Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 1917–1920, March 1982
Evolution

Ochotona princeps (pika) myoglobin: An appraisal of lagomorph phylogeny

(amino acid sequence/mammalian systematics/Leporidae/Ochotonidae/evolution)

HOWARD DENE*, MORRIS GOODMAN*, MALCOLM C. MCKENNA†, AND A. E. ROMERO-HERRERA*

*Department of Anatomy, School of Medicine, Wayne State University, Detroit, Michigan 48201; and †Department of Vertebrate Paleontology, American Museum of Natural History, New York, New York 10024

Communicated by Norman D. Newell, December 21, 1981

ABSTRACT **Myoglobin was purified from skeletal muscle of the pika (*Ochotona princeps*) and its primary structure was determined. This sequence was added to the set of 64 already known vertebrate myoglobin sequences and used to evaluate the phylogenetic position of the pika among other mammals, using a computerized search procedure based on maximum parsimony criteria. The pika is clearly related to the rabbit (*Oryctolagus cuniculus*), with which it is traditionally associated as a member of the order Lagomorpha. A monophyletic group composed of Lagomorpha, Scandentia, and Carnivora is a consistent feature of the dendrograms produced. An association of Carnivora and Lagomorpha casts doubt on the results of those investigations using rabbit antisera in systematic studies and depending on carnivores as an outside reference group for cladogram construction.**

Uncertainties have long persisted regarding the phylogenetic relationships of the mammalian order Lagomorpha (rabbits, pika) (1–3), due partly to the antiquity of the order, partly to an inadequately known fossil record, and partly to the poorly understood diversity within the order. This last point is certainly aggravated by a preoccupation of scientific investigators with laboratory rabbits (*Oryctolagus*) almost to the exclusion of other living lagomorph genera.

There are 12 living genera in this order, which contains a much larger number of species. These genera are distributed unequally between two families. Eleven belong to Leporidae, rabbits and hares, as do more than 20 fossil genera. In contrast, the other surviving lagomorph family, Ochotonidae, contains but a single living genus, *Ochotona*. However, 23 extinct genera are also placed in this family. Clearly, Ochotonidae was formerly a very diverse and important subdivision of the order.

The phylogenetic position of Lagomorpha is, of course, important to the overall pattern of eutherian evolution. Furthermore, conclusions of immunological systematic studies beyond the ordinal level are critically affected by the placement of the rabbit because this animal is the primary source of antibodies used in such studies (4). In order to establish the phylogenetic position of the Lagomorpha, one must attempt to determine which characters are shared by all lagomorphs, which are shared by advanced but not by primitive ones, and which are characters unique to individual branches of the lagomorph phylogenetic tree.

Amino acid sequence information should be able to make a valuable contribution to our phylogenetic understanding of Lagomorpha as protein studies considering more members of the order become available. Combined sequence analysis of data on α- and β-globin, myoglobin, cytochrome *c*, lens α-crystallin, and fibrinopeptides A and B suggests *Tupaia* as the nearest nonlagomorph relative, followed by Primates. The myoglobin sequence for *Oryctolagus cuniculus* has been available since 1976 (5). In an attempt better to understand lagomorph evolution, the primary structure was determined for myoglobin from the pika, *Ochotona princeps*, the living member of the family *Ochotonidae*.

MATERIALS AND METHODS

Extraction and Purification of Myoglobin. Myoglobin was extracted from 500 g of pika (*Ochotona princeps*) skeletal muscle with 750 ml of 2 mM KCN and centrifuged. The supernatant was submitted to ammonium sulfate fractionation (55% saturation), and the precipitated material was removed by centrifugation. After dialysis against dilute KCN (1 mM) the supernatant was concentrated by using an Amicon ultrafiltration unit equipped with a PM 10 membrane. The concentrated sample was then applied to a 2.5 × 180 cm column of Ultrogel AcA 54 (LKB) equilibrated in 50 mM Tris·HCl/2 mM KCN, pH 8.5 (flow rate 15 ml/hr). The myoglobin-containing fraction was then dialyzed and the heme group was removed by using 1.5% HCl in acetone. The myoglobin was further purified by ion-exchange chromatography (CM 23 carboxymethylcellulose, from Whatman), using a linear gradient of 10 to 40 mM Na_2HPO_4 in 1 mM dithiothreitol/8 M urea, pH 6.4 (6).

Determination of Amino Acid Sequence. A portion of myoglobin was subjected to cyanogen bromide (CNBr) cleavage (7). The resulting peptides were separated by gel filtration on a 2.5 × 270 cm column of Sephadex G-75 (Pharmacia), using 0.5% acetic acid for elution (6).

Intact apomyoglobin was digested with trypsin (8) and the resulting soluble peptides were separated by chromatography and electrophoresis on Whatman 3 MM paper (9, 10). The insoluble tryptic peptides were hydrolyzed with pepsin (11) and then the products were separated as the soluble peptides were.

CNBr peptides were subjected to enzymic hydrolysis using chymotrypsin, thermolysin (6), and V8 staphylococcal protease (12). Thermolysin was also used to digest tryptic peptides consisting of residues 17–31, 64–77, 78–96, 79–96, 80–96, and 119–133. In addition, the tryptic peptide containing residues 17–31 was hydrolyzed with Pronase, and the tryptic peptide containing residues 80–96 was digested with cathepsin (13).

Peptides were eluted from preparative chromatography/electrophoresis paper containing 5 mg of apomyoglobin each by using 6 M HCl, hydrolyzed 24 hr at 108°C, and analyzed by using a Beckman 119 CL automatic amino acid analyzer with

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: NR, nucleotide replacement(s).

attached Beckman 126 data programmer. Peptic peptides representing residues 110–115 and 111–115 were also hydrolyzed for 72 hr. Peptides for Edman degradation were eluted with 3% NH_4OH.

The amino acid sequence was established from overlapping enzymic and CNBr peptides as well as by dansyl-Edman degradation (14, 15). Dansyl derivatives were identified by two-dimensional chromatography on thin-layer polyamide plates (16).

RESULTS AND DISCUSSION

The amino acid sequence shown in Fig. 1 for pika skeletal muscle myoglobin was established from 131 overlapping peptides and by 84 steps of dansyl-Edman degradation done on selected peptides. This protein is composed of 153 amino acid residues. Amide and acidic side chains were established on the basis of the electrophoretic mobilities of small peptides at pH 6.5, using Offord's formula (17).

This sequence was added to the set of 64 already known vertebrate myoglobin sequences and used to examine the phylogenetic position of the pika among mammals. The sequence information was analyzed by using maximum parsimony procedures outlined by Goodman *et al.* (18). As expected on the basis of morphological considerations and current opinions on mammalian taxonomy, the rabbit and pika are related. The phylogenetic position of the order Lagomorpha as a whole is less clearly resolved, however.

FIG. 1. Amino acid sequence of pika (*Ochotona princeps*) skeletal muscle myoglobin, obtained from overlapping enzymic and chemically derived peptides and sequential dansyl Edman degradation. ↑, CNBr cleavage; ↑, enzymic hydrolysis; →, dansyl Edman degradation. Peptides: Tp, tryptic; Pe, peptic; Ch, chymotryptic; V8, *Staphylococcus aureus* V8 protease; Pro, Pronase; Th, thermolysin peptides from tryptic peptide residues 17–31 and from CNBr peptide residues 56–131; *Th, thermolysin peptides from tryptic peptides residues 64–77, 78–96, 79–96, 80–96, 119–133; Cath, cathepsin; CNBr, cyanogen bromide peptides. HS, homoserine; HSL, homoserine lactone.

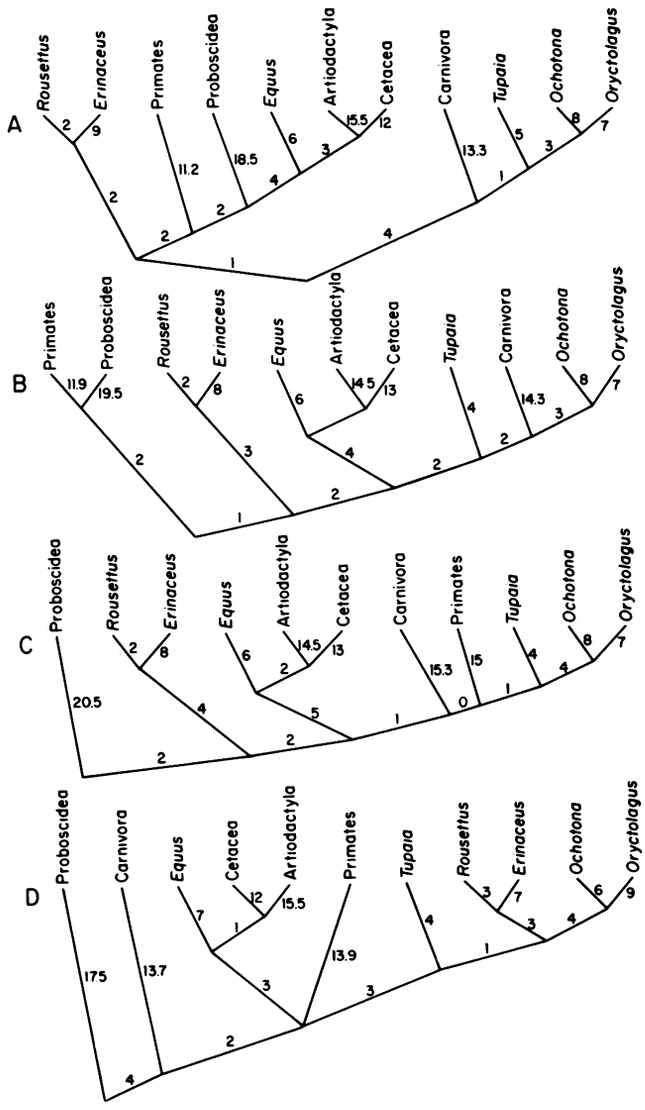

FIG. 2. Dendrograms representing possible arrangements of various eutherian mammal taxa for which myoglobin amino acid sequences are known. In cases for which myoglobin sequences are known for more than one genus in an order (Primates, Proboscidea, Artiodactyla, Cetacea, and Carnivora) only the order is shown. Generic names are given where only one genus in an order has been studied and for the two lagomorphs (*Ochotona* and *Oryctolagus*). The numbers assigned to links in these trees represent the numbers of nucleotide replacements (NR) necessary to account for myoglobin changes between branching points. *C* shows a zero link length between the primate and carnivore branch. Normally this would be represented as a trichotomy. However, because this arrangement is based on information from additional proteins and some of these would require NR along this link, separate branching points for Carnivora and Primates have been retained.

After the addition of the pika primary structure to the already known myoglobin sequences, a relationship of Lagomorpha to Scandentia (*Tupaia*) is consistently found. Either the tree shrew represents the closest relative of the lagomorphs (Fig. 2 *A* and *C*) or it is separated from lagomorphs by one additional branch (Fig. 2 *B* and *D*). In addition to Scandentia, Carnivora sometimes cluster with the same monophyletic group as lagomorphs (Fig. 2 *A–C*). Although the most parsimonious trees found by computer search (Fig. 2 *A* and *B*) exclude Primates from a monophyletic group including *Tupaia* and lagomorphs, one tree requiring six additional nucleotide replacements (NR) suggests its inclusion (Fig. 2*C*).

The two most parsimonious trees so far identified (759 NR) both contain a monophyletic group made up of Lagomorpha (rabbit, pika), Scandentia (*Tupaia*), and Carnivora, but they differ in the branching arrangement within this group (Fig. 2 *A* and *B*). A similar monophyletic group is also favored by combined sequence data for seven different protein chains (α-globin, β-globin, myoglobin, cytochrome *c*, lens α-crystallin, and fibrinopeptides A and B), the only difference being the inclusion of Primates in this group (Fig. 2*C*).

In the dendrogram presented in Fig. 2*A*, the monophyletic group including Lagomorpha, Scandentia, and Carnivora represents the most ancient branch of the eutherian tree. On a tree of equal length (759 NR, Fig. 2*B*) and on the tree favored by combined sequence data (765 NR, Fig. 2*C*), this monophyletic group is joined first by a group composed of Artiodactyla, Perissodactyla, and Cetacea, then by a bat–hedgehog branch, and finally by a branch including the elephants, either by themselves (Fig. 2*C*) or in combination with Primates (Fig. 2*B*). Fig. 2*C* differs most significantly from Fig. 2*B* by the inclusion of Primates in the same monophyletic group as lagomorphs, carnivores, and *Tupaia*.

A fourth alternative is also offered (Fig. 2*D*) that although requiring only two additional NR (761 NR), differs considerably from those discussed above. This tree, unlike other trees of low NR length, depicts a bat–hedgehog branch as the closest relative of lagomorphs and relegates carnivores to a more distant position.

Immunological investigations using chicken antisera have also tended to favor a relationship between Carnivora and Lagomorpha similar to that reported here (19). Because so much immunological work is based on results obtained by using antisera made in rabbits (20–22) and because some investigators use Carnivora as a reference group for construction of phylogenetic trees by using the additive approach (22), the relationship between these two orders is critically important. If they are indeed closely related, rabbit antisera to Carnivora must be recognizing a more restricted and probably very different set of antigenic changes from those recognized by antisera to phylogenetically more distant orders. The relationship between results involving carnivores and those involving more distant groups are thus more complex than usually assumed and, in fact, not necessarily directly comparable.

The earliest known lagomorphs are from the Paleocene of Asia (*Mimotona*) and are very similar to other Tertiary Asian fossil mammals such as eurymylids, early rodents, various anagalid-like animals, and the late Cretaceous genera *Zalambdalestes* and *Barunlestes* (3). Clearly, lagomorphs have been separated from their closest relatives since some time within the Paleocene. The order has undoubtedly been independent of its nearest surviving relatives for at least 60 million years. Because forms representing immediate structural antecedents (*Megalagus* and *Mytonolagus*) of the two surviving lagomorph families existed at the Eocene–Oligocene boundary 38 million years ago (23) and the earliest point at which both palaeolagine leporids and ochotonids can be recognized definitely is about 32 million years ago (24), the time of divergence for the Ochotonidae and Leporidae must be between 32 and 38 million years.

If one favors the dendrogram in Fig. 2*C*, which is based on seven protein chains and is in agreement with one of the two most parsimonious myoglobin trees (Fig. 2*B*) in many respects, Ochotonidae and Leporidae have fixed one NR every 4.3 and 5 million years, respectively. For Lagomorpha as a whole, one NR has been fixed every 3.3 million years, whereas its nearest surviving relative (*Tupaia*) has fixed only one every 16.7 million years.

This work has been supported by National Science Foundation Grants DEB 6719924, DEB 7810717, and PCM 7717644.

1. Wood, A. E. (1975) *Evolution* **11**, 417–425.
2. Van Valen, L. (1964) *Evolution* **18**, 484–491.
3. McKenna, M. C. (1975) in *Phylogeny of the Primates: A Multidisciplinary Approach*, eds. Luckett, W. P. & Szalay, F. S. (Plenum, New York), pp. 21–46.
4. Dene, H., Goodman, M., Prychodko, W. & Matsuda, G. (1980) in *Comparative Biology and Evolutionary Relationships of Tree Shrews*, ed. Luckett, W. P. (Plenum, New York), pp. 269–291.
5. Romero-Herrera, A. E., Lehmann, H. & Castillo, O. (1976) *Biochim. Biophys. Acta* **439**, 51–54.
6. Dene, H., Goodman, M. & Romero-Herrera, A. E. (1980) *Proc. R. Soc. London Ser. B* **207**, 111–127.
7. Gross, E. & Witkop, B. (1962) *J. Biol. Chem.* **237**, 1856–1860.
8. Sick, K., Beale, D., Irvine, D., Lehmann, G., Goodall, P. T. & MacDougall, S. (1967) *Biochim. Biophys. Acta* **140**, 231–242.
9. Ingram, V. M. (1958) *Biochim. Biophys. Acta* **28**, 539–545.
10. Baglioni, C. (1961) *Biochim. Biophys. Acta* **48**, 392–396.
11. Romero-Herrera, A. E. & Lehmann, H. (1974) *Proc. R. Soc. London Ser. B* **186**, 249–279.
12. Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3506–3509.
13. Dene, H., Sazy, J., Goodman, M. & Romero-Herrera, A. E. (1980) *Biochim. Biophys. Acta* **624**, 397–408.
14. Gray, W. R. (1967) *Methods Enzymol.* **11**, 469–475.
15. Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822.
16. Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta* **133**, 369–370.
17. Offord, R. E. (1966) *Nature (London)* **211**, 591–593.
18. Goodman, M., Czelusniak, J., Moore, G. W., Romero-Herrera, A. E. & Matsuda, G. (1979) *Syst. Zool.* **28**, 132–163.
19. Shoshani, J., Goodman, M., Barnhart, M. I., Prychodko, W., Vereshchagin, N. K. & Mikhelson, V. N. (1981) in *The Magadan Baby Mammoth, Collection of Articles*, ed. Vereshchagin, N. K. (Nauka, Leningrad), pp. 191–217.
20. Dene, H., Goodman, M. & Prychodko, W. (1980) *Mammalia* **44**, 211–223.
21. Goodman, M. & Moore, G. W. (1971) *Syst. Zool.* **20**, 19–62.
22. Sarich, V. M. & Cronin, J. (1976) in *Molecular Anthropology: Genes and Proteins in the Evolutionary Ascent of Primates*, eds. Goodman, M. & Tashian, R. E. (Plenum, New York), pp. 141–170.
23. Dawson, M. R. (1970) *Ann. Carnegie Mus.* **41**, 215–230.
24. Simpson, G. G. (1945) *Bull. Am. Mus. Nat. Hist.* **85**, 1–350.