Complete Primary Structure of the Major Component Myoglobin of California Gray Whale (*Eschrichtius gibbosus*)†

Richard A. Bogardt, Jr., Francis E. Dwulet, Lee D. Lehman, Barry N. Jones, and Frank R. N. Gurd*

ABSTRACT: The complete primary structure of the major component myoglobin from the California gray whale, *Eschrichtius gibbosus,* was determined by specific cleavage of the protein to obtain large peptides for degradation by the automatic sequenator. Cleavage at the two methionine residues of the apomyoglobin with cyanogen bromide and at the three arginine residues of the methyl acetimidated protein with trypsin resulted in three and four easily separable peptides, respectively, which when sequenced accounted for 85% of the primary structure. The remainder of the covalent structure was obtained by further digestion of the central cyanogen bromide peptide with trypsin and *S. aureus* strain V8 protease. This protein differs from that of the sperm whale, *Physeter catodon,* at 12 positions, from that of the common porpoise, *Phocoena phocoena,* and the Black Sea dolphin, *Delphinus delphis,* at 14 positions, and from that of the Amazon River dolphin, *Inia geoffrensis,* at 7 positions. All substitutions observed in this sequence fit easily into the tertiary structure of sperm whale myoglobin.

In the preceding paper (Dwulet, et al., 1975) the complete amino acid sequence for Amazon River dolphin, *Inia geoffrensis,* the first Cetacean myoglobin determined on the automatic sequenator, was reported. This paper reports the extension of the above procedures by replacing the apoprotein thermolysin digestion with the fragmentation of the central cyanogen bromide peptide with staphylococcal protease. Completion of the sequence of the myoglobin of California gray whale, *Eschrichtius gibbosus,* increases the number of complete Cetacean myoglobin primary structures known to five, including the above-mentioned Amazon River dolphin, Black Sea dolphin (Kluh and Bakardjieva, 1971), harbor porpoise (Bradshaw and Gurd, 1969), and sperm whale (Edmundson, 1965).

Materials and Methods

Protein Purification. Isolation and purification of California gray whale myoglobin were carried out as previously described (Dwulet et al., 1975).

Peptide Nomenclature. For all cleavage methods the resulting peptides are numbered from the amino terminal to the carboxyl terminal of the completed sequence. The cyanogen bromide fragments are designated by the symbol CB. Peptides isolated from the tryptic cleavage at arginine residues of the methyl acetimidated protein are labeled with the symbol MT. Fragmentation of the middle cyanogen bromide peptide, CB2, with trypsin and staphylococcus protease yielded peptide products labeled as CB2-T and CB2-S, respectively.

Cyanogen Bromide Cleavage. Cleavage of apomyoglobin with cyanogen bromide was accomplished as previously described (Dwulet et al., 1975).

Cleavage at Arginine Residues. Preparation of methyl acetimidated myoglobin (Garner and Gurd, 1975) and digestion of the reacted apoprotein with trypsin (TPCK Worthington) were carried out as previously described (Dwulet et al., 1975).

Cleavage of CB2 (56–131) with Trypsin. Cleavage of the middle cyanogen bromide peptide with trypsin was obtained as previously described (Dwulet et al., 1975).

Cleavage of CB2 (56–131) with Staphylococcal Protease. The central cyanogen bromide peptide was cleaved specifically at glutamic acid residues with an extracellular protease from *Staphylococcus aureus* strain V8 (staphylococcal protease, Miles) according to the method of Houmard and Drapeau (1972). A sample of 17 mg (2 μmol) of CB2 (56–131) was dissolved in 2 ml of 0.1 M ammonium bicarbonate buffer, pH 7.8, 37 °C. To this solution 0.5 ml of the above buffer containing 1 mg (500 units) of staphylococcus protease was added. After 12 h another aliquot of 0.5 ml of buffer containing 1 mg of enzyme was added, and at the end of 24 h the reaction mixture was lyophilized. The resulting mixture was purified on a phosphocellulose column eluted with a linear pyridine–acetate gradient (Bradshaw et al., 1969).

Time Course Hydrolysis of Peptides with Carboxypeptidase C. Digestion of peptides with carboxypeptidase C (Tschesche and Kupfer, 1972; Garner et al., 1974) was carried out on 200–300 nmol of peptide dissolved in 1 ml of 0.05 M sodium citrate buffer, pH 5.7, containing 1.5 units of carboxypeptidase C (Rohm and Haas, Darmstadt). Samples (0.1 ml) were removed at appropriate times and hydrolysis was terminated by the addition of 0.4 ml of diluting buffer, pH 2.0 (Stand In^R, Beckman). Samples were kept frozen until analysis upon an automated Beckman 120B amino acid analyzer (Spackman et al., 1958) with automatic integration performed with a Texas Instruments 980A minicomputer.[6]

Sequencing Procedures. All peptides used to determine this sequence were subjected to automated Edman degradations on a Beckman 890C Sequencer. The fast peptide-DMAA[1] program (071872, Beckman Instruments) was used to sequence peptides of 25 residues or less, and the fast protein-quadrol program (072172C) was used to sequence the protein and the larger peptides. All peptides which had a free ε-amino

† From the Department of Chemistry, Indiana University, Bloomington, Indiana 47401. *Received December 31, 1975.* This is the 74th paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper see Morrow et al. (1976). Thiw work was supported by Public Health Service Research Grant HL-05556. R. A. Bogardt, F. E. Dwulet, and L. D. Lehman were supported by Public Health Service Grant T01 GM 1046-14.

[1] Abbreviations used are: DMAA, dimethylallylamine; m-SPITC, 3-sulfophenyl isothiocyanate, sodium salt.

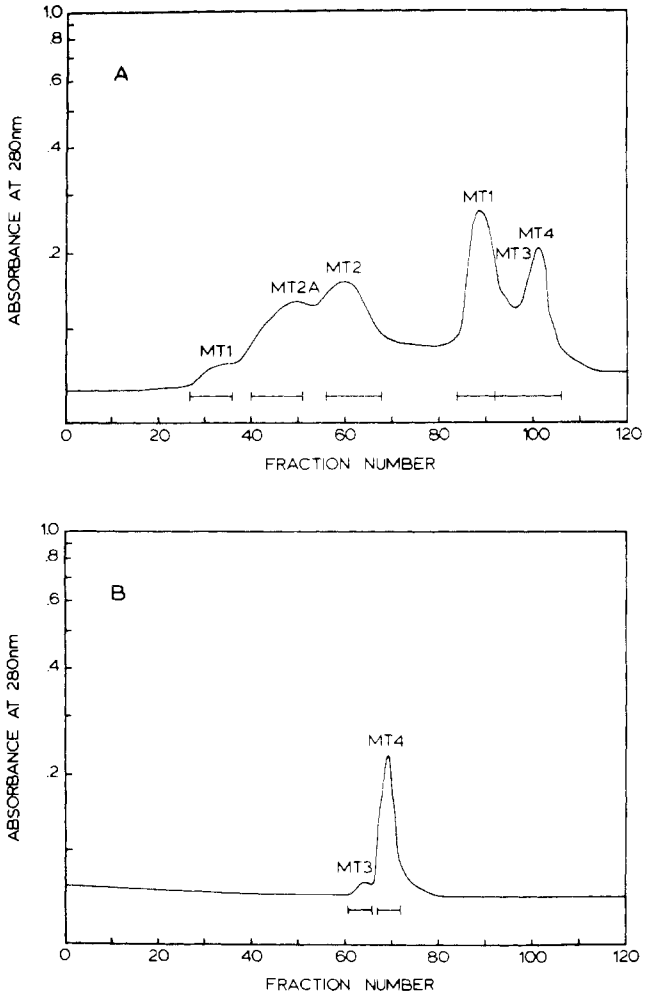

FIGURE 1: (A) Gel filtration pattern for the purification of the peptides obtained by tryptic digestion of the methyl acetimidated apomyoglobin. The peptide mixture was applied to a 2.6 × 200-cm column of Bio-Gel P-10 (200–400 mesh). The peptides were eluted with 10% acetic acid at a flow rate of 30 ml/h and a fraction size of 5 ml. (B) Gel filtration pattern for the repurification of peptides MT3 and MT4 on a 2.6 × 186-cm column of Bio-Gel P-6 (200–400 mesh) eluted with 10% acetic acid at a flow rate of 36 ml/h with a fraction size of 6 ml.

lysine were first coupled with m-SPITC to reduce extraction losses (Dwulet et al., 1975). The amino acid phenylthiohydantoins were identified on a Hewlett-Packard 5711A gas chromatograph, by thin-layer chromatography, or on the amino acid analyzer after acid hydrolysis as previously described (Dwulet et al., 1975).

Results

Amino Acid Composition. The amino acid composition of the principal component of California gray whale myoglobin was obtained from 24-, 48-, and 72-h acid hydrolysates of the ferrimyoglobin. The results are summarized in Table I.

Separation of Cyanogen Bromide Cleaved Peptides. Peptides from the CNBr digest were purified by gel filtration on Bio-Gel P-10 (200–400 mesh) as previously described (Marshall et al., 1974; Dwulet et al., 1975).[2] The amino acid compositions of these peptides are shown in Table II.

Separation of Specifically Cleaved Arginine Peptides. Peptides from the tryptic digest of methyl acetimidated apomyoglobin were initially purified by gel filtration on a column of Bio-Gel P-10 as seen in Figure 1. Compared with the results for myoglobins of other Cetacea (Dwulet et al., 1975),[3,4] a decrease in the yields of peptides MT2 and MT3 was observed repeatedly, along with the appearance of a new peptide MT2A, which resulted from incomplete cleavage at

TABLE I: Amino Acid Composition of *Eschrichtius gibbosus* Myoglobin.

Amino Acid	Number of Residues from Acid Hydrolysates[a]	Number of Residues from the Sequence
Asp	12.1	12
Thr	5.0	5
Ser	5.1	5
Glu	17.0	17
Pro	3.7	4
Gly	10.9	11
Ala	18.2	18
Val	6.1	6
Met	1.9	2
Ile	9.8	10
Leu	18.1	18
Tyr	1.8	2
Phe	6.7	7
Lys	19.9	20
His	11.1	11
Arg	2.8	3
Trp[b]	1.7	2

[a] Acid hydrolyses were performed on ferrimyoglobin for 24, 48, and 72 h at 110 °C with 5.7 N HCl and the values were averaged. The amino acid residues were calculated on the basis of 153 amino acids in the protein. The values of threonine and serine were obtained by extrapolation to zero time. The values of valine, isoleucine, and leucine were the maximum values (72 h). [b] Tryptophan was determined by the method of Liu and Chang (1971).

TABLE II: Amino Acid Composition[a] of the Cyanogen Bromide Cleaved Peptides.

Amino Acid	CB1	CB2	CB3
Asp	5.3 (5)	4.9 (5)	2.0 (2)
Thr	2.1 (2)	2.8 (3)	
Ser	1.1 (1)	3.9 (4)	
Glu	8.1 (8)	6.1 (6)	3.1 (3)
Pro	0.9 (1)	3.1 (3)	
Gly	3.0 (3)	6.2 (6)	1.9 (2)
Ala	5.2 (5)	10.1 (10)	3.0 (3)
Val	3.9 (4)	2.1 (2)	
Ile	2.8 (3)	5.4 (6)	0.9 (1)
Leu	6.9 (7)	8.2 (8)	2.9 (3)
Tyr		0.9 (1)	0.9 (1)
Phe	3.1 (3)	2.0 (2)	1.9 (2)
Lys	6.2 (6)	10.1 (10)	3.9 (4)
His	3.0 (3)	7.8 (8)	
Arg	0.9 (1)	1.0 (1)	0.9 (1)
Trp[b]	1.8 (2)		
Hse	0.8 (1)	0.6 (1)	
Total residues	55	76	22
Yield	74%	58%	80%
Position	1–55	56–131	132–153

[a,b] See footnotes in Table I.

[2] Results of established procedures can be found in supplementary material, as described below.

[3] L. D. Lehman, work in progress.

[4] B. N. Jones, work in progress.

TABLE III: Amino Acid Composition[a] of Peptides Cleaved at the Arginines.

Amino Acid	MT1	MT2A	MT2	MT3	MT4
Asp	4.0 (4)	7.0 (7)	3.9 (4)	3.2 (3)	1.2 (1)
Thr		5.2 (5)	4.8 (5)		
Ser	1.0 (1)	3.9 (4)	4.0 (4)		
Glu	3.9 (4)	11.3 (11)	9.1 (9)	2.0 (2)	2.0 (2)
Pro		3.7 (4)	2.6 (3)	1.2 (1)	
Gly	1.9 (2)	7.1 (7)	5.0 (5)	2.3 (2)	1.9 (2)
Ala	4.1 (4)	11.8 (12)	7.1 (7)	4.8 (5)	2.0 (2)
Val	3.8 (4)	2.1 (2)	1.8 (2)		
Met		1.6 (2)	0.7 (1)	0.9 (1)	
Ile	2.8 (3)	5.3 (6)[b]	5.1 (6)[b]		1.0 (1)
Leu	4.2 (4)	12.8 (13)	10.8 (11)	1.9 (2)	1.1 (1)
Tyr		1.1 (1)	0.9 (1)		0.8 (1)
Phe		5.7 (6)	3.7 (4)	2.1 (2)	1.1 (1)
Lys	1.2 (1)	16.2 (16)	14.8 (15)	1.0 (1)	2.8 (3)
His	1.3 (1)	9.2 (10)	8.8 (9)	0.9 (1)	
Arg	1.1 (1)	1.8 (2)	1.0 (1)	0.9 (1)	
Trp[c]	1.8 (2)				
Total residues	31	108	87	21	14
Yield	61%	31%	67%	71%	97%
Position	1-31	32-139	32-118	119-139	140-153

[a] Amino acid compositions were determined on hydrolysates at 110 °C for 24 h in a sealed evacuated tube containing constant boiling HCl (5.7 N). Destruction of serine, threonine, and tyrosine was not corrected for. The number of residues per molecule of peptide found is given along with the integral values (in parentheses). [b] The value for isoleucine is low because there is an Ile-Ile bond in this peptide which is only partially hydrolyzed in 24 h of hydrolysis. [c] See footnote in Table I.

arginine residue 118. Peptides MT3 and MT4 were separated on a column of Bio-Gel P-6. The amino acid compositions of these peptides are shown in Table III.

Tryptic Peptides of CB2. These peptides were purified as previously described (Dwulet et al., 1975). The amino acid compositions of these peptides are shown in Table IV.

Protease Peptides of CB2. Purification of the peptides resulting from the digestion of CB2 with staphylococcus protease was achieved by ion-exchange chromatography on Cellex P (Bio-Rad) using a linear gradient of pyridine acetate (pH 2.5-5.0), as shown in Figure 2. All peptides were obtained pure. The position and composition of each peptide are reported in Table V.

Sequence Investigation. Only the sequence data necessary to establish the entire primary structure are reported here.

Sequenator Results. The complete primary structure of California gray whale myoglobin is shown in Figure 3. In all the sequenator runs the yields for the phenylthiohydantoins were near the values expected except for those of threonine, serine, and the acids and amides. Reasons for these low yields have been discussed previously (Dwulet et al., 1975). All data presented are uncorrected for these low yields. As was also observed previously, the presence of a Lys-Pro bond at sequence positions 87 and 88 caused an appreciable drop in repetitive yield and increased carryover in continuing cycles in sequenator runs for peptides CB2 and CB2-S4.

Sequenator analysis A (Figure 4) yielded the first 34 amino terminal residues of the intact protein to give three cycles past the first arginine residue. The amino-terminal residues of sequenator analysis B of peptide MT2 overlapped the intact protein analysis and extended the sequence 23 residues to position 57. Sequenator analysis C of peptide CB2 overlapped analysis B and extended the sequence 42 cycles to position 100. Analysis D of peptide CB2-S4 overlapped analysis C starting at position 86 in the sequence and extended the sequence past position 100 to position 104, reconfirming the acid and amide assignments in this area, as well as those of serine and threonine. Peptide CB2-T6 (sequenator analysis E) overlapped analysis D with the only tyrosine residue in the central cyanogen bromide fragment and extended the sequence to position 118, which is also the only arginine residue of CB2.

TABLE IV: Amino Acid Composition[a] of Tryptic Peptides from CB2.

Amino Acid	CB2 T1	CB2 T1A	CB2 T1B	CB2 T2	CB2 T3	CB2 T4A	CB2 T4	CB2 T5	CB2 T6	CB2 T7
Asp	1.1 (1)	1.0 (1)		1.0 (1)					1.0 (1)	2.1 (2)
Thr				1.9 (2)		0.9 (1)	1.0 (1)			
Ser	0.7 (1)	0.6 (1)				1.0 (1)	1.2 (1)		1.6 (2)	
Glu	1.0 (1)	0.8 (1)				3.2 (3)	3.2 (3)		0.9 (1)	1.2 (1)
Pro						1.2 (1)	0.9 (1)	0.7 (1)		1.1 (1)
Gly				3.2 (3)		1.2 (1)	1.0 (1)			2.1 (2)
Ala	1.1 (1)	0.9 (1)		1.1 (1)		3.0 (3)	3.1 (3)		1.1 (1)	4.0 (4)
Val				1.2 (1)					1.2 (1)	
Ile				0.9 (1)				1.8 (2)	2.2 (3)[b]	
Leu	1.0 (1)	1.1 (1)		3.1 (3)		2.1 (2)	2.0 (2)		2.1 (2)	
Tyr									0.9 (1)	
Phe									1.0 (1)	1.0 (1)
Lys	2.7 (3)	1.6 (2)	1.0 (1)	1.1 (1)	2.0 (2)	3.1 (3)	1.8 (2)	2.1 (2)		
His				1.1 (1)		2.9 (3)	2.7 (3)	1.0 (1)	2.0 (2)	0.8 (1)
Arg									0.9 (1)	
Hse										0.6 (1)
Total residues	8	7	1	14	2	18	17	6	16	13
Yield	35%	51%	100%	50%	50%	30%	43%	48%	42%	91%
Position	56-63	56-62	63 or 78	64-77	78-79	79-96	80-96	97-102	103-118	119-131
Pool	CB2 TV	CB2 TIV	CB2 TII	CB2 TIII	CB2 TVI	CB2 TVIII	CB2 TVII-1	CB2 TVII-2	CB2 T insol.	CB2 TI

[a,b] See Table III.

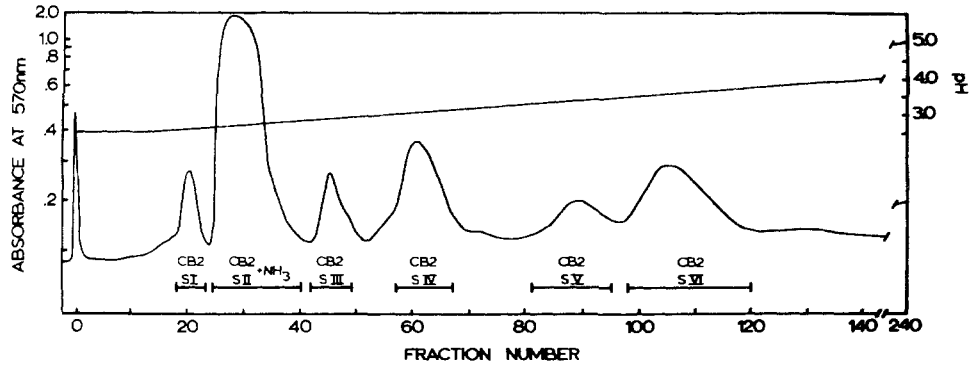

FIGURE 2: The elution pattern of the staphylococcus protease peptides of fragment CB2 on a 1.8 × 18-cm column of Cellex P maintained at 4 °C. The column was equilibrated with 0.05 M pyridine acetate (pH 2.5) and pumped at a flow rate of 30 ml/h with a fraction size of 3 ml. The column was then developed with a 24-h linear gradient of 0.05 M (pH 2.5) to 2.0 M pyridine acetate (pH 5.0). The column was monitored by automatic alkaline hydrolysis and ninhydrin analysis.

```
                5                  10                  15
 1  Val Leu Ser Asp Ala Glu Trp Gln Leu Val Leu Asn Ile Trp Ala
16  Lys Val Glu Ala Asp Val Ala Gly His Gly Gln Asp Ile Leu Ile
31  Arg Leu Phe Lys Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys
46  Phe Lys His Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp
61  Leu Lys Lys His Gly Asn Thr Val Leu Thr Ala Leu Gly Gly Ile
76  Leu Lys Lys Lys Gly His His Glu Ala Glu Leu Lys Pro Leu Ala
91  Gln Ser His Ala Thr Lys His Lys Ile Pro Ile Lys Tyr Leu Glu
106 Phe Ile Ser Asp Ala Ile Ile His Val Leu His Ser Arg His Pro
121 Gly Asp Phe Gly Ala Asp Ala Gln Ala Ala Met Asn Lys Ala Leu
136 Glu Leu Phe Arg Lys Asp Ile Ala Ala Lys Tyr Lys Glu Leu Gly
151 Phe Gln Gly
```

FIGURE 3: The amino-acid sequence of California gray whale myoglobin. The hyphens between the amino acid residues have been omitted for clarity.

Analysis F of CB2-S5B overlapped analysis E starting at position 110 and extended the sequence to position 123, overlapping arginine 118. Analysis G (peptide MT3) also overlapped peptide CB2-S5B and extended the sequence from residue 119 past methionine 131 to sequence position 134. The final sequenator analysis H of CB3 overlapped analysis G starting at position 132 and extended to the carboxyl terminus of the protein at position 153.

Because of the low repetitive yield of sequenator peptide analysis D, despite clear cut results with little carryover, the carboxyl terminal sequence of peptide CB2-S4 was reconfirmed by time course digestion with carboxypeptidase C (Figure 5).

Discussion

The present report is the second in a series[3,4] of complete Cetacean myoglobin sequences determined by automated Edman degradation. The strategy employed is the isolation of a minimal number of peptides by specific chemical or enzymatic cleavage. The cleavages at the two methionine and the three arginine residues allowed for the automated sequencing of 85% of the myoglobin with five sequenator runs, as previously discussed (Dwulet et al., 1975). The emphasis here will be on the enzymatic fragmentation of the middle cyanogen bromide fragment, CB2, which allowed for the determination of the remaining 15% of the sequence with the necessary overlaps to be completed with three peptides: CB2-T6, CB2-S4, and CB2-S5B.

TABLE V: Amino Acid Composition[a] of Staphylococcus Protease Peptides of CB2.

Amino Acid	CB2 S1	CB2 S2	CB2 S3	CB2 S4	CB2 S5A	CB2 S5B
Asp		1.9 (2)			1.2 (1)	1.9 (2)
Thr		1.9 (2)		1.2 (1)		
Ser	0.8 (1)			0.8 (1)	1.0 (1)	0.9 (1)
Glu	1.0 (1)	1.0 (1)	1.1 (1)	1.9 (2)		0.9 (1)
Pro				1.8 (2)		1.2 (1)
Gly		3.7 (4)				2.2 (2)
Ala	1.0 (1)	0.9 (1)	0.9 (1)	2.1 (2)		4.9 (5)
Val		1.0 (1)				1.1 (1)
Ile		1.1 (1)		1.8 (2)	1.0 (1)	0.9 (2)[b]
Leu		3.8 (4)		3.0 (3)		1.3 (1)
Tyr				0.8 (1)		
Phe					0.9 (1)	0.8 (1)
Lys	1.1 (1)	5.1 (5)		4.2 (4)		
His		3.2 (3)		2.2 (2)		2.8 (3)
Arg						0.9 (1)
Hse						0.7 (1)
Total residues	4	24	2	20	4	22
Yield	47%	38%	26%	22%	25%	30%
Position	56–59	60–83	84–85	86–105	106–109	110–131
Pool	CB2 SIII	CB2 SVI	CB2 SII	CB2 SV	CB2 SI	CB2 SIV

[a,b] See Table III.

As the tryptic digest of CB2 was discussed in the previous paper in this series, the discussion here will center around the staphylococcus protease digestion of CB2. Drapeau et al. (1972) reported the isolation of an enzyme from the culture filtrates of *S. aureus* strain V8 that cleaved at acid residues with a specificity similar to that of trypsin for basic residues. It was further reported that under controlled conditions the protease would specifically cleave at glutamic acid bonds with the single exception of Asp–Gly bonds (Houmard and Drapeau, 1972). Welling et al. (1975), however, have reported cleavage with this protease at Asp–Ala and Asp–Asn bonds in the structural determination of dromedary and kangaroo ribonucleases. Such an Asp–Ala cleavage was found at positions 109 and 110 in gray whale myoglobin. This cleavage was the only uncharacteristic cleavage found in the digest. Peptide yields were similar to those found from a staphylococcus protease cleavage on horse myoglobin (Houmard and Drapeau, 1972).

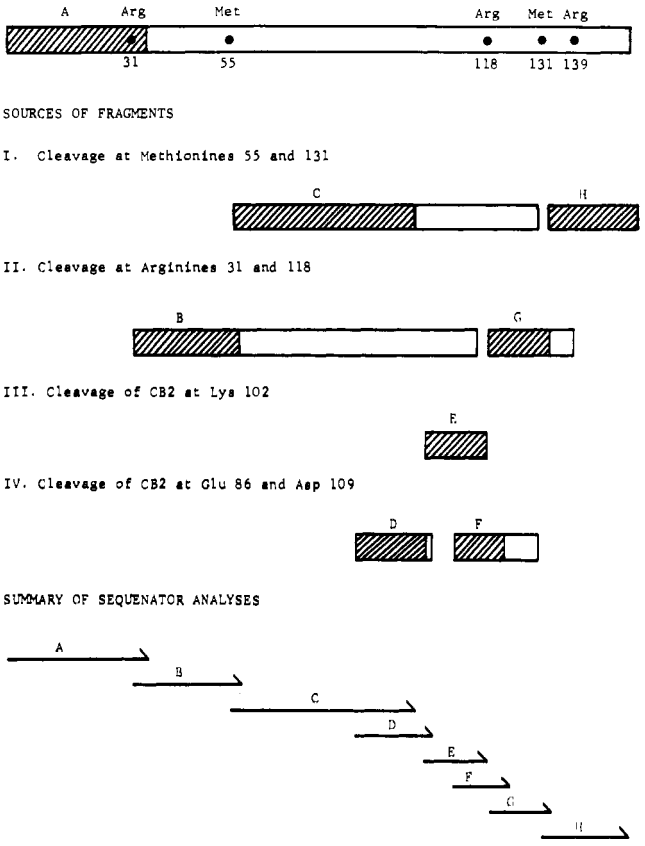

FIGURE 4: Diagrammatic summary of fragments generated from the California gray whale myoglobin for sequenator analysis. The top bar represents the whole myoglobin and the residues that are important for its fragmentation. The capital letters A–H identify the sequenator analyses in the order in which they are described in the test. A hatched section in each horizontal bar indicates the segment of sequence determined by that. analysis. A summary of overlaps is shown in the lower portion by the labeled arrows.

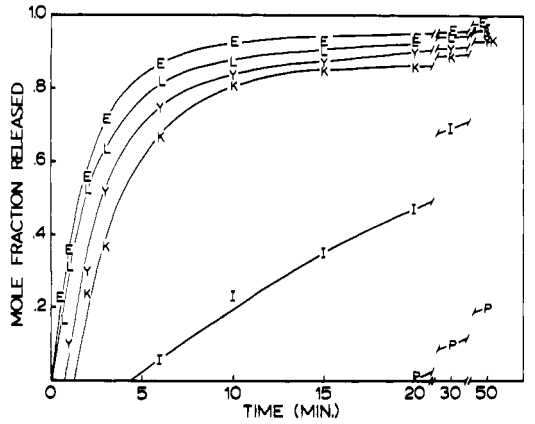

FIGURE 5: Time course enzymatic digestion of peptide CB2-S4 with carboxypeptidase C. Conditions are described in the text. The sequence found was: Pro-Ile-Lys-Tyr-Leu-Glu. Points on the curves are represented by the appropriate amino acid single letter code.

The sequence of California gray whale myoglobin is compared in Figure 6 with the known cetacean myoglobins, those of sperm whale, Amazon River dolphin, common porpoise, and the Black Sea dolphin. As seen from the difference matrix in Figure 7, gray whale myoglobin has the closest similarity in sequence to another Balaenoptera whale, the minke, and to the Amazon River dolphin, a fresh water species. The sequence of gray whale myoglobin will be examined here in comparison

Residue Number	1	4	5	12	13	15	21	26	28	35	45
Species											
Gray Whale	Val	Asp	Ala	Asn	Ile	Ala	Val	Gln	Ile	Gly	Lys
Sperm Whale	Val	Glu	Gly	His	Val	Ala	Val	Gln	Ile	Ser	Arg
A.R. Dolphin	Gly	Asp	Gly	Asn	Ile	Gly	Leu	Gln	Val	Gly	Lys
Porpoise	Gly	Glu	Gly	Asn	Val	Gly	Leu	Gln	Val	Gly	Lys
Dolphin	Gly	Asp	Gly	Asn	Val	Gly	Val	Glu	Ile	Gly	Lys

Residue Number	54	66	74	83	85	109	121	122	129	151	152
Species											
Gray Whale	Glu	Asn	Gly	Glu	Glu	Asp	Gly	Asp	Ala	Phe	Gln
Sperm Whale	Glu	Val	Ala	Glu	Glu	Glu	Gly	Asn	Gly	Tyr	Gln
A.R. Dolphin	Glu	Asn	Gly	Glu	Glu	Glu	Gly	Asp	Ala	Phe	His
Porpoise	Glu	Asn	Gly	Glu	Asn	Glu	Ala	Glu	Gly	Phe	His
Dolphin	Asp	Asp	Ala	Asp	Glu	Glu	Ala	Gln	Gly	Phe	His

FIGURE 6: Comparison of the amino acid sequences of Cetacean myoglobins whose sequences have been completed to date. Only those positions in which differences occur are reported. All other positions are conserved. A. R. dolphin is Amazon River dolphin.

	GRAY WHALE	MINKE WHALE	AMAZON RIVER DOLPHIN	COMMON PORPOISE	BLACK SEA DOLPHIN	
SPERM WHALE	12	14	15	15	14	
GRAY WHALE		3	7	14	14	
MINKE WHALE			10	13	14	
AMAZON RIVER DOLPHIN				7	11	
COMMON PORPOISE					11	

FIGURE 7: Difference matrix for Cetacean myoglobins obtain by summing the number of different amino acids between pairs of proteins.

to the seven differences in sequence from the Amazon River dolphin. These will be referred to with the residue found in California gray whale myoglobin given first, after the position number, followed by the homologous Amazon River dolphin residue in parentheses.

Position 1 Valine (Glycine). Glycine is the common amino-terminal residue found in the majority of known myoglobins. Gray whale is the third published case of a valine residue (Edmundson, 1965; Edman and Begg, 1967). Valine is a common amino-terminal residue for Balaenoptera whales (Edman and Begg, 1967).[3,5]

Position 5 Alanine (Glycine). This is the first complete myoglobin sequence to report alanine at position 5. Only the partial sequences of the myoglobins of the coelacanth, *Latimeria chalumnae* (Chauvet and Acher, 1972), and the

[5] M. T. Rothgeb, work in progress.
[6] R. A. Bogardt, work in progress.

humpback whale, *Megaptera novaeangliae* (Edman and Begg, 1967), have this residue. The alanine residue has been found commonly in myoglobins of other Balaenoptera whales.[3,5]

Position 15 Alanine (Glycine). This residue appears only in the myoglobin sequences from whales.

Position 21 Valine (Leucine) and Position 28 Isoleucine (Valine). These two positions appear to be examples of Fitch's "covarions" (Fitch and Markowitz, 1970). Covarions are concomitantly variable codons represented by a limited set of amino acids within a protein structure. In the majority of myoglobin sequences known, an increase in side-chain volume of residue 21 is accompanied by a decrease in side-chain volume at position 28. These residues are not in direct contact with each other within the tertiary structure of the sperm whale protein (Watson, 1969). Residue 21 in sperm whale myoglobin lies in a crevice and residue 28 is found to be inaccessible to solvent. Position 28, furthermore, is a residue involved with one of the seven internal cavities described by Lee and Richards (1971).

Position 109 Aspartic Acid (Glutamic Acid). This is a common residue in the Balaenoptera whales[3,5] and in the myoglobin sequences of ox (Han et al., 1970), sheep (Han et al., 1972), and pig (Floc'h et al., 1973).

Position 152 Glutamine (Histidine). Glutamine is the common residue for this position. Histidine at this position is found only in myoglobins from porpoises and dolphins, the pinnipedia such as harbor seal (Bradshaw and Gurd, 1969), and California sea lion (Vigna et al., 1974), and the ox. This substitution is the only charge change seen between the California gray whale and the Amazon River dolphin myoglobins.

All of the above changes are considered conservative. All comparisons made are in terms of the main component myoglobin from skeletal muscle of the given species. The yields of minor myoglobin components from the California gray whale muscle tissue are comparable to those from the sperm whale (Hapner et al., 1968; Garner et al., 1974).

Acknowledgments

The authors are grateful to Mrs. K. Verhey, Mrs. L. Gurd, Miss C. Bowman, and Miss K. Summers for their excellent technical help. We are grateful to Drs. William H. Garner and Edward Mitchell for advice. The cooperation is gratefully acknowledged of the National Marine Fisheries Services in making material available during a survey of the gray whale stock.

Supplementary Material Available

Tables and figures containing additional data on structure determination as noted in the text (14 pages). Ordering information is given on any current masthead page.

References

Bradshaw, R. A., Garner, W. H., and Gurd, F. R. N. (1969), *J. Biol. Chem. 244,* 2149.

Bradshaw, R. A., and Gurd, F. R. N. (1969), *J. Biol. Chem. 244,* 2167.

Chauvet, J., and Acher, R. (1972), *FEBS Lett. 28,* 16.

Drapeau, G. R., Baily, Y., and Houmard, J. (1972), *J. Biol. Chem. 247,* 6720.

Dwulet, F. E., Bogardt, R. A., Jones, B. N., Lehman, L. D., and Gurd, F. R. N. (1975), *Biochemistry 14,* 5336.

Edman, P., and Begg, J. (1967), *Eur. J. Biochem. 1,* 80.

Edmundson, A. B. (1965), *Nature (London) 205,* 389.

Fitch, W. M., and Markowitz, E. (1970), *Biochem. Genet. 4,* 579.

Floc'h, R., Dautrevaux, M., and Han, K. (1973), *Biochemie 55,* 95.

Garner, M. H., Garner, W. H., and Gurd, F. R. N. (1974), *J. Biol. Chem. 249,* 1513.

Garner, W. H., and Gurd, F. R. N. (1975), *Biochem. Biophys. Res. Commun. 63,* 262.

Han, K., Dautrevaux, M., Chaila, X., and Biserte, G. (1970), *Eur. J. Biochem. 16,* 465.

Han, K., Tetaert, D., Maschetto, Y., Dautrevaux, M., and Kapeyou, C. (1972), *Eur. J. Biochem. 27,* 585.

Hapner, K. D., Bradshaw, R. A., Hartzell, C. R., and Gurd, F. R. N. (1968), *J. Biol. Chem. 243,* 683.

Houmard, J., and Drapeau, G. R. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 3506.

Kluh, I., and Bakardjieva, A. (1971), *FEBS Lett. 17,* 31.

Lee, B., and Richards, F. M. (1971), *J. Mol. Biol. 55,* 379.

Liu, T. Y., and Chang, Y. T. (1971), *J. Biol. Chem. 246,* 2842.

Marshall, R. C., Jones, W. C., Vigna, R. A., and Gurd, F. R. N. (1974), *Z. Naturforsch. Teil C: 29,* 90.

Morrow, J. S., Matthew, J. B., Wittebort, R. J., and Gurd, F. R. N. (1976), *J. Biol. Chem.,* (in press).

Spackman, D. H., Moore, S., and Stein, W. H. (1958), *Anal. Chem. 30,* 1190.

Tschesche, H., and Kupfer, S. (1972), *Eur. J. Biochem. 26,* 33.

Vigna, R. A., Gurd, L. J., and Gurd, F. R. N. (1974), *J. Biol. Chem. 249,* 4144.

Watson, H. C. (1969), *Prog. Stereochem. 4,* 299.

Welling, G. W., Groen, G., and Beintema, J. J. (1975) *Biochem. J. 147,* 505.