p 345.
Lanchantin, G. F., Friedmann, J. A., and Hart, D. W. (1968), *J. Biol. Chem. 243,* 476.
Lee, J. C., and Timasheff, S. N. (1974), *Arch. Biochem. Biophys. 165,* 268.
Magnusson, S., Petersen, T. E., Sottrup-Jensen, L., and Claeys, H. (1975), in Proteases and Biological Control, Vol. 2, Reich, E., Rifkin, D. B., and Shaw, E., Ed., Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, p 123.
Magnusson, S., Sottrup-Jensen, L., Petersen, T. E., Morris, H. R., and Dell, A. (1974), *FEBS Lett. 44,* 189.
Mattock, P., and Esnouf, M. P. (1973), *Nature (London), New Biol. 242,* 90.
Moore, S., and Stein, W. H. (1963), *Methods Enzymol. 6,* 819.
Nelsestuen, G. L., Zytokoviez, T. H., and Howard, J. B. (1974), *J. Biol. Chem. 249,* 6347.
Nemerson, Y., and Clyne, L. P. (1974), *J. Lab. Clin. Med. 83,* 301.
Østerud, B., and Flengsrud, R. (1975), *Biochem. J.* 145, 469.
Pirkle, H., McIntosh, M., Theodor, I., and Vernon, S. (1973), *Thromb. Res. 2,* 461.
Proctor, R. R., and Rapaport, S. I. (1961), *Am. J. Clin. Pathol. 36,* 212.
Rosenberg, J. S., Beeler, D. L., and Rosenberg, R. D. (1975b), *J. Biol. Chem. 250,* 1607.
Rosenberg, J. S., McKenna, P. W., and Rosenberg, R. D. (1975a), *J. Biol. Chem. 250,* 8883.
Schmer, G. (1972), *Hoppe-Seyler's Z. Physiol. Chem. 353,* 810.
Segrest, J. P., and Jackson, R. L. (1972), *Methods Enzymol. 28,* 54.
Shapiro, S. S., and Waugh, D. F. (1966), *Thromb. Diath. Haemorrh. 16,* 469.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30,* 1190.
Stenflo, J. (1976), *J. Biol. Chem. 251,* 355.
Stenflo, J., Fernlund, P., Egan, W., and Roepstorff, P. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 2730.
Teller, D. C., Horbett, J. A., Richards, E. G., and Schachman, H. K. (1969), *Ann. N.Y. Acad. Sci. 164,* 66.
Thompson, A. R., Ericsson, L. H., and Enfield, D. L. (1974), *Circulation, Suppl. III, 49–50,* 292.
Titani, K., Fujikawa, K., Enfield, D. L., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 3082.
Walz, D. A., and Seegers, W. H. (1974), *Biochem. Biophys. Res. Commun. 60,* 717.
Warren, L. (1959), *J. Biol. Chem. 234,* 1971.
Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406.
Wright, I. (1959), *J. Am. Med. Assoc. 170,* 325.
Yphantis, D. A. (1964), *Biochemistry 3,* 297.

The Complete Amino Acid Sequence of the Major Component Myoglobin from the Arctic Minke Whale, *Balaenoptera acutorostrata*†

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

ABSTRACT: The complete primary structure of the major component myoglobin from the Arctic minke whale, *Balaenoptera acutorostrata,* was determined by specific cleavage of the protein to obtain large peptides which are readily degraded by the automatic sequencer. Over 80% of the amino acid sequence was established from the three peptides resulting from the cleavage of the apomyoglobin at the two methionine residues with cyanogen bromide along with the four peptides resulting from the cleavage of the methyl acetimidated apomyoglobin at the three arginine residues with trypsin. The further digestion of the central cyanogen bromide peptide with trypsin and *S. aureus* strain V8 protease enabled the determination of the remainder of the covalent structure. This myoglobin differs from that of the dwarf sperm whale, *Kogia simus,* at 16 positions, and the common dolphin, *Delphinus delphis,* at 14 positions, from that of the common porpoise, *Phocaena phocaena,* and the bottlenosed dolphin, *Tursiops truncatus,* at 13 positions, from that of the Amazon River dolphin, *Inia geoffrensis,* at 10 positions, and from that of California gray whale, *Eschrichtius gibbosus,* at 3 positions. 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 sequence of the myoglobin from the Amazon River dolphin (Dwulet et al., 1975), California gray whale (Bogardt et al., 1976), and the Atlantic bottlenosed dolphin (Jones et al., 1976) were reported. These sequences of Cetacean myoglobins were determined by automatic Edman degradation. This paper reports the application of the peptide fragmentation and analytical procedures that were used in these papers in determining the complete amino acid sequence of the major component myoglobin from the Arctic minke whale. Completion of this sequence extends the number of known Cetacean myoglobin sequences to seven, including the above mentioned Amazon River dolphin, California gray whale, and Atlantic bottlenosed dolphin, Black Sea

† From the Department of Chemistry, Indiana University, Bloomington, Indiana 47401. *Received September 21, 1976.* This is the 78th paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper, see Jones et al. (1977). This work was supported by United States Public Health Service Research Grant HL-05556. L.D.L., F.E.D., and R.A.B. were supported by United States Public Health Service Grant T01 GM 1046-14.

‡ Present address: Scripps Clinic and Research Foundation, La Jolla, California 92037.

TABLE I: Amino Acid Composition of *Balaenoptera acutorostrata* Myoglobin.

Amino Acid	No. of Residues from Acid Hydrolysates[a]	No. of Residues from the Sequence
Asp	11.0	11
Thr	4.9	5
Ser	5.1	5
Glu	17.1	17
Pro	4.1	4
Gly	10.1	10
Ala	19.0	19
Val	6.0	6
Met	2.0	2
Ile	9.8	10
Leu	18.1	18
Tyr	2.0	2
Phe	6.8	7
Lys	20.0	20
His	11.9	12
Arg	3.1	3
Trp[b]	1.9	2

[a] Ferrimyoglobin samples were hydrolyzed for 24, 48, and 72 h and duplicate analyses were performed on each hydrolysate. The values obtained for each residue were averaged, except for serine, threonine, valine, isoleucine, and leucine. 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).

dolphin (Kluh and Bakardjieva, 1971), common porpoise (Bradshaw and Gurd, 1969), and sperm whale (Edmundson, 1965).

Experimental Section

Materials

The principal component of Arctic minke whale myoglobin was isolated from muscle tissue as described by Hapner et al. (1968). Phosphate buffer, pH 6.4, 0.1 ionic strength, was used to effect the purification of the crude homogenate on CM-50 Sephadex. The homogeneity of the purified myoglobin was 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).

Methyl acetimidate hydrochloride was prepared according to the method of Hunter and Ludwig (1962). Preparation of 3-SPITC[1] was by the procedure of Dwulet and Gurd (1976). Tos-PhCH$_2$Cl treated trypsin was purchased from Worthington. Staphylococcal protease was obtained from Miles Laboratories Ltd. Carboxypeptidase C was purchased from Rohm and Haas, Darmstadt, Germany. Sequencer reagents of "Sequencer" grade were obtained from Beckman Instruments. All other chemicals were reagent grade.

Methods

Peptide Nomenclature. For all cleavage methods the resulting peptides are numbered from the amino terminal to the carboxyl terminal of the complete sequence. The cyanogen bromide fragments are designated by the symbol CB, peptides resulting from the cleavage with trypsin at the arginine residues in the methyl acetimidated protein are given the symbol MT, and the tryptic and staphylococcal protease peptides obtained by the digestion of the middle cyanogen bromide peptide, CB2, are labeled TCB2 and PCB2, respectively.

Specific Enzymatic and Chemical Cleavage. The procedures used in the cyanogen bromide cleavage, tryptic digestion of the methyl acetimidated apomyoglobin, tryptic digestion of CB2 (56–131), and staphylococcal protease digestion of CB2 (56–131) were as described previously (Dwulet et al., 1975; Bogardt et al., 1976; Jones et al., 1976).

Amino Acid Analysis. Acid hydrolysis was performed with constant boiling HCl at 110 °C unless otherwise specified. The amino acids were analyzed by the method of Spackman et al. (1958) on a Model 121 Beckman amino acid analyzer interfaced with a Texas Instruments 980A minicomputer which performed the identification and integration of the amino acid chromatograms (Bogardt, 1977). Tryptophan was determined by the method of Liu and Chang (1971).

Automated Edman Degradations. Automated Edman degradations were performed on a Beckman 980C sequencer. The sequencing techniques and the methods for the identification of the phenylthiohydantoin amino acids used in this sequence determination are identical with those previously reported by Dwulet et al. (1975).

Carboxypeptidase C Time Course Hydrolysis. Time course hydrolysis with carboxypeptidase C was performed as previously described (Bogardt et al., 1976).

Results

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

Peptide Separation. The peptides resulting from the cyanogen bromide cleavage, tryptic digestion of the methyl acetimidated apomyoglobin, tryptic digestion of CB2 (56–131), and staphylococcal protease digestion of CB2 (56–131) were purified as described previously (Dwulet et al., 1975; Bogardt et al., 1976; Jones et al., 1976). The amino acid analyses of all peptides obtained were found to be in good agreement with the expected values and the results can be found in the supplementary material.[2]

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

Sequencer Results. The complete primary structure of Arctic minke whale myoglobin is shown in Figure 1. The sequence strategy used to obtain the primary structure is outlined diagrammatically in Figure 2. Sequencer analysis A (Figure 2) represents the first 34 amino terminal residues obtained by the automatic Edman degradation of the intact apomyoglobin. Peptide MT2 (32–118) in sequencer analysis B provided a 3-residue overlap with the intact protein analysis and extended the sequence 27 residues to position 61. Sequencer analysis C of peptide CB2 (56–131) provided an overlap of sequencer analysis B of 6 residues and extended the sequence 28 residues to position 89. Analysis D yielded the entire sequence of peptide PCB2-4 (86–105), which overlapped analysis C by 4 residues and extended the sequence 16 residues to position 105. Peptide TCB2-6 (103–118) in sequencer analysis E overlapped analysis D with the only tyrosine in CB2 (52–131) and extended the sequence to position 118, which is the only arginine in CB2 (56–131). Analysis F of PCB2-6 (110–122) extended the se-

[1] Abbreviations used are: 3-SPITC, 3-sulfophenyl isothiocyanate, sodium salt; Tos-PhCH$_2$Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone; NMR, nuclear magnetic resonance.

[2] Results of established procedures can be found in supplementary material as described in the paragraph at the end of this paper.

	5				10					15					
1	Val	Leu	Ser	Asp	Ala	Glu	Trp	His	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	Ala	Glu	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 1: The amino acid sequence of arctic minke whale myoglobin. The hyphens between the amino acid residues have been omitted for clarity.

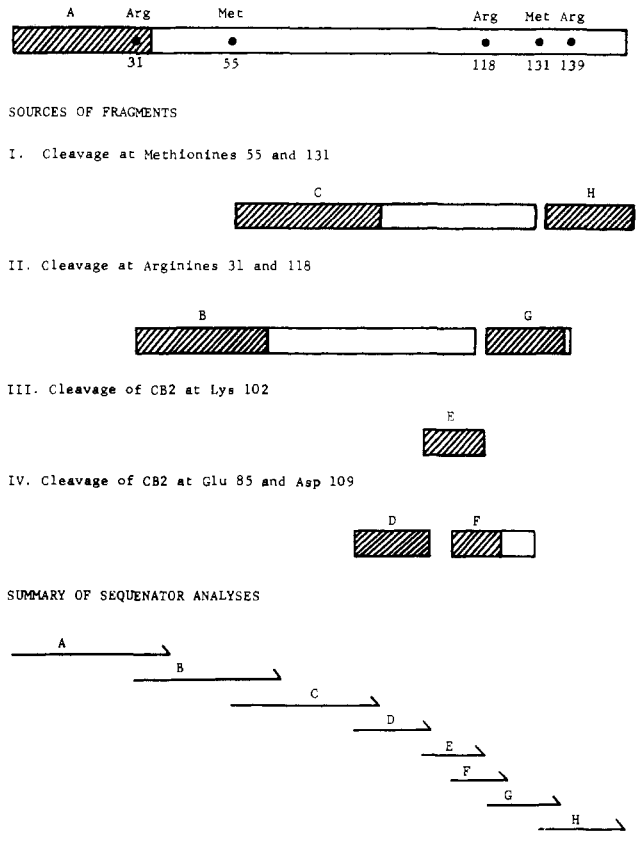

FIGURE 2: Diagrammatic summary of fragments generated from the arctic minke whale myoglobin for sequenator analysis. The top bar represents the whole myoglobin and the residues that are important for its fragmentation. The capital letters A–H identify the sequenator analyses in the order in which they are described in the text. A hatched section in each horizontal bar indicates the segment of the sequence determined by that analysis. A summary of overlaps is shown in the lower portion by the labeled arrows.

Residue Number	1	4	5	8	12	13	15	21	26	28	35	45
Species												
Minke Whale	Val	Asp	Ala	His	Asn	Ile	Ala	Val	Gln	Ile	Gly	Lys
A. B. Dolphin	Gly	Asp	Gly	Gln	Asn	Val	Gly	Leu	Gln	Val	Gly	Lys
A. R. Dolphin	Gly	Asp	Gly	Gln	Asn	Ile	Gly	Leu	Gln	Val	Gly	Lys
Common Porpoise	Gly	Glu	Gly	Gln	Asn	Val	Gly	Leu	Gln	Val	Gly	Lys
B. S. Dolphin	Gly	Asp	Gly	Gln	Asn	Val	Gly	Val	Glu	Ile	Gly	Lys
Gray Whale	Val	Asp	Ala	Gln	Asn	Ile	Ala	Val	Gln	Ile	Gly	Lys
Sperm Whale	Val	Glu	Gly	Gln	His	Val	Ala	Val	Gln	Ile	Ser	Arg

Residue Number	54	66	74	83	85	109	121	122	129	144	131	152
Species												
Minke Whale	Glu	Asn	Gly	Glu	Glu	Asp	Ala	Glu	Ala	Ala	Phe	Gln
A. B. Dolphin	Asp	Asn	Ala	Asp	Glu	Glu	Ala	Glu	Gly	Ala	Phe	His
A. R. Dolphin	Glu	Asn	Gly	Glu	Glu	Glu	Gly	Asp	Ala	Ala	Phe	His
Common Porpoise	Glu	Asn	Gly	Glu	Asn	Glu	Ala	Glu	Gly	Thr	Phe	His
B. S. Dolphin	Asp	Asp	Ala	Asp	Glu	Glu	Ala	Gln	Gly	Ala	Phe	His
Gray Whale	Glu	Asn	Gly	Glu	Glu	Asp	Gly	Asp	Ala	Ala	Phe	Gln
Sperm Whale	Glu	Val	Ala	Glu	Glu	Glu	Gly	Asn	Gly	Ala	Tyr	Gln

FIGURE 3: 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. B. dolphin is Atlantic bottlenosed dolphin, A. R. dolphin is Amazon River dolphin, and B. S. dolphin is Black Sea dolphin.

	GRAY WHALE	MINKE WHALE	COMMON PORPOISE	COMMON DOLPHIN	BOTTLENOSED DOLPHIN	AMAZON RIVER DOLPHIN
SPERM WHALE	12	14	15	14	15	15
GRAY WHALE		3	14	14	14	7
MINKE WHALE			13	14	13	10
COMMON PORPOISE				11	6	7
COMMON DOLPHIN					5	11
BOTTLENOSED DOLPHIN						7

FIGURE 4: Difference matrix for Cetacean myoglobins obtained by summing the number of different amino acids between pairs of proteins.

quence to position 122, overlapping arginine 118. Peptide MT3 (119–139) was used in analysis G to overlap peptide PCB2-6 (110–122) and extend the sequence to position 135. The final sequencer analysis H of peptide CB3 (132–153) overlapped analysis G starting at position 132 and extended the sequence to the carboxyl terminus of the protein at position 153.

Because of the low repetitive yield of sequenator analysis F, despite clearcut results with little carryover, the carboxyl terminal sequence of PCB2-6 (110–122) was determined by time course digestion with carboxypeptidase C.[2] This procedure reconfirmed the amino acid sequence around arginine 118.

In all sequencer runs the yields for phenylthiohydantoins were similar to those previously discussed (Dwulet et al., 1975).

Discussion

The present report is the fourth in a series[3,4] of complete cetacean myoglobin sequences determined by automated Edman degradation. The information derived from these sequence investigations has been used in the study of electrostatic interactions within the myoglobin molecule (Shire et al., 1975), the comparison of the oxygen equilibria of myoglobins from different species (Shire, 1974), and the interpretation of proton NMR results in which pK_a values of individual histidine residues were assigned (Botelho, 1975). The sequence data have also been used to develop a computer model of cetacean phylogenetics (Bogardt, 1977).

The sequence of minke whale myoglobin is compared in Figure 3 with the known cetacean myoglobins. As seen in the difference matrix in Figure 4, minke whale myoglobin has the closest similarity in sequence to another *Balaenoptera* whale, California gray whale, from which it differs by only three amino acids. The sequence of minke whale myoglobin will be examined here in comparison to the differences in sequence from the California gray whale. These will be referred to with the residue found in minke whale myoglobin, followed by the homologous California gray whale residue in parentheses.

Position 8 Histidine (*Glutamine*). Glutamine is the common residue for this position. Histidine at this position has not been found in other cetacean myoglobins. Histidine has been found at this position in the pinnipedia such as harbor seal (Bradshaw and Gurd, 1969). The presence of a histidine at this position should enable minke whale to serve as a model to show what role position 8 plays in denaturation by cupric ion (Hartzell et al., 1968) and the corresponding renaturation process (Marks et al., 1971). This substitution is the only charge change between California gray whale and minke whale myoglobins.

Position 121 Alanine (*Glycine*). Both alanine and glycine are common in other cetacean myoglobins at this position.

Position 122 Glutamic Acid (*Aspartic Acid*). These acid residues and their amides, glutamine and asparagine, are common at this position, as can be seen in Figure 3.

The above changes are conservative and the analogous residues of sperm whale myoglobin are found on the surface of the molecule (Watson, 1969). These changes are compatible with the sperm whale myoglobin three-dimensional structure and no significant change in the backbone conformation of the minke whale myoglobin is expected to occur.

[3] F. E. Dwulet, work in progress.

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

Acknowledgments

The authors are grateful to Mrs. K. Verhey and Mrs. D. Embry for their excellent technical help. We are grateful to Dr. Edward Mitchell and Dr. William Garner for helpful advice.

Supplementary Material Available

Experimental material including elution profiles, sequencer repetitive yield plots, carboxypeptidase C time course plot, and amino acid composition tables (21 pages). Ordering information is given on any current masthead page.

References

Bogardt, R. A. (1977), Ph.D. Thesis, Indiana University.

Bogardt, R. A., Dwulet, F. E., Lehman, L. D., Jones, B. N., and Gurd, F. R. N. (1976), *Biochemistry 15,* 2597.

Botelho, L. H. (1975), Ph.D. Thesis, Indiana University.

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

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

Dwulet, F. E., and Gurd, F. R. N. (1976), *Anal. Biochem. 76,* 530.

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

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

Hartzell, C. R., Bradshaw, R. A., Hapner, K. D., and Gurd, F. R. N. (1968), *J. Biol. Chem. 243,* 690.

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

Jones, B. N., Vigna, R. A., Dwulet, F. E., Bogardt, R. A., Lehman, L. D., and Gurd, F. R. N. (1976), *Biochemistry 15,* 4418.

Jones, W. C., Jr., Rothgeb, T. M., and Gurd, F. R. N. (1976), *J. Biol. Chem. 251,* 7452.

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

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

Marks, R. H. L., Cordes, E. H., and Gurd, F. R. N. (1971), *J. Biol. Chem. 246,* 466.

Shire, S. J. (1974), Ph.D. Thesis, Indiana University.

Shire, S. J., Hanania, G. I. H., and Gurd, F. R. N. (1975), *Biochemistry 14,* 1352.

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

Teale, F. W. J. (1959), *Biochim. Biophys. Acta 35,* 543.

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