I). Apparently, the pattern of complementary charge interactions is rather similar for the two complexes.

One of the most significant features of the hypothetical complex between cytochrome b_5 and cytochrome c proposed by Salemme is that the hemes of the two proteins are nearly coplanar, and the closest approach between the two is 8.4 Å. Since there are no amino acid side chains between the two hemes which could function as intermediate electron carriers, electron transfer from cytochrome b_5 to cytochrome c would appear to involve a direct transfer of electrons from heme edge to heme edge, possibly by a thermally activated tunnelling mechanism (Hopfield, 1974; Jortner, 1976).

References

Argos, P., and Mathews, F. S. (1975), *J. Biol. Chem. 250,* 747.

Borgese, N., and Meldolesi, J. (1976), *FEBS Lett. 63,* 231.

Ferguson-Miller, S., Brautigan, D. L., and Margoliash, E. (1976), *J. Biol. Chem. 251,* 1104.

Hopfield, J. J. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3640.

Jortner, J. (1976), *J. Chem. Phys. 64,* 4860.

Matlib, M. A., and O'Brien, P. J. (1976), *Arch. Biochem. Biophys. 173,* 27.

Rogers, M. J., and Strittmatter, P. (1974), *J. Biol. Chem. 249,* 895.

Salemme, F. R. (1976), *J. Mol. Biol. 102,* 563.

Smith, H., Staudenmayer, N., and Millett, F. (1977), *Biochemistry 16* (preceding paper in this issue).

Smith, L., Davies, H. C., and Nava, M. (1974), *J. Biol. Chem. 249,* 2904.

Spatz, L., and Strittmatter, P. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 1042.

Staudenmayer, N., Ng, S., Smith, M. B., and Millett, F. (1977), *Biochemistry 16,* 600.

Staudenmayer, N., Smith, M. B., Smith, H. T., Spies, F. K., and Millett, F. (1976), *Biochemistry 15,* 3198.

Strittmatter, P. (1964), On Rapid Mixing and Sampling Techniques in Biochemistry, Chance, B., Eisenhardt, R. H., Gibson, Q. H., and Lunberg-Holm, K. K., Ed., New York, N.Y., Academic Press, pp 71–84.

Strittmatter, P., Rogers, M. J., and Spatz, L. (1972), *J. Biol. Chem. 247,* 7188.

Swanson, R., Trus, B. L., Mandel, N., Mandel, G., Kallai, O., and Dickerson, R. E. (1977), *J. Biol. Chem. 252,* 759.

Yu, C. A., Yu, L., and King, T. E. (1973), *J. Biol. Chem. 248,* 529.

Complete Primary Structure of the Major Component Myoglobin of Pacific Common Dolphin (*Delphinus delphis*)†

Chi-Chin Wang, Romualdo Avila, Barry N. Jones, and Frank R. N. Gurd*

ABSTRACT: The complete amino acid sequence of the major component myoglobin from Pacific common dolphin, *Delphinus delphis,* was determined by the automatic Edman degradation of several large peptides obtained by specific cleavages of the protein. More than 80% of the covalent structure was established by the degradation of the apomyoglobin and five peptides from: (1) cyanogen bromide cleavage at the two methionine residues, (2) trypsin cleavage of the acetimidated apomyoglobin at the three arginine residues, and (3) 2-*p*-nitrophenylsulfenyl-3-methyl-3′-bromoindolenine cleavage at the two tryptophan residues. The rest of the sequence was determined by use of the peptides prepared from further digestion of the central cyanogen bromide peptide with staphylococcal protease and trypsin. The primary structure of this myoglobin proved identical with that from the Atlantic bottlenosed dolphin, *Tursiops truncatus,* but showed four substitutions with respect to the sequence reported for the Black Sea dolphin which has also been given the designation *Delphinus delphis.*

This report presents the application of the automatic Edman degradation procedure to the determination of the primary structure of myoglobin from the Pacific common dolphin, *Delphinus delphis.* The general strategy and analytical procedures had been established and utilized for the determination of several myoglobin sequences recently reported (Dwulet et al., 1975), 1977; Bogardt et al., 1976; Jones et al., 1976; Lehman et al., 1977). In addition, the protein was cleaved at the tryptophan residues (Fontana, 1972) to produce a long peptide to assist in confirming some amino acid substitutions found between this protein and that from the same named species from the Black Sea (Kluh and Bakardjieva, 1971). Other Cetacean myoglobin sequences reported include common porpoise (Bradshaw and Gurd, 1969) and sperm whale (Edmundson, 1965).

Experimental Section

Materials

The major component of Pacific common dolphin myoglobin was isolated from frozen muscle tissue by the procedure of Hapner et al. (1968). Phosphate buffer, pH 6.5, ionic strength 0.1, was used at room temperature to fractionate the crude

† From the Department of Chemistry, Indiana University, Bloomington Indiana 47401. *Received May 19, 1977.* This is the 83rd paper in a series dealing with coordination complexes and catalytic properties of proteins and related substances. For the preceding paper, see Rothgeb and Gurd (1978). This work was supported by Public Health Service Research Grant HL-05556. R. Avila is supported by Latin American Scholarship Program of American Universities.

TABLE I: Amino Acid Composition of Pacific Common Dolphin Myoglobin.

Amino acid	No. of residues from acid hydrolysis[a]	No. of residues from the sequence
Asp	11.9	12
Thr	5.2	5
Ser	5.0	5
Glu	15.8	16
Pro	3.9	4
Gly	13.1	13
Ala	16.8	17
Val	5.8	6
Met	1.9	2
Ile	7.8	8
Leu	19.3	19
Tyr	1.9	2
Phe	7.0	7
Lys	19.9	20
His	12.2	12
Arg	2.9	3
Trp	1.9	2

[a] Samples were hydrolyzed with 5.7 N HCl in the usual manner for 24, 48, and 72 h. The values for serine and threonine are obtained by extrapolating to zero time. Isoleucine value was obtained by extrapolating to 96 h. The results obtained for all other residues were averaged.

extract and purify the principal component on C-50 CM[1]-Sephadex. The homogeneity of the purified myoglobin was verified by cellulose-acetate electrophoresis at pH 9.2 and 5.2. Apomyoglobin was prepared by the method of Teale (1959), as applied by Hapner et al. (1968).

BNPS-Skatole was prepared by the procedures of Fontana et al. (1966) and Omenn et al. (1970). Methyl acetimidate hydrochloride was prepared according to the method of Hunter and Ludwig (1962) and 3-sulfophenyl isothiocyanate sodium salt was prepared according to the procedure of Dwulet and Gurd (1976). TPCK-treated trypsin was purchased from Worthington Biochemical Laboratories and staphlococcal protease was obtained from Miles Laboratories Ltd. Sequenator reagents were all "Sequencer" grade obtained from Beckman Instruments. All other chemicals were the highest grade available.

Methods

All the enzymatic and chemical cleavages and peptide fractionation procedures were as described previously (Jones et al., 1976; Lehman et al., 1977).

Cleavage at Tryptophan Residues with BNPS-Skatole. Apomyoglobin (103 mg, 6 μmol) was dissolved in 3 mL of 50% acetic acid and added with stirring to a solution of BNPS-skatole (43 mg, 120 μmol) in 3 mL of glacial acetic acid. The mixture was allowed to stir at room temperature for 24 h. β-Mercaptoethanol [0.12 mL (2% v/v)] was added and the solution was incubated at 37 °C for 4 h to terminate the oxidation reaction and reduce any methionine sulfoxide. The mixture was then dialyzed exhaustively with deionized water followed by 5% acetic acid. The bright yellow precipitate of BNPS-skatole was separated by centrifugation. Purification of the cleavage product, peptide 15–153, was effected by recycling the supernatant through a Sephadex G-50 (Superfine) column (2.6 × 200 cm) three times. The flow was established with 10% acetic acid at a rate of 15 mL/h.

[1] Abbreviations used are: BNPS-Skatole, 2-*p*-nitrophenylsulfenyl 3-methyl-3′-bromoindolenine; Tos-PheCH$_2$Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone; CM, carboxymethyl.

```
                    5                  10                  15
  1   Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Val 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 Asp Met Lys Ala Ser Glu Asp
 61   Leu Lys Lys His Gly Asn Thr Val Leu Thr Ala Leu Gly Ala Ile
 76   Leu Lys Lys Lys Gly His His Asp 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   Ala Glu Phe Gly Ala Asp Ala Gln Gly 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 1: The complete amino acid sequence of Pacific common dolphin myoglobin.

Results

Amino Acid Composition. Table I summarizes the composition obtained from 24, 48, and 72 h hydrolysis followed by amino acid analysis, and from the sequence count.

Peptide Purification. The separation profile of the BNPS-skatole cleaved apomyoglobin consists of two major peaks.[2] By use of the amino acid composition and cyanate method (Stark and Smyth, 1963) for the N-terminal residue determination, BS1 was identified as uncleaved myoglobin and BS2 was identified as peptide 15–153. All the other peptides were generated and purified as described in the preceding papers of this series. The fractionation profiles and peptide compositions are presented in the supplementary materials (see paragraph at the end of this paper concerning supplementary material).

Sequence Investigations. Only the sequence data necessary to establish the complete primary structure are reported.

Sequenator Results. The complete sequence of Pacific common dolphin is presented in Figure 1. The strategy used is shown in Figure 2. The bar on the top represents the whole myoglobin sequence and the residues subject to the cleavage reactions are marked. The hatched sections in each bar represent the numbers of residues analyzed by the sequenator in each peptide. The numbers of residues represented by each hatched section are: A, 28; B, 24; C, 27; D, 44; E, 20; F, 16; G, 13; H, 18; and I, 22.

Discussion

Identity with Myoglobin of Atlantic Bottlenosed Dolphin. The sequence shown in Figure 1 is identical with that reported for the major myoglobin component of the Atlantic bottlenosed dolphin (Jones et al., 1976). The results of observations comparing different Cetacean myoglobins with respect to histidine proton magnetic resonance absorption (Botelho, 1975) and denaturation by copper ion[3] also fit with the identity. The two myoglobin preparations, once mixed, could not be separated electrophoretically or chromatographically. Two-dimensional peptide maps (Bennett, 1967) of the tryptic peptides from the whole protein and from each purified cyanogen bromide fragment showed completely identical patterns between these two species. The close evolutionary relationship between these two species implied by the identity of the major myoglobin component is significant. The differences from other Cetacean myoglobin sequences are discussed by Jones et al. (1976).

[2] The peptide isolation profiles, peptide compositions, and repetitive yield plots of each sequenator analysis are represented in the supplementary materials.

[3] R. Avila, work in progress.

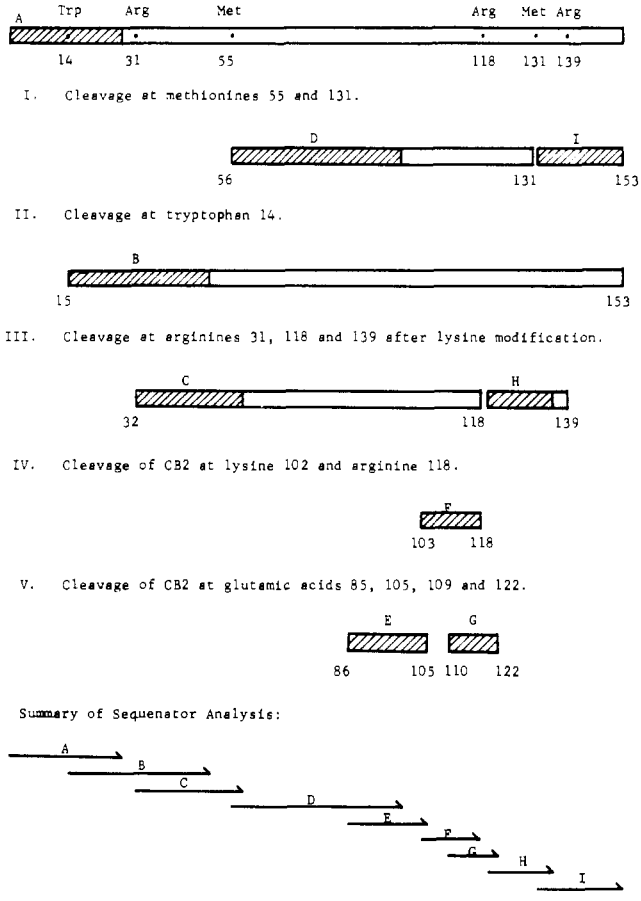

FIGURE 2: Summary of the peptides prepared from Pacific common dolphin myoglobin for sequenator analysis. See text for the number of residues analyzed in each of the hatched sections.

Comparison with Sequence of Black Sea Dolphin. The sequence shown in Figure 1 differs from that of the Black Sea dolphin (Kluh and Bakardjieva, 1971), also described as *Delphinus delphis,* at not less than 4 residue positions.[4] The differences occur at residues 21, 28, 66, and 122 and have been confirmed for the Pacific common dolphin myoglobin on samples prepared from a second animal. Figure 3 compares the assignments at these positions for a number of Cetacea. These different positions will be discussed here by referring first to the amino acid found in Pacific common dolphin myoglobin, then the sequence number, followed in parentheses by the residue reported for the Black Sea dolphin myoglobin.

Leucine-21 (Valine) and Valine-28 (Isoleucine). These two residues have been recognized (Bogardt et al., 1976) as examples of covarions (Fitch and Markowitz, 1970), which are concomitantly variable codons represented by a limited set of amino acids within a protein. Figure 3 shows that, with the exception of the myoglobin of the dwarf sperm whale, the residues 21 and 28 follow one of the two patterns compared here. The leucine–valine pair, found in the Pacific common dolphin, is characteristic of the small dolphins and porpoises. The valine–isoleucine pair, found in the Black Sea dolphin, is characteristic of the larger Cetaceans. These relationships are particularly interesting when the sequence data are used to develop phylogenetic relationships among these Cetaceans (Bogardt, 1977).

Residue Number	21	28	66	122
Species				
Pacific Common Dolphin	Leu	Val	Asn	Glu
Black Sea Dolphin	Val	Ile	Asp	Gln
Atlantic Bottlenosed Dolphin	Leu	Val	Asn	Glu
Amazon River Dolphin	Leu	Val	Asn	Asp
Common Porpoise	Leu	Val	Asn	Glu
Minke Whale	Val	Ile	Asn	Glu
Gray Whale	Val	Ile	Asn	Asp
Sperm Whale	Val	Ile	Val	Asn[a]
Dwarf Sperm Whale	Ile	Ile	Val	Asp

[a] Assigned by Romero Herrera and Lehman (1974) as aspartic acid

FIGURE 3: Comparison of four amino acid sequence positions of Cetacean myoglobins which have been completed to date.

Asparagine-66 (Aspartic Acid). The usual amino acid found at this position (Figure 3) is asparagine. The aspartic acid reported for the Black Sea dolphin represents a single base change from asparagine, as does the valine found in the sperm whale and dwarf sperm whale.

Glutamic Acid-122 (Glutamine). Figure 3 shows that both acids and amides occur at this position. In the case of the Pacific common dolphin the existence of glutamic acid at this position was further supported by the rather high yield of cleavage at this point by staphylococcal protease.[2]

Cleavage at Tryptophan Residues. This cleavage was obtained in satisfactory yield to provide the large peptide 15–153. Automatic sequencing of this peptide allowed reconfirmation of the region in which lower repetitive yields are encountered in the analysis of the unmodified apoprotein. It was useful for the confirmation of the assignments of residues 21 and 28 discussed above. This same large peptide forms the basis for a semisynthetic reconstruction of myoglobin (Wang et al., 1977).

Acknowledgments

We are grateful to Mrs. K. Verhey and Mr. K. Wong for their excellent technical help, Mrs. D. Embry for typing this manuscript, and Mr. K. Sexton for supplying the animal tissue. We are also grateful for the useful discussions and advice offered by Dr. G. Hanania, Dr. E. Mitchell, Dr. F. Dwulet, Mr. R. Bogardt, and Mr. L. Lehman.

Supplementary Material Available

Experimental results including elution profiles of peptides, amino acid compositions, and sequenator repetitive yield plots are provided (22 pages). Ordering information is given on any current masthead page.

References

Bennett, J. C. (1967), *Methods Enzymol. 11,* 330.

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

Bogardt, R. A., Jr., 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,* 267.

Dwulet, F. E., Bogardt, R. A., Jones, B. N., Lehman, L. D.,

[4] A fifth point of difference is uncertain. Kluh and Bakardjieva (1971) first assigned position 26 as glutamine and on subsequent repetition were only able to detect glutamic acid. It was thought possible that prolonged exposure to an alkaline medium might have caused deamidation.

and Gurd, F. R. N. (1975), *Biochemistry 14,* 5336.
Dwulet, F. E., and Gurd, F. R. N. (1976), *Anal. Biochem. 76,* 530.
Dwulet, F. E., Jones, B. J., Lehman, L. D., and Gurd, F. R. N. (1977), *Biochemistry 16,* 873.
Edmundson, A. B. (1965), *Nature (London) 205,* 389.
Fitch, W. M., and Markowitz, E. (1970), *Biochem. Genet. 4,* 579.
Fontana, A. (1972), *Methods Enzymol. 25,* 419.
Fontana, A., Marchiori, F., Rocchi, R., and Pajetta, P. (1966), *Gazz. Chim. Ital. 96,* 1301.
Hapner, K. D., Bradshaw, R. A., Hartzell, C. R., and Gurd, F. R. N. (1968), *J. Biol. Chem. 243,* 683.
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.
Kluh, J., and Bakardjieva, A. (1971), *FEBS Lett. 17,* 31.
Lehman, L. D., Dwulet, F. E., Bogardt, R. A., Jones, B. N., and Gurd, F. R. N. (1977), *Biochemistry 16,* 706.
Omenn, G. S., Fontana, A., and Anfinsen, C. B. (1970), *J. Biol. Chem. 245,* 1895.
Romero Herrera, A. E., and Lehman, H. (1974), *Biochim. Biophys. Acta 336,* 318.
Rothgeb, T. M., and Gurd, F. R. N. (1978), *Methods Enzymol.* (in press).
Stark, G. R., and Smyth, D. G. (1963), *J. Biol. Chem. 238,* 214.
Teale, F. W. J. (1959), *Biochim. Biophys. Acta 35,* 543.
Wang, C. C., Garner, W. H., and Gurd, F. R. N. (1977), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 36,* 890.

Covalent Structure of Cartilage Collagen. Amino Acid Sequence of Residues 363–551 of Bovine α1(II) Chains†

William T. Butler,* John Edward Finch, Jr., and Edward J. Miller

ABSTRACT: The covalent structures of α1(II)-CB11-C6, a chymotryptic peptide from the COOH terminus of α1(II)-CB11, and of α1(II)-CB8 from bovine nasal cartilage collagen are reported. These structures represent residues 363–551 of the bovine α1(II) chain. The sequence displays the repeating Gly-X-Y sequence characteristic of the triple helical portions of all α chains. Another phenomenon observed here, which is also true for other collagen α chains, was the occurrence of phenylalanyl and leucyl residues exclusively in the X positions of the repetitive triplet structure. When the amino acid sequence for this segment of α1(II) was compared with that of α1(I), the level of identity was 73%, a figure slightly lower than that for residues 1–162 at the NH_2-terminal triple-helical region (Butler, W. T., Miller, E. J., and Finch, J. E., Jr. (1976), *Biochemistry 15,* 3000). Three sites of occurrence of glycosylated hydroxylysines in the α1(II) chain were identified by the present studies. Two of these are galactosylhydroxylysines while the other site is a mixture of glucosylgalactosylhydroxylysine and galactosylhydroxylysine. One of the monosaccharides (at residue 408) and the mixture of mono- and disaccharides (at position 531) occur in positions occupied by lysines in the α1(I) chain. The other monosaccharide occurs in a site (residue 420) present as arginine in α1(I). A comparison of the sequences reported for residues 360–660 of α1(I), α2, and α1(III) chains with that of α1(II) reported here along with other unpublished data for α1(II) revealed 65 residues which are identical in these chains and, thus, are possibly "invariant." The frequency of occurrence of lysine, arginine, glutamic acid, and phenylalanine as invariant residues was higher than expected from their overall contents in collagen.

The type II collagen of hyaline cartilages is made of three α1(II) chains, each containing approximately 1050 amino acids. The α1(II) chains are similar in amino acid composition to the α1(I) and α2 chains of type I collagen and to the α1(III) chains of type III collagen (Miller, 1976) but have much higher levels of hydroxylysine-bound glucose and galactose (Miller, 1971; Trelstad et al., 1970; Miller, 1976). Recent studies from our laboratory have also shown that the majority of α1(II) hydroxylysine glycosides occurs in positions which are occupied by lysyl residues in the α1(I) chains (Butler et al., 1974a, 1976). The α1(II) chains were shown to be identical with α1(I) chains in about 80% of the positions (Butler et al., 1974a, 1976). One unusual discovery was that at least two distinct α1(II) chains are present in the bovine nasal septum (Butler et al., 1977). It was observed that three positions of α1(II) display sequence heterogeneity; that is, each is occupied by two amino acids. The chains were tentatively called α1(II)Major and α1(II)Minor to reflect the relative amounts. At present no information is available on the distribution or significance of the two chains.

In this publication we present data showing the primary structure at the COOH-terminal end of α1(II)-CB11, representing residues 363–402[1] of the triple-helical portion of α1(II), and the complete structure of α1(II)-CB8 (residues 403–551) (see the review by Miller, 1976, for clarification of

† From the Institute of Dental Research and the Department of Biochemistry, University of Alabama in Birmingham, University Station, Birmingham, Alabama 35294. *Received June 7, 1977.* This research was sponsored by U.S. Public Health Service Grant DE-02670 from the National Institute of Dental Research.

[1] Numbering begins with the NH_2-terminal triple helical portion of the collagen α chains (Hulmes et al., 1973; Fietzek and Kühn, 1976).