively large number of differences in sequence from other myoglobins, including that from *H. portusjacksoni*, it proved necessary to examine many proteolytic enzyme digests to establish the sequence and the order of alignment of tryptic peptides. Consequently many peptide sequences were confirmed by several degradation experiments and the chances of errors were minimized.

The *N*-terminal peptide T1 was readily located on a peptide map by the Ehrlich stain for tryptophan, and in subsequent maps by its fluorescence under ultraviolet light. It is acetylated and not capable of degradation by the Edman method. Its amino acid sequence was originally determined on a small amount of unblocked peptide that was detected by Ehrlich stain on peptide maps. Whether this unblocked material occurs naturally or is formed in the isolation was not investigated. The sequence was confirmed, as previously described, using gas chromatography–mass spectrometry and the known specificity of trypsin for positioning the lysine residue.

Tryptic peptide 2A was insoluble and was purified from the 'core' fraction after ionophoresis at pH 1·8 followed by chromatography in the second dimension. Only the first four residues were sequenced directly by the dansyl–Edman procedure, but chymotryptic, thermolytic and staphylococcal protease peptides enabled deduction of the sequence. The absence of a basic amino acid suggested a chymotryptic-like cleavage had released residues from the *C*-terminus, later shown to be four residues due to splitting of an -Asn-Ile bond. A similar chymotryptic-like activity in TPCK-trypsin, possibly inherent, had cleaved an adjacent -Leu-Ala bond in *H. portusjacksoni* myoglobin, releasing in this case a pentapeptide.

There is only one arginine residue in gummy shark myoglobin and peptide C6 provided the overlap with tryptic peptide T3. The tryptic peptides T4–T8 can be aligned from the sequence and analysis of overlapping peptides E7, Th11, C11, E9 and C12 which account for residues 1–65. As reported in the previous paper (Fisher and Thompson 1979), the specificity of the staphylococcal protease used in our experiments resulted in hydrolysis of bonds other than glutamyl bonds, and the presence of partially deamidated glutaminyl residues may account for the splitting of the 60-61 bond shown as -Gln-Ala-.

Tryptic peptide T8 proved difficult to isolate pure, the deamidated fraction apparently being more soluble and appearing as an acidic peptide on peptide maps. Chymotryptic digestion of this material gave two peptides, T8C1 and T8C2, corresponding to a splitting of the -Leu-Ser- (65-66) bond. The amino acid analysis of the two fragments together with the analyses and sequence data for Th16 and Th17 helped complete the sequence of T8.

Further alignment of tryptic peptides T9–T13 can be deduced from the analysed peptides Th17, C13, Th20, C16 and Th21, which extends the sequence to residue 102.

There were no overlap peptides to link T13 and T14 and overall compositions of the myoglobin and its cyanogen bromide fragments, together with homology to the *H. portusjacksoni* sequence which has no tryptic-sensitive residue in this region, were used.

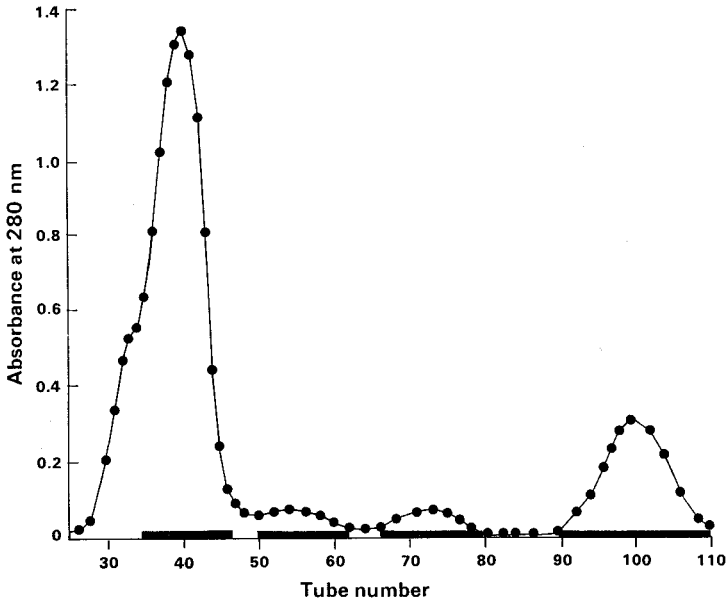


Fig. 4. Gel filtration of a cyanogen bromide digest of gummy shark apomyoglobin on a column of Sephadex G50 (120 by 2·5 cm) run in 5% formic acid. Fraction size 4·9 ml. Fractions pooled are shown by bars. Samples from peak tubes 40, 55, 74 and 100 were taken for hydrolysis and amino acid analyses.

The remainder of the alignments of tryptic peptides can be deduced by working from the *C*-terminal T17 peptide, which contains no basic amino acid and is strongly homologous to other myoglobin sequences. Peptides C20 and E11 link the large T14 peptide to T15 and T16, and a cyanogen bromide fragment links part of T14 through to T17.

Peptide T14 was insoluble in pH 6·5 pyridine–acetate buffer and either remained at the origin or streaked from the origin during peptide mapping. Direct sequencing after elution with 60% (v/v) pyridine placed residues 103–114, the remainder of the sequence being deduced from smaller thermolytic and chymotryptic peptides, and fragments obtained by cyanogen bromide cleavage of methionyl bonds following residues 114 and 119. These methionine residues are in the same positions determined for *H. portusjacksoni* myoglobin. With Met-Tyr and Met-Asn sequences, compared

with Met-Thr and Met-Gln sequences in the earlier work, better yields of the cyanogen bromide fragments were obtained.

The fractionation of a cyanogen bromide digest is shown in Fig. 4. Amino acid analyses of material in peak tubes are shown in Table 2 compared with the expected values from the sequence. Fraction 4, the low molecular weight peptide, sequenced Tyr-Pro-Ser-Glu-Hse corresponding to residues 115–119. Fractions 2 and 3 appeared to be the same, fraction 2 being a dimer as in this experiment the cysteine residue had not been carboxymethylated. Peptide mapping of tryptic digests gave peptides T14B, T15, T16 and T17 from both fractions. Pepsin digestion of fraction 3 gave the peptide Asn (-Ala-Gln-Val-Gln-Ala-Ala) which was analysed $Asp_{0.8}\,Glu_{2.0}\,Gly_{0.4}\,Ala_{2.4}\,Val_{0.9}$. The presence of glycine in this is possibly due to polymorphism at position 125 due to different animals being used for pooled red muscle. In addition, during dansyl–Edman degradations of fraction 3, and of a neutral tryptic peptide (residues 120–127) from peptide maps, both glycine and alanine were detected at the residue-125 step.

Table 2. Amino acid composition of cyanogen bromide fragments of gummy shark myoglobin

Samples analysed were from peak tubes of Fig. 4 which were hydrolysed at 110°C for 24 h under vacuum with 6 M HCl containing 0·1% phenol. Values are given as mole per mole of protein and have not been corrected for losses during hydrolysis and incomplete hydrolysis. The amino acid composition determined from sequence data is shown in parentheses

	Tube 40	CN1	Tube 55	Tube 74	CN3	Tube 100	CN2
Lysine	12·5	(12)	2·0	3·0	(3)		
Histidine	5·0	(5)					
Arginine	1·0	(1)					
Aspartic acid	12·9	(11)	3·1	2·9	(3)		
Threonine	8·2	(8)	0·4		(0)		
Serine	7·6	(9)	2·2	2·0	(2)	1·0	(1)
Glutamic acid	11·2	(11)	5·4	5·1	(5)	1·0	(1)
Proline	6·0	(5)				1·0	(1)
Glycine	4·4	(4)	2·3	2·1	(2)		
Alanine	10·6	(11)	6·0	5·9	(6)		
Valine	8·1	(9)	1·2	0·9	(1)		
Isoleucine	7·0	(8)	1·9	1·8	(3)		
Leucine	9·8	(10)	0·5		(0)		
Tyrosine	2·2	(2)	1·1	1·0	(1)	0·9	(1)
Phenylalanine	4·2	(4)	2·8	3·0	(3)		
Cysteine[A]	0	(0)	+	+	(1)		
Tryptophan	0·8	(2)					
Homoserine[B]	+	(1)				+	(1)
Total		113			30		5

[A] Detected as *S*-carboxymethylcysteine.

[B] Detected as homoserine plus its lactone.

Amino Acid Sequence of Gummy Shark Myoglobin

The complete amino acid sequence of shark myoglobin is shown in Table 1 with the amide groups assigned to particular residues. These assignments were made predominantly by ionophoretic mobility at pH 6·4 (Sanger *et al.* 1955; Ambler 1963; Offord 1966), supplemented by the evidence from the specificity of staphylococcal

protease for glutamic acid residues. The peptides and basis for assignments are included in the supplementary data.*

Amino Acid Differences between Shark Myoglobins and Myoglobins from other Animals

The number of differences between myoglobins of known amino acid sequence are shown in Table 3. The lack of four amino terminal residues in shark myoglobins has been taken as a single genetic event and counted as one difference only for the first four residues.

Table 3. Amino acid differences between myoglobins

Matrix of observed differences between myoglobins compared over 156 residues. Shark and tuna myoglobins lack four amino terminal residues; this has been taken as a single genetic event and counted as one difference only for the first four residues

	Human	Bovine	Pig	Dolphin	Harbour porpoise	Sperm whale	Red kangaroo	Opossum	Echidna	Platypus	Chicken	Penguin	Tuna	*Heterodontus* sp.	*Mustelus* sp.
Human															
Bovine	29														
Pig	10	23													
Dolphin	22	25	19												
Harbour porpoise	20	23	18	10											
Sperm whale	24	31	21	13	15										
Red kangaroo	20	34	16	27	28	29									
Opossum	15	33	14	28	28	30	14								
Echidna	25	36	23	27	25	29	22	21							
Platypus	25	36	22	28	27	30	23	19	9						
Chicken	35	45	35	38	39	42	34	37	38	37					
Penguin	45	48	42	46	46	48	42	46	46	44	28				
Tuna	80	85	81	80	80	82	83	81	81	79	79	84			
Heterodontus sp.	83	87	85	86	85	84	88	85	86	85	85	87	83		
Mustelus sp.	87	89	87	87	86	88	89	90	89	87	91	94	90	68	

Discussion

Because of the small amount of red muscle in the gummy sharks, pooled samples from different sharks were used. Variation in the amino acid residue at position 125 was found, with either glycine or alanine present. This variation probably reflects polymorphism in the population, as has been noted for other proteins such as carboxypeptidase A (Walsh *et al.* 1966) and the γ-chain of human haemoglobin (Schroeder and Huisman 1974). Partial loss of amide groups was also noted, mostly involving

* Supplementary data to this paper are deposited with and can be obtained from the Editor-in-Chief, Editorial and Publications Service, CSIRO, P.O. Box 89, East Melbourne, Vic. 3002. The data deposited contain the following material: amino acid analyses of shark apomyoglobin; amino acid analyses of tryptic peptides; elution profiles and peptide maps for tryptic, chymotryptic, thermolytic and staphylococcal protease digests; amino acid analyses of some chymotryptic, thermolytic and staphylococcal protease peptides; peptides and their net charge used for assignment of side chain amide and acidic groups.

glutaminyl sequences, and this raised problems in the assignment of amides. For example, the isolation of a peptide from a staphylococcal protease digest with a *C*-terminal glutamic acid residue will only unequivocally assign an acidic residue if the specificity of the enzyme is exact for glutamic acid residues and no deamidation has occurred during the preparation of the protein and the isolation of peptides. Similar considerations apply when measuring mobilities, as a deamidated peptide will occupy a different position on a peptide map, and unless both forms are detected wrong deductions are possible. The partial loss of amide groups has been reported by others working on myoglobin (Edmundson 1965; Satterlee *et al.* 1969; Romero-Herrera and Lehmann 1974; Fisher and Thompson 1979) and probably arise during storage or preparation procedures.

Gummy shark myoglobin shows many similarities to *H. portusjacksoni* myoglobin. Both myoglobins are 148 residues in length with an acetylated *N*-terminal residue, equivalent by the evidence of homology to residue 5 of myoglobins containing the normal 153 residues. Both myoglobins have a deletion in the CD interhelical region, which by homology has been placed at CD7 for gummy shark myoglobin. In an earlier paper (Fisher and Thompson 1979) the gap was placed at CD8 rather than CD7 although the isoleucine residue in *H. portusjacksoni* myoglobin is only one minimum base change from either leucine or lysine, which have been the invariant residues at CD7 and CD8 respectively in all previously sequenced myoglobins, and it now seems that the gap residue corresponds better to CD7.

In Table 4 a comparison of the amino acid sequence of the shark myoglobins is made with previously determined myoglobin sequences. The sequences are taken from Dayhoff (1972, 1973, 1976, 1978), Castillo and Lehmann (1977), Castillo *et al.* (1978), Romero-Herrera *et al.* (1978) and Brown *et al.* (1979). It can be seen that of the 68 differences between the two shark myoglobins there are 32 residues that are identical in the shark myoglobins but different from human myoglobin residues in corresponding positions. Of these 32 residues there are 24 unique to the shark myoglobins.

In the *H. portusjacksoni* shark myoglobin there were changes in sequence that prevented the formation of four of the seven salt bridges that help to stabilize the secondary and tertiary structure (Romero-Herrera *et al.* 1978). A similar loss of four salt bridges is noted for gummy shark myoglobin. The loss of intersegmental hydrogen bonding residues such as occurred with the other shark myoglobin is less.

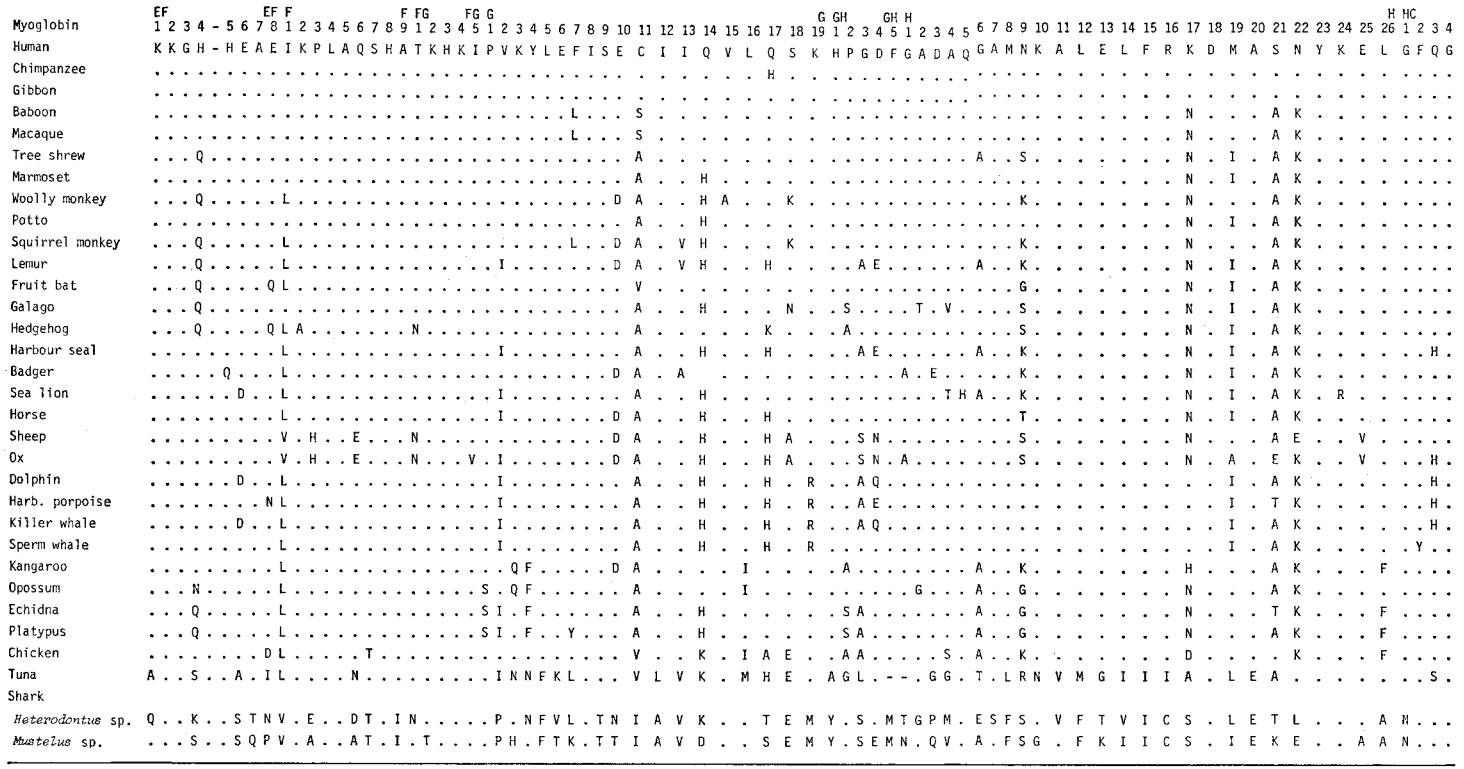
The most striking change in a residue related to the functional role of myoglobin is the absence of the E7 distal histidine residue in gummy shark myoglobin, glutamine being the replacement residue. A similar replacement has been reported previously only in α-globin from the opossum (Stenzel 1974). Of the other residues in contact with the haem group (Perutz 1976), only F7 threonine is a change from the usual residue of serine. The striking changes in residues at C2, E13 and EF1 that were noted in the Port Jackson shark are not present in gummy shark myoglobin, which has the usual residues in these positions.

With no fish myoglobin sequences for comparison when the sequence determined for *H. portusjacksoni* was reported (Fisher and Thompson 1979), the large number of changes in sequence compared with other myoglobins suggested a slow rate of change in the shark myoglobin. The number of differences between the myoglobins of gummy shark given in this paper and other myoglobins is slightly greater than the corresponding numbers for *H. portusjacksoni*. This would suggest a date of divergence no more

Table 4. Amino acid sequences of myoglobins

Helical regions A–H and interhelical regions NA, AB, CD...etc. are numbered. The one letter code and alignment follow Dayhoff (1972). The sequence of human myoglobin is given in full, dashes indicate deletions or insertions in longer globins. Dots indicate identity with the residue in human myoglobin. Ac represents an acetyl group. References are given in the text

	NA	A		B	C	CD	D	E
Myoglobin	1 2 3	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16		1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	1 2 3 4 5 6 7	1 2 3 4 5 6 7 8	1 2 3 4 5 6 7	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Human	G-L	SDGEWQLVLNVWGKVE	A	DIPGHGQEVLIRLFKG	HPETLEK	FDKFKHLK	SEDEMKA	AEDLKKHGATVL-TALGGILK
Chimpanzee	...	....	.	....	....	....	....	....
Gibbon	...	....	.	...S............	....	....	....	....
Baboon	...	....	.	...S............	....	....	....	....
Macaque	...	....	.	...S............	....	....	....	........V............
Tree shrew	...	....	.	.VA.............	....	....	T......	........N....S.......
Marmoset	...	......S.........	.	.LA.....I.....TA	....	.....N..	TP.....	........V............
Woolly monkey	...	..........I.....	.	...S.......S....	....	....	....	..E.....V............
Potto	...	......S.........	.	.LA.....I.....TA	....	.....N..	TP.....	........V............
Squirrel monkey	...	..........I.....	.	...S.......S....	....	....	....	..E.....T............
Lemur	...	....	.	.VG...........T.	....	....	TA.....	........T............
Fruit bat	...	....	.	....	....	....	....	....
Galago	...	.........KI.....	.	.LA....D......TA	....	.....N..	TA.....	........V............
Hedgehog	...	....	.	...............D	....	....	......S	........T............
Harbour seal	...	.....H..........	T	.LA............S	....	....	...D.RR	....R...N............
Badger	...	....	.	.LA.............	....	....	......G	........N............
Sea lion	...	..........I.....	.	.LV.............	....	....	......R	........K............
Horse	...	......Q.........	.	..A...........T.	....	....	T.A....	........TV...........
Sheep	...	..........A.....	.	.VA...........T.	....	....	T.A....	........N............
Ox	...	......A...A.....	.	.VA...........T.	....	....	T.A....	........N............
Dolphin	...	....	.	.VA....DI.......	....	....	T.AD...	........D........A...
Harb. porpoise	...	.E..............	.	.LA....D........	....	....	T.A....	........N............
Killer whale	...	....	.	.LA....DI.......	....	....	T.AD...	........N........A...
Sperm whale	V..	.E.......H..A...	.	.VA....DI......S	....	..R.....	T.A....	........V........A...
Kangaroo	...	..........I.....	T	.VG....D........	....	....	....	........I........N...
Opossum	...	..........A.....	.	....	....	....	....	.................N...
Echidna	...	.........K......	T	..T....D.......T	....	....	T......	.A......GV.......S...
Platypus	...	.........K......	G	.L.............T	....	.....G..	T......	.A......G........N...
Chicken	...	..Q...Q..TI.....	.	..A...H...M...HD	.....DR	.....G..	T.PD..G	........Q........AQ..
Tuna	...	.AcADFDA..KC..P..	.	.YTTM.GL..T....E	....QKL	.P..AG-I	AQADIAG	NAAISA.......KK..EL..
Shark								
Heterodontus sp.	...	.AcT..EH.NK..AV..	P	...AV.LAI.L....E	.K..KDL	.P...E-I	PVQQLGN	N...R...V....R...N...
Mustelus sp.	...	.AcVD.EK.NS..SA..	S	.LTAI..NI.L...EQ	Y..SQNH	.P...N-K	.LG.L.D	TA.I.AQADT...S...N.V.


Myoglobin
Human
Chimpanzee
Gibbon
Baboon
Macaque
Tree shrew
Marmoset
Woolly monkey
Potto
Squirrel monkey
Lemur
Fruit bat
Galago
Hedgehog
Harbour seal
Badger
Sea lion
Horse
Sheep
Ox
Dolphin
Harb. porpoise
Killer whale
Sperm whale
Kangaroo
Opossum
Echidna
Platypus
Chicken
Tuna
Shark
Heterodontus sp.
Mustelus sp.

recent than that of *H. portusjacksoni*, favouring the idea of mosaic radiation from the time of divergence, even though *Mustelus* sp. is representative of 'modern' sharks (Romer 1966). More recently, the amino acid sequence of tuna myoglobin has been completed (Brown *et al.* 1979; Rice *et al.* 1979). The number of differences between this sequence and the shark myoglobin sequences (approx. 85) is greater than between the shark sequences (68 residues), in keeping with the different characteristics and evolution of the cartilaginous and bony fishes. Of the 24 positions that were unique to the elasmobranch myoglobins only four positions have identical residues in tuna myoglobin, at positions CD3, G5, H15 and H20, but there are 11 additional positions where the tuna myoglobin shows a different residue from that present in the sharks or other myoglobins. For the myoglobins in Table 4 there are only 27 invariant residues; these are associated with internal, haem-contact, salt-linked or hydrogen-bonded residues.

Acknowledgments

This work was supported in part by the Australian Research Grants Committee. We are indebted to Drs Terry Gorman and Ken Graham, N.S.W. State Fisheries, for supplies of sharks; to Mr R. G. Mann and Miss P. A. Davidson for skilled technical assistance; and to Drs K. D. Barrow and A. M. Duffield for the n.m.r. and mass spectrograph analyses.

References

Air, G. M., and Thompson, E. O. P. (1969). Studies on marsupial proteins. II. Amino acid sequence of the β-chain of haemoglobin from the grey kangaroo, *Macropus giganteus*. *Aust. J. Biol. Sci.* **22**, 1437–54.

Air, G. M., and Thompson, E. O. P. (1971). Studies on marsupial proteins. IV. Amino acid sequence of myoglobin from the red kangaroo *Megaleia rufa*. *Aust. J. Biol. Sci.* **24**, 75–95.

Ambler, R. P. (1963). The amino acid sequence of *Pseudomonas* cytochrome C-551. *Biochem. J.* **89**, 349–78.

Beard, J. M., and Thompson, E. O. P. (1971). Studies on marsupial proteins. V. Amino acid sequence of the α-chain of haemoglobin from the grey kangaroo, *Macropus giganteus*. *Aust. J. Biol. Sci.* **24**, 765–86.

Brown, W. D. (1961). Chromatography of myoglobin on diethylaminoethyl cellulose columns. *J. Biol. Chem.* **236**, 2238–40.

Brown, W. D., Watts, D. A., and Rice, R. H. (1979). The amino acid sequence of yellowfin tuna myoglobin. *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **38**, 345.

Castillo, O., Jones, L. T., and Lehmann, H. (1978). The myoglobin of an echidna (*Tachyglossus aculeatus aculeatus*). *Biochim. Biophys. Acta* **533**, 289–92.

Castillo, O., and Lehmann, H. (1977). The myoglobin of the fruit bat (*Rousettus aegypticus*). *Biochim. Biophys. Acta* **492**, 232–6.

Clegg, J. B., Naughton, M. A., and Weatherall, D. J. (1966). Abnormal human haemoglobins. Separation and characterisation of the α- and β-chains by chromatography and the determination of two new variants. *J. Mol. Biol.* **19**, 91–108.

Dayhoff, M. O. (1972). 'Atlas of Protein Sequence and Structure.' Vol. 5. Plus Supplements 1 (1973), 2 (1976) and 3 (1978). (National Biomedical Research Foundation: Silver Spring.)

Dreyer, W. J., and Bynum, E. (1967). High-voltage paper electrophoresis. In 'Methods in Enzymology'. (Ed. C. H. W. Hirs.) Vol. XI. pp. 32–9 (Academic Press: New York.)

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

Ellman, G. L. (1959). Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **82**, 70–7.

Fisher, W. K., Nash, A. R., and Thompson, E. O. P. (1977). Haemoglobins of the shark, *Heterodontus portusjacksoni*. III. Amino acid sequence of the β-chain. *Aust. J. Biol. Sci.* **30**, 487–506.

Fisher, W. K., and Thompson, E. O. P. (1976). Studies on monotreme proteins. VII. Amino acid sequence of myoglobin from the platypus, *Ornithorhynchus anatinus*. *Aust. J. Biol. Sci.* **29**, 57–72.

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

Grant, E. M. (1978). 'Guide to Fishes.' (Department of Harbours and Marine: Brisbane.)

Nash, A. R., Fisher, W. K., and Thompson, E. O. P. (1976). Haemoglobins of the shark, *Heterodontus portusjacksoni*. II. Amino acid sequence of the α-chain. *Aust. J. Biol. Sci.* **29**, 73–97.

Offord, R. E. (1966). Electrophoretic mobilities of peptides on paper and their use in the determination of amide groups. *Nature* (*London*) **211**, 591–3.

Perutz, M. F. (1976). Structure and mechanism of haemoglobin. *Br. Med. Bull.* **32**, 195–208.

Perutz, M. F., Kendrew, J. C., and Watson, H. C. (1965). Structure and function of haemoglobin. II. Some relations between polypeptide chain configuration and amino acid sequence. *J. Mol. Biol.* **13**, 669–78.

Rice, R. H., Watts, D. A., and Brown, W. D. (1979). Sequences of the soluble tryptic peptides from myoglobin of yellowfin tuna (*Thunnus albacares*). *Comp. Biochem. Physiol.* **62B**, 481–7.

Romer, A. S. (1966). 'Vertebrate Palaeontology.' 3rd Edn. (Chicago University Press.)

Romero-Herrera, A. E., and Lehmann, H. (1974). The amino acid sequence of human myoglobin and its minor fractions. *Proc. R. Soc. London Ser. B* **186**, 249–79.

Romero-Herrera, A. E., Lehmann, H., Joysey, K. A., and Friday, A. E. (1978). On the evolution of myoglobin. *Trans. R. Soc. London Ser. B* **283**, 61–163.

Sanger, F., Thompson, E. O. P., and Kitai, R. (1955). The amide groups of insulin. *Biochem. J.* **59**, 509–18.

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

Schroeder, W. A., and Huisman, T. H. J. (1974). Multiple cistrons for fetal hemoglobin in man. *Ann. N.Y. Acad. Sci.* **241**, 70–9.

Stenzel, P. (1974). Opossum Hb chain sequence and neutral mutation theory. *Nature* (*London*) **252**, 62–3.

Thompson, E. O. P., Hosken, R., and Air, G. M. (1969). Studies on marsupial proteins. 1. Polymorphism of haemoglobin of the grey kangaroo, *Macropus giganteus*. *Aust. J. Biol. Sci.* **22**, 449–62.

Walsh, K. A., Ericsson, H., and Neurath, H. (1966). Bovine carboxypeptidase A variants resulting from allelomorphism. *Proc. Natl Acad. Sci. U.S.A.* **56**, 1339–44.

Manuscript received 7 January 1980, accepted 6 February 1980