PENGUIN (*APTENODYTES FORSTERI*) MYOGLOBIN. A 70 RESIDUE N-TERMINAL SEQUENCE

S. PEIFFER, M. DECONINCK, C. PAUL, J. DEPRETER, A.G. SCHNEK and J. LÉONIS
Laboratoire de Chimie Générale I, Faculté des Sciences, Universite Libre de Bruxelles, Belgique

Received 29 June 1973
Revised version received 23 August 1973

1. Introduction

Since Edmundson [1] determined the amino acid sequence of the sperm whale myoglobin in 1965, this protein has been examined in several other species (land and marine mammals, gastropod).

With the purpose of elucidating evolutionary and genetic informations, comparative structural studies have been extended to a new zoological class: the birds. The partial covalent structure of chicken myoglobin was determined [2] first.

This paper presents the 70 residue N-terminal sequence of the penguin myoglobin. Residues have been positioned after sequence data obtained on peptides isolated from tryptic (T), chymotryptic (C), tryptic following maleylation of the protein (TM) and cyanogen bromide (CNBr) digestions of the globin. These results have been confirmed by stepwise degradation of the whole globin and of a large TM peptide using a sequencer.

2. Materials and methods

The main component of penguin myoglobin was prepared from muscles as previously described [3].

After having removed the heme moiety from the protein, the globin was denatured by heat (5 min, 100°C). It was submitted to tryptic hydrolysis by adding the enzyme to a 10 mg/ml substrate solution (4% w/w) and allowing the enzymatic digestion to be processed at 40°C for 150 min at pH 8.5 The globin was subjected to chymotryptic hydrolysis (pH 8.0, 40°C, 180 min) with an enzyme/substrate ratio of 3:100.

The separations of tryptic and chymotryptic peptides were performed by column chromatography (150 × 1.4 cm) on Dowex 50W × 2 resin at 40°C using a double linear gradient of pyridine acetate (pH 2.5 to 5.1 and 0.1 to 2.0 N in pyridine).

Each fraction was purified by high voltage paper electrophoresis at pH 3.6 or 6.5.

Maleylated myoglobin [4] was digested with trypsin at an enzyme/substrate ratio of 4:100 (w/w) for 240 min at 40°C and peptides were isolated on a 1.4 × 150 cm column of Sephadex G-25 in 0.01 N NH_3. After the removal of the maleyl groups by incubation in 0.1 N acetic acid for 48 hr at 45°C, some peptides were purified anew by high voltage paper electrophoresis at pH 6.5.

Reaction with CNBr was carried out as described by Gross and Witkop [5] and the isolation of peptides was performed by gel filtration on Sephadex G–25 and/or G–50 on a 150 × 1.4 cm column.

Amino acid compositions of the peptides were obtained with an automatic amino acid analyzer (Beckman Unichrom) and peptide sequences were determined by the combined dansyl–Edman technique [6, 7]. Dansyl amino acids were identified by two dimensional chromatography on polyamide sheets of 5.0 × 5.0 cm using solvents reported by Woods and Wang [8].

The C-terminal residues were identified by amino

acid analysis after kinetic hydrolysis with carboxypeptidases A and/or B.

The whole globin and a maleylated peptide (TM_3) were submitted to automatic Edman degradation [9] on the Beckman sequencer (model 890). The identification of the PTH amino acids was realized by gas-chromatography [10] (Beckman CG 45 chromatograph) and in some cases by automatic analysis of the amino acid liberated by acid hydrolysis (6 N HCl, 120°C, 16 hr) of the PTH derivative.

3. Results and discussion

The N-terminal sequence of penguin myoglobin obtained by the chymotryptic, tryptic and CNBr peptides overlaps and by automatic sequential degradation is presented in table 1.

A few particular points might be outlined:

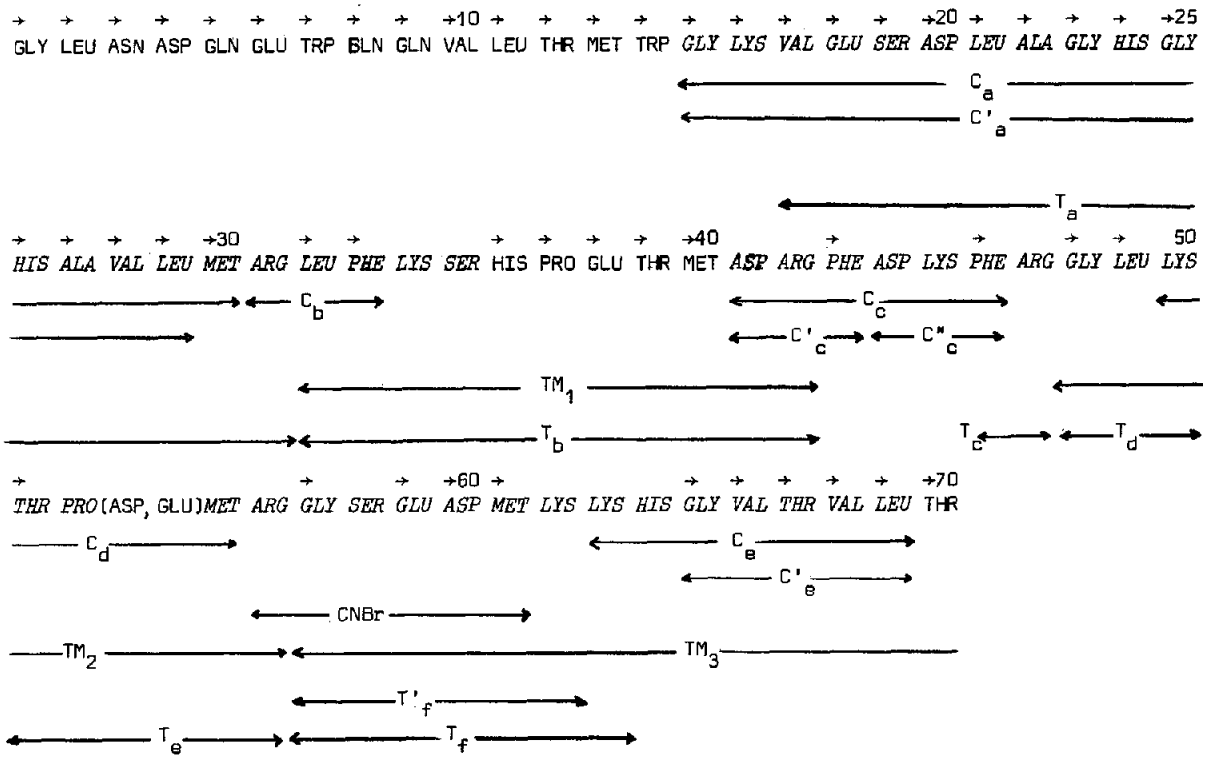
i) Since penguin myoglobin contains 8 residues of methionine (instead of 2 or 3 for mammalian globins), the gel filtration of the CNBr digest produced a complex mixture from which only one peptide was obtained in a satisfactory pure state after high voltage electrophoresis on paper;

ii) The peptide TM_3 isolated from the tryptic digest of maleylated myoglobin is a large fragment containing 42 residues ranging from the positions 57 to 98. Its N-terminal sequence was partially obtained by automatic sequential degradation;

iii) The residues Asp and Glu in positions 53 and 54 were positioned by homology with the mammalian myoglobin sequences already determined.

Table 1

Determination of the N–terminal sequence of penguin myoglobin.



← C, T, TM, CNBr → chymotryptic, tryptic, tryptic from maleylated myoglobin or cyanogen bromide fragments.

→ Amino acids identified by automatic sequential degradation.

Residues identified as dansyl derivatives or by carboxypeptidases are set in italic print.

The partial amino acid sequence of the penguin globin differs from sperm whale myoglobin at 25 sites on the first 70 residues and from chicken myoglobin at 8 sites on the first 47 residues.

When compared with all other species studied, penguin protein presents 10 substitutions never observed before. These replacements are Asn for Ser at position 3, Met for Val, Ala, Ile at position 13, Ser for Ala, Thr at position 19, Ala for Asp or Glu at position 27, Met for Leu at position 40, Arg for Lys at position 47, Gly for His at position 48, Pro for Glu at position 52, Gly for Ala, Arg at position 57 and Met for Leu at position 61.

Although it is difficult to discuss the evolution of globins with incomplete data, it clearly appears, on the basis of the amount of substitution, that the penguin protein is markedly different from mammalian myoglobins.

An examination of the amino acid sequence of this penguin hemoprotein in relation to the three dimensional structure of sperm whale myoglobin shows that the replacements occur both at external and internal positions of the molecule. The differences observed inside are generally interchanges between very similar residues, essentially hydrophobic amino acids.

Furthermore, the variable residues observed among the two proteins can be considered as differing most frequently by a single point mutation [11]. Only the replacements of Gly–5, His–12, His–48 and Glu–52 in sperm whale by Gln, Thr, Gly and Pro respectively in penguin are consequent upon the exchange of two bases in the codon triplet.

None of the substitutions observed seems to be of such a nature as to induce a drastic change in the spatial conformation exhibited by the sperm whale molecule. The particular substitution, of a Pro for Glu at the second residue of the D α helix (position 52) is perfectly compatible with the progression of the α helix [12].

Comparison of the amino acid sequences of the penguin protein with myoglobins from other species and physicochemical studies in solution on penguin, chicken and sperm whale myoglobins [3] confirm this conservation of structure.

Acknowledgements

Support from the Fonds National de la Recherche Scientifique and from the Fonds de la Recherche Fondamentale Collective is gratefully acknowledged.

The authors are greatly indebted to Professor F.W. Putnam of the Indiana University and to Professor J. Dirkx and Dr. C. Monier from the Université de Mons for the loan of a sequencer.

References

[1] Edmundson, A.B. (1965) Nature 205, 883.
[2] Deconinck, M., Depreter, J., Paul, C., Peiffer, S., Schnek, A.G., Putnam, F.W. and Léonis, J. (1972) FEBS Letters 23, 279
[3] Deconinck, M., Peiffer, S., Schnek, A.G. and Léonis, J. (1972) Biochimie 54, 969.
[4] Butler, P.J.G., Harris, J.I., Hartley, B.S. and Leberman, R. (1969) Biochem. J. 112, 679.
[5] Gross, E. and Witkop, B. (1962) J. Biol. Chem. 237, 1856.
[6] Dopheide, T.A.A., Moore, S. and Stein, W. (1967) J. Biol. Chem. 242, 1833.
[7] Gray, W.R. and Hartley, B.S. (1963) Biochem. J. 89, 59.
[8] Woods, K.R. and Wang, K.T. (1967) Biochim. Biophys. Acta 133, 369.
[9] Edman, P. and Begg, C. (1967) Eur. J. Biochem. 1, 80.
[10] Pisano, J.J. and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597.
[11] Nirenberg, M.W., Leder, P., Bernfield, M., Brimacombe, R., Trupin, J., Rottman, F. and O'Neal, C. (1965) Proc. Natl. Acad. Sci. U.S. 53, 1161.
[12] Flory, J.P. (1969) Statistical Mechanics of Chain Molecules, p. 270, Interscience.