The Complete Amino Acid Sequence of the Major Component Myoglobin of Amazon River Dolphin (*Inia geoffrensis*)†

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

ABSTRACT: The complete amino acid sequence of the major component myoglobin from Amazon River dolphin, *Inia geoffrensis*, was determined by specific cleavage of the protein to obtain large peptides which are readily degraded by the automatic sequencer. Three easily separable peptides were obtained by cleaving the protein with cyanogen bromide at the methionine residues and four peptides were obtained by cleaving the methyl-acetimidated protein with trypsin at the arginine residues. From these peptides over 85% of the sequence was completed. The remainder of the sequence was obtained by fragmentation of the large cyanogen bromide peptide with trypsin. This protein differs from that of the common porpoise, *Phocoena phocoena*, at seven positions, from that of the common dolphin, *Delphinus delphis*, at 11 positions, and from that of the sperm whale, *Physeter catodon*, at 15 positions. By comparison of this sequence with the three-dimensional structure of sperm whale myoglobin it appears that those residues close to the heme group are most conserved followed by those in nonhelical regions and lastly by those in the helical segments. All of the substitutions observed in this sequence fit easily into the three-dimensional structure of the sperm whale myoglobin.

In preceding papers the complete amino acid sequences were reported for California sea lion myoglobin (Vigna et al., 1974) and for harbor seal and porpoise myoglobin (Bradshaw and Gurd (1969)). This paper reports the peptide fragmentation and analytical procedures used to determine the complete sequence of the main component myoglobin from another species of Cetacea, *Inia geoffrensis* (Amazon River dolphin). Completion of this sequence extends the number of complete Cetacean myoglobin sequences to four, including sperm whale (Edmundson, 1965), harbor porpoise (Bradshaw and Gurd, 1969), and common dolphin (Kluh and Bakardjieva, 1971). Specific chemical and enzymatic cleavages at methionine and arginine residues produced peptides which were readily degraded in the automatic sequencer. Tryptic peptides of the middle cyanogen bromide fragment were also isolated to complete the sequence of that fragment. In addition, the apoprotein was digested with thermolysin to obtain all necessary overlap peptides and to reconfirm serine and threonine positions.

Materials and Methods

Protein Purification. The principal component of Amazon River dolphin myoglobin was isolated from frozen muscle tissue as previously described by Hapner et al. (1968). Phosphate buffer, pH 6.4, $\mu = 0.1$, was used to effect the purification of the crude homogenate on CM-50 Sephadex. The homogeneity of the purified myoglobin was then shown by polyacrylamide gel electrophoresis at pH 9.2 and 5.2. Apomyoglobin was prepared by the method of Teale (1959), as applied by Hapner et al. (1968).

Preparation of m-SPITC.[1] The sodium salt of this sulfonated isothiocyanate was prepared essentially according to the method of Dyson (1938). To 10 g (0.058 mol) of metanilic acid in 30 ml of water was added 29 ml (0.058 mol) of 2 *N* sodium hydroxide. The solution was then made acidic to litmus paper with glacial acetic acid and 7 g (0.059 mol) of thiophosgene was added with rapid stirring. The reaction was allowed to progress for 2 hr, after which the precipitate was filtered off and recrystallized from hot water. The yield was 20–25% of theoretical.

Peptide Nomenclature. Peptides obtained from the thermolysin digest of the apomyoglobin are designated by TH and numbered from their elution positions. For all other cleavage methods the peptides are numbered from the amino terminus to the carboxyl terminus of the completed sequence. The cyanogen bromide fragments are designated by the symbol CB, the peptides isolated following tryptic cleavage at arginine residues in the methyl-acetimidated protein are given the symbol MT, and the tryptic peptides from CB 2 are labeled CB2-T.

Cyanogen Bromide Cleavage. Apomyoglobin was cleaved with cyanogen bromide according to the method of Tsao et al. (1974). To 170 mg (10 μmol) of apomyoglobin in 15 ml of deoxygenated 70% formic acid was added 163 mg (150-fold excess/methionine) of cyanogen bromide. The solution was stirred in the dark at 4°C for 48 hr after which time the reaction mixture was diluted with water and lyophilized. The peptide mixture was then dissolved in 10% acetic acid and purified by gel filtration.

Preparation of Methyl-Acetimidated Myoglobin. Methyl acetimidate hydrochloride was prepared according to the method of Hunter and Ludwig (1962). To improve its stability, the reagent was stored under vacuum in a desiccator at −20°C. The methyl acetimidation of myoglobin (Garner and Gurd, 1975) was accomplished by a modification of the method of Reynolds (1968). To 250 mg (13 μmol) of fer-

† From the Department of Chemistry, Indiana University, Bloomington, Indiana 47401. *Received June 18, 1975.* This is the 72nd paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper see Oster et al. (1975). This work was supported by Public Health Service Research Grant HL-05556. F.E.D., R.A.B., and L.D.L. were supported by Public Health Service Grant TO1 GM 1046-14.

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

Table I: Amino Acid Composition of *Inia geoffrensis* Myoglobin.

Amino Acid	From Acid Hydrolysates[a]	From the Sequence
Asp	11.2	11
Thr	5.0	5
Ser	5.2	5
Glu	17.0	17
Pro	3.7	4
Gly	14.1	14
Ala	15.6	16
Val	4.9	5
Met	1.9	2
Ile	8.1	9
Leu	19.5	19
Tyr	1.8	2
Phe	6.7	7
Lys	19.9	20
His	11.9	12
Arg	2.8	3
Trp[b]	1.7	2

[a] Acid hydrolyses were performed on ferrimyoglobin for 24, 48, and 72 hr 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 hr). [b] Tryptophan was determined by the method of Liu and Chang (1971).

rimyoglobin dissolved in 10 ml of water and kept at 15°C was added a solution of 1.5 g (50-fold excess/amino group) of methyl acetimidate hydrochloride dissolved in 6.9 ml of 2 *N* sodium hydroxide. The pH of the mixture was then adjusted to 10.0 and the reaction allowed to proceed for 1 hr at which time the pH was lowered to 8.5 and the reaction was left overnight. The protein was separated from the reactants by gel filtration on Sephadex G-10 in 0.1 *N* Tris buffer (pH 10.0). The protein was deionized and the heme removed as with the ferrimyoglobin. The methyl-acetimidated apomyoglobin was lyophilized and stored at −20°C. The yield was approximately 180 mg of protein.

Cleavage with Trypsin at the Arginines of the Methyl-Acetimidated Apomyoglobin. To 90 mg (5 μmol) of methyl-acetimidated apomyoglobin dissolved in 5 ml of freshly deionized 5 *M* urea was added 1.8 mg (2%, w/w) of trypsin (TPCK Worthington). The pH was adjusted with base to 8.0 and the temperature maintained at 25°C. After 2 hr another 1.8 mg of trypsin was added and the reaction was allowed to progress for a total of 5 hr. The reaction was stopped by dropping the pH to 3.0 with acetic acid. The digest was then desalted in a Bio-Gel P-2 column and the peptides were purified by gel filtration.

Cleavage of CB 2 (56–131) with Trypsin. To a solution of 17 mg (2 μmol) of CB 2 (56–131) in 4 ml of water maintained at pH 8.1 and 25°C was added 0.34 mg (2%, w/w) of trypsin. After 3 hr another 0.34 mg of trypsin was added and at the end of 6 hr the pH was lowered to 6.0. The reaction mixture was centrifuged and the insoluble material washed three times with water. The supernatant phase and the water washes were combined and applied to a phosphocellulose column and eluted with a linear pyridine acetate gradient (Bradshaw et al., 1969).

Cleavage of Apomyoglobin with Thermolysin. Apomyoglobin (300 mg; 17 μmol) was digested with thermolysin as previously described by Vigna et al. (1974). The peptides were first purified on AG 50-X8 and, when necessary, repurified on AG 50-X4 using linear pyridine acetate gradients.

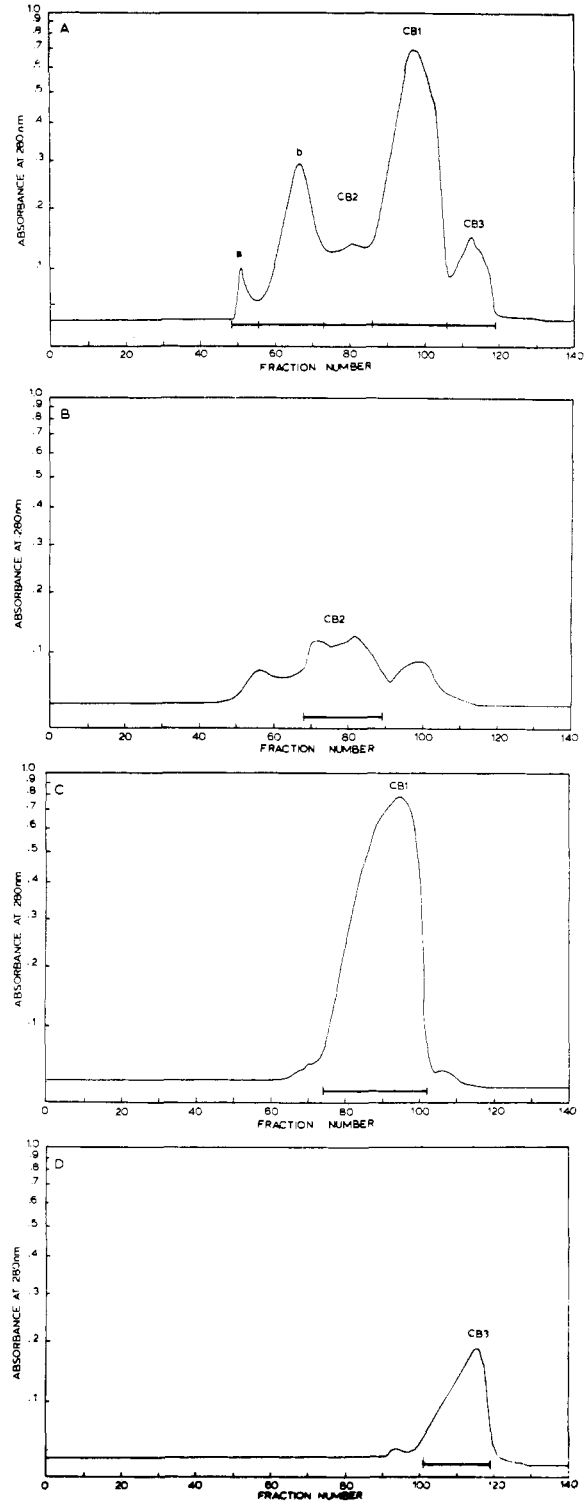

FIGURE 1: (A) Gel filtration pattern for the initial purification of the peptides obtained from the cyanogen bromide cleavage of Amazon River dolphin apomyoglobin. The peptide mixture was applied to a 2.6 cm × 195 cm column of Bio-Gel P-10 (200–400 mesh) and eluted with 10% acetic acid at a flow rate of 36 ml/hr and a fraction size of 6 ml. Peak a is uncleaved protein and peak b is a mixture of partially cleaved protein. (B) Peptide CB 2, (C) peptide CB 1, and (D) peptide CB 3 were all repurified using the same conditions as in the initial purification.

Sequencing Techniques. All the peptides used in this sequence determination were subjected to automated Edman degradations using a Beckman Model 890C sequencer. The fast peptide-DMAA program (071872, Beckman Instru-

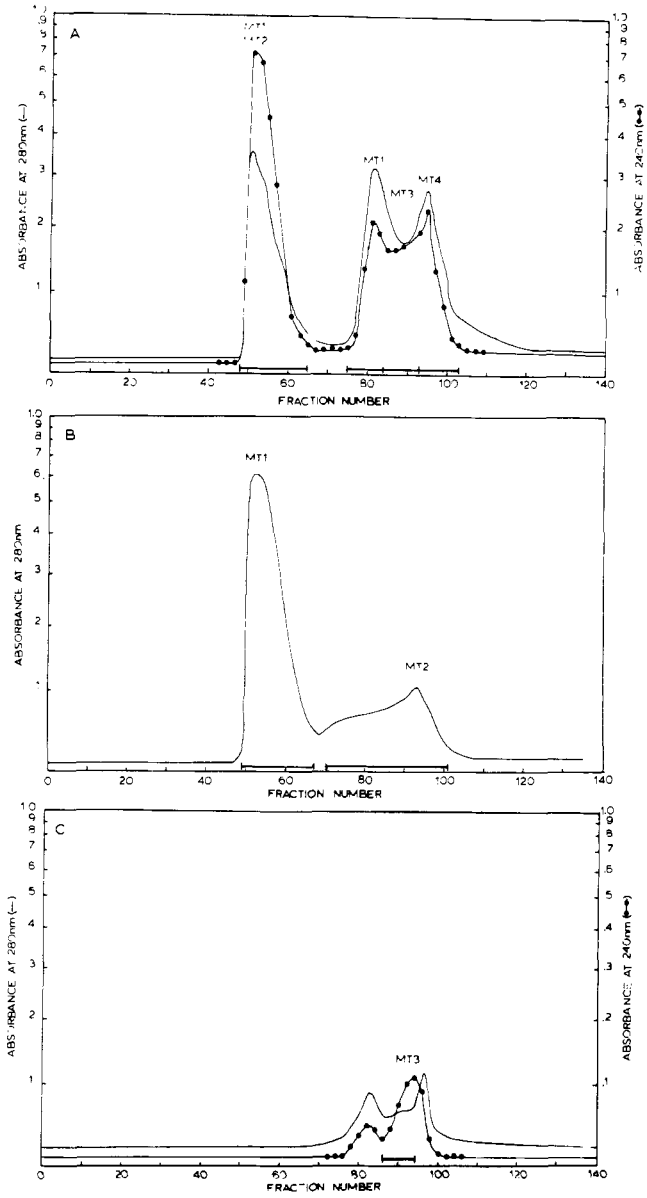

FIGURE 2: (A) Gel filtration pattern for the initial purification of the peptides obtained by tryptic digestion of the methyl-acetimidated apomyoglobin. The peptide mixture was applied to a 2.6 cm × 186 column of Bio-Gel P-6 (200–400 mesh) and eluted with 10% acetic acid at a flow rate of 36 ml/hr and a fraction size of 6 ml. (B) Peptides MT 1 and MT 2 were separated from each other by gel filtration on a 2.6 cm × 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/hr and a fraction size of 5 ml. (C) Gel filtration pattern for the repurification of peptide MT 3 on the original P-6 column keeping all of the original parameters constant.

ments) was used to sequence peptides of 25 residues or less and the fast protein-quadrol program (072172C) was used to sequence the NH_2-terminal region of the larger peptides.

All peptides which had free ϵ-aminolysine residues were first coupled with m-SPITC to decrease extraction losses. This was accomplished by first drying the peptide in the sequencer cup and then adding 1 mg of m-SPITC dissolved in 0.5 ml of DMAA buffer. The peptide was coupled for 1 hr and the buffer was removed by vacuum evaporation. When the peptide was dry, the desired program was started.

The amino acid phenylthiohydantoins were identified by gas chromatography on 2 mm × 4 ft columns of 10% SP400 on Chromosorb WHP in a Hewlett-Packard 5711A gas chromatograph as described by Pisano et al. (1972). Residues which were found to be acids or amides were also checked by thin-layer chromatography (Inagami and Murakami, 1972), and residues which gave no peaks on the gas chromatograph were reconverted to the free amino acid by 6 *N* HCl (Van Orden and Carpenter, 1964) and the residue was determined on the amino acid analyzer (Spackman et al., 1958).

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

Amino Acid	CB 1	CB 2	CB 3
Asp	5.1 (5)	4.2 (4)	2.0 (2)
Thr	2.1 (2)	2.9 (3)	
Ser	1.2 (1)	3.8 (4)	
Glu	8.0 (8)	7.3 (7)	2.0 (2)
Pro	1.0 (1)	2.9 (3)	
Gly	6.0 (6)	6.3 (6)	1.9 (2)
Ala	3.1 (3)	10.0 (10)	3.0 (3)
Val	2.9 (3)	1.9 (2)	
Ile	1.9 (2)	5.4 (6)	1.0 (1)
Leu	7.7 (8)	8.0 (8)	3.0 (3)
Tyr		0.9 (1)	0.9 (1)
Phe	2.9 (3)	2.0 (2)	2.0 (2)
Lys	6.0 (6)	10.2 (10)	3.9 (4)
His	3.0 (3)	8.0 (8)	1.0 (1)
Arg	1.0 (1)	1.1 (1)	1.1 (1)
Trp[b]	1.8 (2)		
Hse	0.4 (1)	0.4 (1)	
Total residues	55	76	22
Yield (%)	64	53	74
Position	1–55	56–131	132–153

[a,b] See footnotes in Table I.

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

Amino Acid	MT 1	MT 2	MT 3	MT 4
Asp	4.0 (4)	2.9 (3)	3.0 (3)	1.1 (1)
Thr		4.5 (5)		
Ser	1.0 (1)	4.0 (4)		
Glu	3.9 (4)	9.5 (10)	2.2 (2)	1.2 (1)
Pro		2.4 (3)	0.8 (1)	
Gly	4.6 (5)	5.4 (5)	2.3 (2)	1.9 (2)
Ala	2.4 (2)	7.1 (7)	5.1 (5)	2.2 (2)
Val	2.7 (3)	1.8 (2)		
Met		0.7 (1)	0.6 (1)	
Ile	1.6 (2)	5.1 (6)[b]		0.9 (1)
Leu	4.6 (5)	10.2 (11)	2.1 (2)	1.3 (1)
Tyr		0.9 (1)		0.8 (1)
Phe		3.6 (4)	1.8 (2)	1.0 (1)
Lys	1.2 (1)	15.5 (15)	1.1 (1)	2.7 (3)
His	1.3 (1)	9.3 (9)	1.1 (1)	0.9 (1)
Arg	1.1 (1)	1.0 (1)	0.9 (1)	
Trp	1.8 (2)			
Total Residues	31.	87	21	14
Yield (%)	61	78	88	91
Position	1–31	32–118	119–139	140–153

[a] Amino acid compositions were prepared on hydrolysates at 110°C for 24 hr 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 mole 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 hr of hydrolysis.

Results

Amino Acid Composition. The amino acid composition of the principal component of Amazon River dolphin myo-

Table IV: Amino Acid Composition[a] of Tryptic Peptides from CB 2.

Amino Acid	CB2-T1	CB2-T1A	CB2-T1B	CB2-T2	CB2-T3	CB2-T4A	CB2-4	CB2-T5	CB2-T6	CB2-T7
Asp	1.1 (1)	1.0 (1)		1.0 (1)						2.0 (2)
Thr				1.8 (2)		1.1 (1)	0.9 (1)			
Ser	1.0 (1)	0.6 (1)				1.1 (1)	1.2 (1)		1.6 (2)	
Glu	1.2 (1)	1.0 (1)				3.1 (3)	3.2 (3)		2.2 (2)	1.2 (1)
Pro						0.8 (1)	1.0 (1)	0.8 (1)		1.3 (1)
Gly				3.2 (3)		1.2 (1)	1.1 (1)			2.0 (2)
Ala	1.0 (1)	1.0 (1)		1.1 (1)		2.8 (3)	3.1 (3)		1.2 (1)	3.4 (3)
Val				0.9 (1)					1.2 (1)	
Ile				0.9 (1)				2.2 (2)	2.1 (3)[b]	
Leu	1.1 (1)	0.9 (1)		3.1 (3)		2.1 (2)	2.0 (2)		2.1 (2)	
Tyr									0.9 (1)	
Phe									1.0 (1)	1.1 (1)
Lys	2.7 (3)	1.7 (2)	1.0 (1)	1.1 (1)	2.0 (2)	3.0 (3)	1.8 (2)	1.9 (2)		
His				1.0 (1)		2.7 (3)	2.7 (3)	1.0 (1)	1.9 (2)	
Arg									0.9 (1)	
Hse										0.6 (1)
Total Residues	8	7	1	14	2	18	17	6	16	11
Yield (%)	16	67	88	70	51	26	42	45	41	72
Position	56–63	56–62	63 or 78	64–77	78–79	79–96	80–96	97–102	103–118	119–139
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.

globin was obtained from 24-, 48-, and 72-hr hydrolysates of the ferrimyoglobin. The results are summarized in Table I.

Separation of Cyanogen Bromide Cleaved Peptides. The peptides obtained from the CNBr digest were initially purified on a column of Bio-Gel P-10 as can be seen in Figure 1. The separation obtained was comparable with that obtained by Marshall et al. (1974). The peptides were repurified by gel filtration on the same P-10 column. The amino acid compositions of these peptides are shown in Table II.

Cleavage at Arginine Residues. Methyl-acetimidated apomyoglobin was cleaved by trypsin at its arginine residues and initially purified by gel filtration on a column of Bio-Gel P-6 as seen in Figure 2. Peptides MT 1 and 2 were separated on a column of P-10 and peptide MT 3 was repurified by gel filtration on the original P-6 column. The amino acid compositions of these peptides are shown in Table III.

Tryptic Peptides of CB2. Partial purification of the soluble peptides in a tryptic digest of CB 2 was achieved by ion-exchange chromatography on Cellex P using a linear gradient of pyridine acetate (pH 2.5 → 5.0).[2] All of the peptides except CB2-T4 and CB2-T5 were obtained pure. These two peptides were separated by gel filtration on Bio-Gel P-4. The position and composition of these peptides are reported in Table IV.

Thermolysin Peptides. After the apomyoglobin was digested with thermolysin the peptides were isolated on AG 50-X8 using a linear pyridine acetate gradient as can be seen in Figure 3. After analysis all pools which contained serine, threonine, tyrosine, or arginine were repurified on AG 1-X8 or AG 50-X4. The amino acid compositions for these peptides are summarized in Table V.

Sequence Investigations. Only the sequence data necessary to establish the entire sequence are reported here.

Sequenator Results. The complete amino acid sequence of Amazon River dolphin myoglobin is shown in Figure 4. In all the sequencer runs the yields for the phenylthiohydantoins quantitated by gas chromatography[2] were near the values expected except for serine and threonine and the acids and amides. The values for serine and threonine can be low because of dehydration and polymerization and the values for the amides can be low because of absorption of these polar molecules onto the liquid phase of the gas chromatography columns. The reason for the low values of aspartic and glutamic acid may be incomplete silylation.

In all the sequencer runs the results were as expected except in the sequencer run on peptide CB 2. In this peptide at positions 32 and 33 (positions 87 and 88 in the completed sequence) there is a Lys-Pro bond. When these residues were reached there was an appreciable drop in the repetitive yield along with an increase in the carryover from previous cycles. No explanation can be offered at present. However, these results are very reproducible, having been seen in the same peptide from other myoglobins.

Sequenator analysis A (Figure 5) yielded the first 36 amino terminal residues of the intact protein to give five cycles past the first arginine residue. The amino terminal residues of sequenator analysis B on peptide MT 2 overlapped the intact protein sequence at these five residues and extended it 24 residues to position 60. In a similar manner the sequenator analysis C of peptide CB 2 overlapped the sequenator analysis B by 5 residues and extended it 38 cycles to residue 98 in the completed sequence.

The remaining methyl-acetimidated peptide (MT3) was sequenced 17 rounds in sequenator analysis D to overlap the carbonyl terminal of CB 2 and the amino terminal of CB 3. Finally in sequenator analysis E peptide CB 3 was completely sequenced to complete the carboxyl terminal of the protein. The sequence of residues 99–118 was completed by conventional techniques on tryptic and thermolytic peptides.

Discussion

The present report deals with the first completed Cetacean myoglobin sequence determined largely by automated Edman degradation. The strategy differs from that of the

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

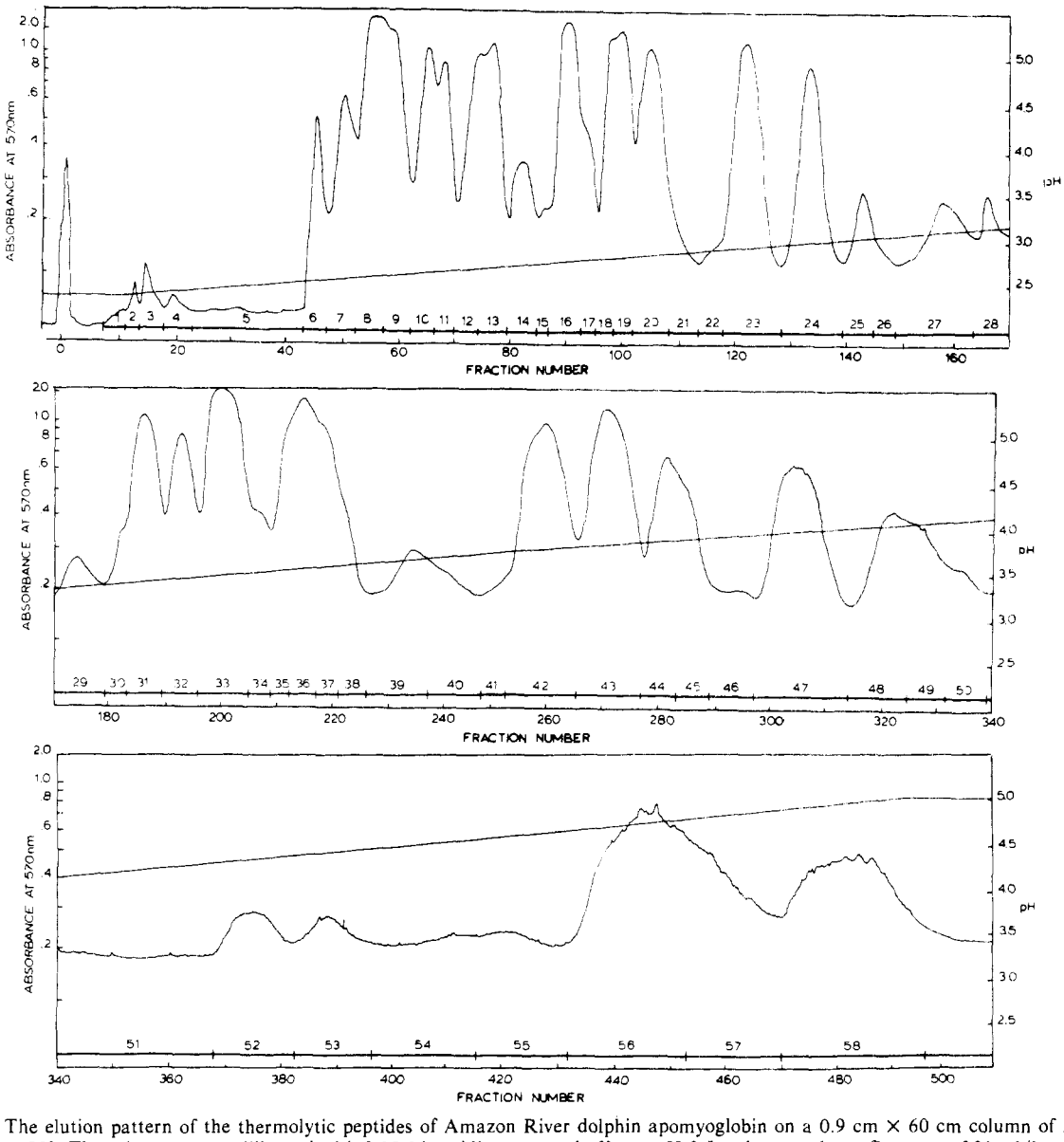

FIGURE 3: The elution pattern of the thermolytic peptides of Amazon River dolphin apomyoglobin on a 0.9 cm × 60 cm column of AG 50-X8 maintained at 55°. The column was equilibrated with 0.05 *M* pyridine acetate buffer at pH 2.5 and pumped at a flow rate of 30 ml/hr with a fraction size of 3 ml. The column was then developed with a 48-hr linear gradient of 0.05 *M* pyridine acetate (pH 2.5) to 2.0 *M* pyridine acetate (pH 5.0). The column was monitored by automatic alkaline hydrolysis and ninhydrin analysis.

classical methods applied to Cetacean myoglobins (Edmundson, 1965; Bradshaw and Gurd, 1969; Kluh and Bakardjieva, 1971) in which several different cleavage patterns were used to isolate large numbers of overlapping peptides. Here the pattern has naturally emphasized the isolation of large peptides when possible. The cleavages at the two methionine residues and the three arginine residues have allowed the automated sequencing of 85% of the protein with five sequencer runs: whole protein, MT 2, CB 2, MT 3, and CB 3. The rest of the protein sequence was determined by more classical techniques.

Methyl acetimidation has been used previously by Henderson et al. (1973) to assist in protein sequencing. Unlike the results obtained by Reynolds (1968) with ribonuclease, it was found in the present work that the methyl-acetimidated apomyoglobin had quite different solubility properties from the unmodified apoprotein. Apomyoglobin is soluble at all pH values whereas the modified apoprotein is almost completely insoluble in the range above pH 6.0. It could be retained in solution with a minimum of 5 *M* urea, a concentration compatible with trypsin activity (Delaage and Lazdunski, 1968). Complete modification with methyl acetimidate occurs readily with ferrimyoglobin (Garner and Gurd, 1975), whereas all attempts at treating the apomyoglobin directly in the present work resulted in a product which was only partially modified and insoluble.

The isolated peptides were not treated to remove the acetimidyl groups (Ludwig and Byrne, 1962) because it was found that the repetitive yields for these peptides were approximately 2% better than for peptides of similar size treated with m-SPITC. In addition, reconversion of the phenylthiohydantoin lysine derivatives yielded free lysine rather than the ϵ-methylacetimidyl derivative.

The results reported here agree in general with those found by Singhal and Atassi (1971) for sperm whale myoglobin modified with citraconic anhydride. There are only two differences. The first is that in the present case peptide MT 1 (1–31) appears in two chromatographic pools (Figure 2), representing an aggregated and an unaggregated form. The second is the absence of a cleavage within the last frag-

Table V: Amino Acid Composition[a] of Thermolysin Peptides.

Amino Acid	8-6	9-1	9-3	10	23-1	24	30	45-1	42-2	42-1	47-2	50	54-1	56-2	56-3
Asp	1.1 (1)		0.9 (1)				1.0 (1)		2.2 (2)				1.0 (1)		1.1 (1)
Thr		1.0 (1)				0.9 (1)	1.1 (1)			0.9 (1)				1.1 (1)	
Ser	0.9 (1)		0.7 (1)	0.9 (1)	0.7 (1)				0.7 (1)						0.9 (1)
Glu	2.0 (2)		1.1 (1)	1.1 (1)	1.1 (1)	1.7 (2)		1.1 (1)	1.1 (1)	1.1 (1)					
Pro									1.2 (1)	1.0 (1)				0.9 (1)	1.0 (1)
Gly	2.0 (2)						1.2 (1)		2.2 (2)	1.0 (1)				1.1 (1)	0.9 (1)
Ala		0.8 (1)	1.0 (1)	1.0 (1)	1.0 (1)	0.9 (1)			2.2 (2)						
Val		1.1 (1)													
Met															
Ile				0.9 (1)							1.0 (1)	1.0 (1)		1.2 (1)	
Leu	1.1 (1)	1.0 (1)			1.0 (1)	1.0 (1)			1.0 (1)			1.1 (1)			0.9 (1)
Tyr								0.9 (1)			1.0 (1)				
Phe									1.0 (1)	1.0 (1)			1.0 (1)		
Lys						0.9 (1)		1.0 (1)		1.2 (1)	1.0 (1)		1.0 (1)	1.8 (2)	
His					1.0 (1)		1.1 (1)		1.8 (2)	1.0 (1)				1.0 (1)	2.1 (2)
Arg									1.0 (1)			1.0 (1)	1.0 (1)		1.0 (1)
Trp	+(1)														
Total	8	4	4	4	5	6	4	3	14	7	3	3	4	7	8
Residues															
Yield (%)	31	26	32	37	20	40	6	57	26	56	36	20	23	34	23
Position	1–8	68–71	57–60	107–110	89–93	49–54	64–67	146–148	115–128	33–39	101–103	30–32	138–141	94–100	115–122

[a] See footnote *a* in Table III.

```
                5                  10                  15
 1  Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Ile Trp Gly
16  Lys Val Glu Ala Asp Leu Ala Gly His Gly Gln Asp Val 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 Glu 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 His Gly
```

FIGURE 4: The amino aicd sequence of Amazon River dolphin myoglobin. The hyphens between the amino acid residues have been omitted for clarity.

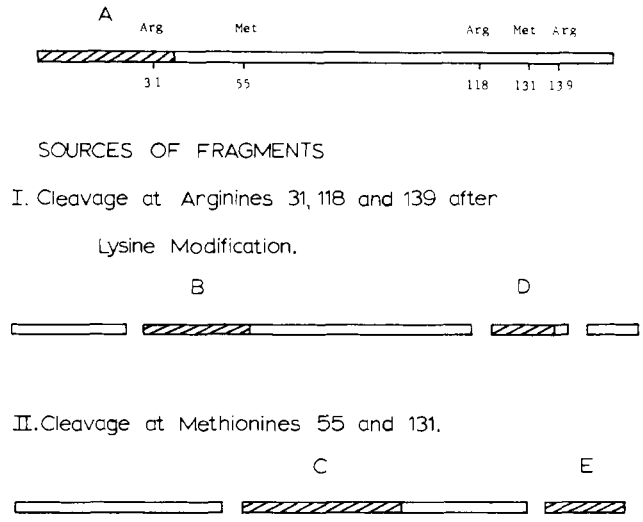

FIGURE 5: Diagrammatic summary of fragments generated from Amazon River dolphin myoglobin for sequenator analysis. The top bar represents the whole myoglobin and the residues that are important for its fragmentation. The capital letters A–E identify the sequenator analyses in the order in which they are described in the text. A hatched section in each horizontal bar indicates the segment of sequence determined by that analysis.

ment at lysine-147 (Singhal and Atassi, 1971) due to incomplete protection by citraconylation.

A useful observation that has come out of this work is that the m-SPITC promoted higher repetitive yields than the well-established para isomer (Birr et al., 1970). The m-SPITC was introduced as an alternative Edman reagent capable of more facile cyclization.[3]

The sequence of the Amazon River dolphin myoglobin is compared in Figure 6 with those of the sperm whale, the common porpoise, and the common dolphin. As emphasized in the difference matrix shown in Figure 7, the Amazon River dolphin myoglobin is similar in sequence to the other dolphins represented in Figure 6, and in addition the bottlenose dolphin, *Tursiops truncatus*.[4] As pointed out previously for the common or harbor porpoise myoglobin (Bradshaw and Gurd, 1969) the substitutions appear closely compatible with the crystalline three-dimensional structure of the sperm whale protein (Watson, 1969). It is instructive to ex-

[3] F. E. Dwulet, J. S. Scarlett III, and W. H. Garner, in preparation.

[4] R. A. Vigna and B. N. Jones, work in progress.

Species	Residue Number 1	4	12	13	15	21	26	28	35	45	54
A. R. Dolphin	Gly	Asp	Asn	Ile	Gly	Leu	Gln	Val	Gly	Lys	Glu
Sperm Whale	Val	Glu	His	Val	Ala	Val	Gln	Ile	Ser	Arg	Glu
Common Porpoise	Gly	Glu	Asn	Val	Gly	Leu	Gln	Val	Gly	Lys	Glu
B. S. Dolphin	Gly	Asp	Asn	Val	Gly	Val	Glu	Ile	Gly	Lys	Asp

Species	Residue Number 66	74	83	85	121	122	129	144	151	152
A. R. Dolphin	Asn	Gly	Glu	Glu	Gly	Asp	Ala	Ala	Phe	His
Sperm Whale	Val	Ala	Glu	Glu	Gly	Asn	Gly	Ala	Tyr	Gln
Common Porpoise	Asn	Gly	Glu	Asn	Ala	Glu	Gly	Thr	Phe	His
B. S. Dolphin	Asp	Ala	Asp	Glu	Ala	Gln	Gly	Ala	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.

	SPERM WHALE	BLACK SEA DOLPHIN	COMMON PORPOISE	AMAZON RIVER DOLPHIN	
	15	5	6	7	BOTTLENOSED DOLPHIN
		14	15	15	SPERM WHALE
			11	11	BLACK SEA DOLPHIN
				7	COMMON PORPOISE

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

amine the seven differences between the myoglobins of the Amazon River dolphin and the common porpoise from the North Atlantic Ocean. These will be referred to with the residue in the Amazon River dolphin myoglobin given first followed by the homologous common porpoise residue in parentheses.

4 Aspartic Acid (Glutamic Acid). Aspartic acid has been found in other myoglobins such as that of California sea lion (Vigna et al., 1974), and the substitution may be considered a conservative one.

13 Isoleucine (Valine). This is the first report of isoleucine at position 13 in a porpoise or dolphin myoglobin, but it is common in the Balaenoptera whales (Edman and Begg, 1967).[5]

85 Glutamic Acid (Asparagine). This is a common residue at the position (Figure 6), and the substitution constitutes the only charge change between the two proteins under comparison.

121 Glycine (Alanine), 122 Aspartic Acid (Glutamic Acid), 129 Alanine (Glycine), 144 Alanine (Threonine). All of these changes are conservative and are seen in some other myoglobin whose sequence has already been reported. For example, the harbor seal myoglobin shows identity with the Amazon River dolphin myoglobin at positions 122 and 129 (Bradshaw and Gurd, 1969), and other identities are given in Figure 6. It should be noted that all these comparisons are made in terms of the main component of myoglobin from skeletal muscle of the given species (Garner et al., 1974; Parkhurst and LaGow, 1975). The yield of minor myoglobin components from the Amazon River dolphin appeared relatively sparse.

[5] R. A. Bogardt, L. D. Lehman, and T. M. Rothgeb, work in progress.

Acknowledgments

The authors are grateful to Mrs. K. Verhey, Mrs. L. Gurd, Miss C. Bowman, and Miss K. Summers for their excellent technical help. Mr. K. Sexton helped with the procurement of the animal material after the dolphin died at the Sea World, San Diego. We are grateful to Drs. William H. Garner and Edward Mitchell for advice.

Supplementary Material Available

Tables and figures containing additional data as noted in the text will appear following these pages in the microfilm edition of this volume of the journal. Photocopies of the supplementary material from this paper only or microfiche (105 × 148 mm, 24× reduction, negatives) containing all of the supplementary material for the papers in this issue may be obtained from the Business Office, Books and Journals Division, American Chemical Society, 1155 16th St., N.W., Washington, D.C. 20036. Remit check or money order for $4.50 for photocopy or $2.50 for microfiche, referring to code number BIO-75-5336.

References

Birr, C., Reitel, C., and Wieland, T. (1970), *Angew. Chem. Int. Ed. Engl. 9,* 731.

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.

Delaage, M., and Lazdunski, M. (1968), *Eur. J. Biochem. 4,* 378.

Dyson, G. M. (1938), *Recl. Trav. Chim. Pays-Bas 58,* 1026.

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

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

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.

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

Henderson, L. E., Henriksson, D., and Nyman, P. O. (1973), *Biochem. Biophys. Res. Commun. 52,* 1388.

Hunter, M. J., and Ludwig, M. L. (1962), *J. Am. Chem. Soc. 84,* 3491.

Inagami, T., and Murakami, K. (1972), *Anal. Biochem. 47,* 501.

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

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

Ludwig, M. L., and Byrne, R. (1962), *J. Am. Chem. Soc. 84,* 4160.

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

Oster, O., Neireiter, G. W., Clouse, A. O., and Gurd, F. R. N. (1975), *J. Biol. Chem.* (in press).

Parkhurst, L. J., and LaGow, J. (1975), *Biochemistry 14,* 1200.

Pisano, J. J., Bronzert, T. J., and Brewer, H. B. (1972), *Anal. Biochem. 45,* 43.

Reynolds, J. H. (1968), *Biochemistry 7,* 3131.

Singhal, R. P., and Atassi, M. Z. (1971), *Biochemistry 10,* 1756.
Spackman, D. H., Moore, S., and Stein, W. H. (1958), *Anal. Chem. 30,* 1190.
Teale, F. W. J. (1959), *Biochim. Biophys. Acta 35,* 543.
Tsao, D., Morros, D. H., Azari, P., Tengerdy, R. P., and Phillips, J. L. (1974), *Biochemistry 13,* 403.
Van Orden, H. O., and Carpenter, F. H. (1964), *Biochem. Biophys. Res. Commun. 14,* 399.
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.

Isolation of Soluble Elastin from Lathyritic Chicks. Comparison to Tropoelastin from Copper Deficient Pigs†

Judith Ann Foster,*,‡ Robert Shapiro, Paul Voynow, George Crombie, Barbara Faris, and Carl Franzblau

ABSTRACT: Tropoelastin was isolated from the aortas of chicks rendered lathyritic by treatment with β-aminopropionitrile. The soluble elastin was judged homogeneous by sodium dodecyl sulfate polyacrylamide gel electrophoresis and possessed an estimated molecular weight of 70000. Automated sequential analysis revealed that the N-terminal region of the chick tropoelastin is very homologous to tropoelastin isolated from copper-deficient piglets. N-terminal analysis of a trypsin digest of chick tropoelastin showed that tyrosine frequently is found adjacent to lysine residues. This positioning of tyrosine residues may be significant in terms of a possible regulatory role in elastin cross-link formation.

Significant progress has been made in the study of elastin structure with the recent isolation of a soluble "precursor" of elastin, referred to as tropoelastin. The isolation of this soluble elastin from any source requires inhibition of lysine-derived cross-linkages (Franzblau and Lent, 1969) which are thought to contribute to the extremely insoluble character of mature elastin (Partridge, 1962). The first step in the biosynthesis of elastin cross-links is the oxidative deamination of lysyl residues by the enzyme lysyl oxidase (Pinnell and Martin, 1968). Inhibition of the enzyme by either copper deficiency (Weissman et al., 1963) or β-aminopropionitrile (BAPN)[1] treatment (Sykes and Partridge, 1972) has resulted in the isolation of a non-cross-linked, soluble elastin which is readily amenable to structural studies. The most extensive information currently available on tropoelastin has come from material isolated from copper-deficient piglets. This tropoelastin has a molecular weight of approximately 72000 (Sandberg et al., 1969) and possesses the same amino acid composition as insoluble porcine elastin, with the only difference being that the precursor elastin contains few, if any, cross-linkages but does possess a high content of lysine residues relative to the insoluble elastin. Foster et al. (1973) have published extensive primary sequence data on tryptic peptides derived from porcine tropoelastin which indicate an elastin primary structure distinct from that of collagen and possessing repeating units of tetra-, penta-, and hexapeptides.

Recently, Sykes and Partridge (1974) have reported on the isolation of soluble elastin from the aortas of BAPN-treated chicks. This tropoelastin has a molecular weight of 57000, which differs significantly from that isolated from the porcine source.

Since the isolations of these tropoelastin samples differed in the method of enzyme inhibition, the purification procedure, and animal source, the discrepancy in molecular weights could be attributable to any of these differences. An alternative explanation for the molecular weight difference is suggested by the finding of several laboratories that proteolysis of soluble elastin occurs during isolation procedures. Consequently, various enzyme inhibitors have been employed to protect the tropoelastin from proteolytic attack (Sandberg et al., 1975; Narayanan and Page, 1974).

The present communication reports on the isolation and characterization of tropoelastin from the aortas of BAPN-treated chicks and compares this elastin to that derived from a copper-deficient, porcine source.

Materials and Methods

Preparation of Tropoelastin. The procedure followed was a modification of that described by Sykes and Partridge (1974). The method actually is based on one first reported by Smith et al. (1972). The major modification introduced was the addition of protease inhibitors during all but the final steps of the procedure in order to prevent tropoelastin degradation. One thousand chicks were raised on a diet of commercial starting feed supplemented with 0.1% (w/w) β-aminopropionitrile fumarate (BAPN) beginning on the day of hatching. During the feeding period, there was a mortality of approximately 10%. After 6 days the chicks were sacrificed and their aortas removed and immediately placed in a wash solution containing 20 m*M* disodium ethylenediaminetetraacetate (EDTA), 5 m*M* BAPN, and 5 m*M*

† From the Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts 02118. *Received June 18, 1975.* This investigation was supported by National Institutes of Health Grants HL 15964, A01-HR-42934, and HL 17415.

‡ Established Investigator of the American Heart Association. To whom correspondence should be addressed.

[1] Abbreviations used are: BAPN, β-aminopropionitrile; EDTA, disodium ethylenediaminetetraacetate; MalNEt, *N*-ethylmaleimide; Pth, phenylthiohydantoin derivatives; TPCK, L-tosylamido-2-phenylethyl chloromethyl ketone.