

## Research Article

# Taxonomy of Mitochondrial Cytochrome B Proteins of the Same Amino Acid Sequence Length

Alexander A. Zamyatnin,<sup>1</sup> Tatiana A. Belozerskaya <sup>1</sup> and Andrey A. Zamyatnin Jr. <sup>2,3,4</sup>

<sup>1</sup>A.N. Bach Institute of Biochemistry, Federal Research Center of Biotechnology, Russian Academy of Sciences, Moscow 119071, Russia

<sup>2</sup>Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russia

<sup>3</sup>Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia

<sup>4</sup>Department of Biotechnology, Sirius University of Science and Technology, Sochi, Russia

Correspondence should be addressed to Andrey A. Zamyatnin Jr.; [zamyat@belozersky.msu.ru](mailto:zamyat@belozersky.msu.ru)

Received 31 August 2021; Accepted 26 October 2021; Published 11 November 2021

Academic Editor: Piotr Minkiewicz

Copyright © 2021 Alexander A. Zamyatnin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prior to this study, we discovered a protein characterized by many different amino acid sequences with the same number of amino acid residues. This turned out to be a unique cytochrome b, in which 1048 molecules out of 1689 contain 379 amino acid residues. A detailed study of the occurrence of this protein in living organisms at different taxonomic levels (from biological domains to biological orders of animals) has been carried out in the work presented here. We found that the main part of all b cytochromes is present in eukaryotes (99.2%), in biological kingdoms (95.9% in animals), in biological phylums (97.5% in chordates), and in biological classes (79.7% in mammals). Withal, this protein, containing 379 amino acid residues and characterized by many different amino acid sequences, is found only in eukaryotes (100%), only in animals (100%) and mainly in mammals (81.1%). Thus, a representative that has cytochrome b with a corresponding number of amino acid residues has not yet been identified among archaea and prokaryotes, while it is common in representatives of different biological types, classes, and orders of animals. It is believed that the structural diversity of a given protein within the same length and its one function of participation in the process of electron transfer relate to the physicochemical features of the extra- and intramembrane fragments of the polypeptide chain of this protein.

## 1. Introduction

The length (number of amino acid residues) of natural peptide structures varies over a very wide range. Thus, according to data from the protein-peptide database Swiss-Prot (<https://www.uniprot.org/uniprot/?query=reviewed:yes>), the minimum peptide structure consists of two, and the largest one consists of several tens of thousands of amino acid residues. Despite the fact that more than 500,000 natural amino acid sequences have already been described in detail (and despite their wide range of sizes), the length of protein molecules is rarely the subject of special consideration among other physicochemical characteristics.

Known investigations are mainly devoted to study of the shape of the distribution of peptide structures along the

length within the natural range of the number of amino acid residues [1–4]. At the same time, various techniques are used to smooth out the irregularities of such distributions and to obtain the corresponding fitted mathematical expressions. Thus, peaks presented in the natural distribution are excluded from consideration, for which no explanation is given.

We have previously drawn attention to peaks in the distributions of different biological kingdoms and domains and have identified a unique protein which is represented by more than 1000 different amino acid sequences with a single length of 379 amino acid residues [5, 6]. This turned out to be mitochondrial cytochrome b, identified in numerous representatives of different biological species of animals. A detailed analysis of the occurrence of this protein in

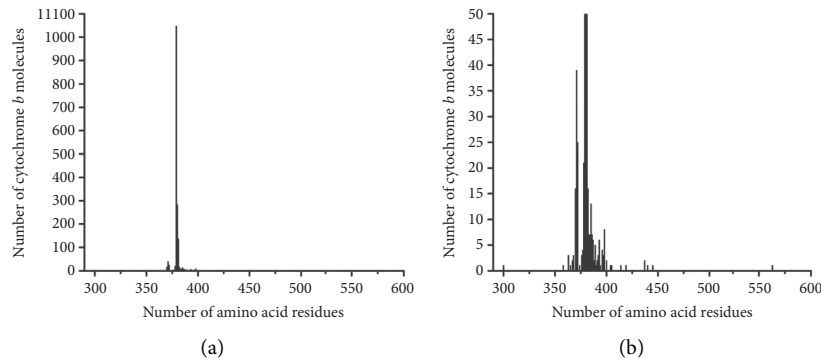


FIGURE 1: Distribution of the number of all identified cytochrome b molecules, according to data from the Swiss-Prot database on the number of amino acid residues. (a) Data at the ordinate, illustrating the entire maximum peak of the distribution. (b) The same data with a truncated ordinate and the same abscissa to identify all small peaks.

TABLE 1: Content of the number of cytochrome b amino acid sequences in the swiss-prot database in various biological domains.

Biological domains	Number of molecules	Number of amino acid residues
<i>Archaea</i>	1	563
<i>Prokaryota</i> (bacteria)	12	398, 400, 405, 414, 419, 437, 440, 445
<i>Eukaryota</i>	1 676	300–404
All	1 689	300–563

individual taxonomic groups of living organisms has been carried out in this work.

## 2. Methods

We used data from the Swiss-Prot protein-peptide database [7], which is a part of the UniProt database [8], in our investigation. The ability to perform different procedures and analyses is provided on the UniProt database website (<https://www.uniprot.org/uniprot/?query=reviewed:yes>). Thus, the UniProt database program tools were applied to our work highlighting all and specific amino acid residue sequences (option: Search), exclusion of sequence fragments (option: Sequence > Fragment > Sequence complete), extraction of specific sequences with a given number of amino acid residues  $p$  (option: Sequence > Sequence length > from  $p1$  to  $p2$ ), exclusion of identical sequences (option: Protein page > Similar proteins > 100% identity), and sorting protein names and sequences by different characteristics (standard UniProt table sorting).

At the time of the study, this contained information about 562,755 amino acid sequences, obtained for representatives of archaea, prokaryotes, and eukaryotes. The minimum number of amino acid residues (2) in it contains two oligopeptides [9, 10], and the maximum (35,213) contains one protein (mouse titin [11]). The Swiss-Prot database contains data not only on complete amino acid sequences but also on protein fragments. However, the database service makes it possible to exclude incomplete sequences. Data on 553,531 sequences were identified in the database after an appropriate procedure for the elimination of fragments. All b cytochromes known to date were isolated from these. The isolation and processing of data on these proteins has already been described in detail by us [6].

## 3. Results

At the first stage, we identified all cytochrome b molecules of all living organisms from the entire dataset of the Swiss-Prot database. As a result, we were able to ascertain the distribution of 1689 of these cytochromes, according to the number of amino acid residues. It turned out that the number of amino acid residues contained in them can vary from 300 to 563. These data are illustrated in Figure 1(a), in which one large peak (corresponding to 1048 molecules containing 379 amino acid residues) stands out. In addition to this peak, 283 molecules made up of 380 residues, and 136 molecules made up of 381, can be identified. These are clearly visible in Figures 1(a) and 1(b). These figures also show a small number of molecules containing both more and fewer than 379 amino acid residues.

Let us consider these data in more detail in taxonomic groups at different levels. At the first (highest) level, we selected b cytochromes from the domains of archaea, prokaryotes, and eukaryotes [12]. Most known cytochrome b molecules were detected in eukaryotes (99.2%), as can be seen from the data in Table 1.

The same results are graphically presented in Figure 2. From the data in this figure and Table 1, it also follows that b cytochromes containing 379 amino acid residues are found only in eukaryotes (100%); the regions where these lengths occur in prokaryotes and eukaryotes overlap slightly, and the value of 563 amino acid residues for a single representative of archaea goes far beyond the values for both eukaryotes and prokaryotes. Thus, most of the known cytochrome b molecules are found in eukaryotes, and they contain 379 amino acid residues in most cases.

Next, we analyzed data for 1676 b cytochromes of various biological kingdoms (only the eukaryotic domain).

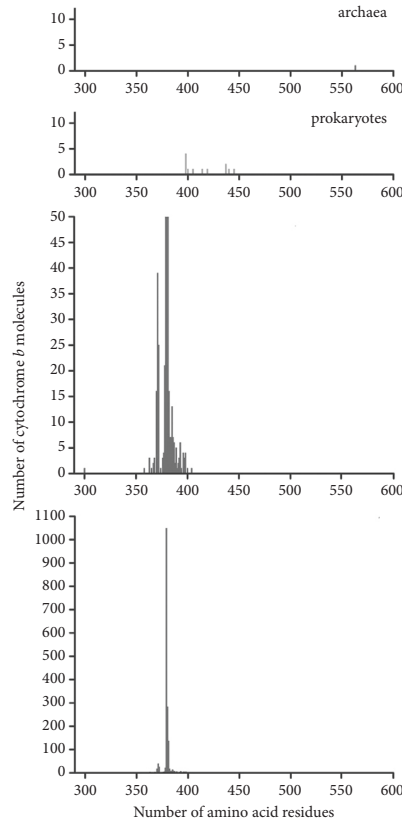


FIGURE 2: Distribution of the number of cytochrome *b* molecules, according to data from the Swiss-Prot database on the number of amino acid residues in various biological domains. The data for eukaryotes are presented at two scales of the ordinate—a truncated one, revealing small peaks throughout the entire region where lengths of amino acid residues exist, and a full one, where the entire maximum peak of the distribution is visible.

TABLE 2: Content of the number of cytochrome *b* amino acid sequences in the Swiss-Prot database in various biological kingdoms of the eukaryote (*Eukaryota*) domain.

Biological kingdoms	Number of molecules	Number of amino acid residues
<i>Metazoa</i> (animals)	1 608	358–398
<i>Fungi</i> (fungi)	38	382, 383, 384, 385, 386, 387, 388, 389, 390, 400
<i>Viridiplantae</i> (plants)	12	381, 388, 392, 393, 394, 397, 398, 404
<i>Alveolata</i>	9	300, 363, 368, 376, 382, 391
<i>Euglenozoa</i>	3	363, 371, 372
<i>Amoebozoa</i>	3	385, 387, 389
<i>Rhodophyta</i>	2	381, 384
<i>Stramenopiles</i>	1	383
<b>All <i>Eukaryota</i></b>	<b>1 676</b>	<b>300–404</b>

The results are shown in Table 2 in descending order according to the number of molecules detected. In this case, 95.9% of the molecules were identified in animals, and 65% (1048) of the molecules contained 379 amino acid residues. Figure 3 shows how the intervals of the number of amino acid residues in animal fungi and plants relate to one another. It is clear that the regions where the lengths of amino acid residues occur in animals and fungi, as well as in animals and plants, do not overlap, while the overlap for plants and fungi is almost complete. The main result is that all *b* cytochromes of both plants and fungi contain more amino acid residues than the *b* cytochromes of animals. The size of

the area where the lengths of cytochrome *b* occur in representatives of species not included in the considered kingdoms (300–391) turned out to be rather wide, overlapping with the area characteristic of animals and slightly overlapping with the same area in plants and fungi. In addition, none of them contained 379 amino acid residues; i.e., all proteins with this number of amino acid residues were found only in animal representatives.

All *b* cytochromes of animals representing the biological phylum were considered at the next taxonomic level. In this case (Table 3 and Figure 4), it turned out that the overwhelming number of molecules of this protein is

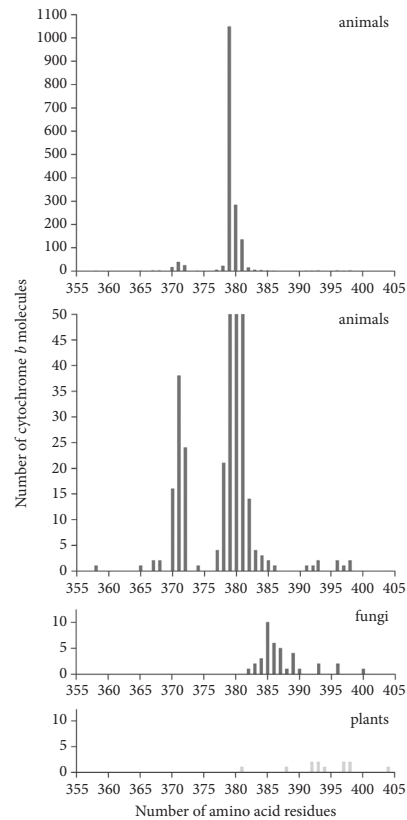


FIGURE 3: Distribution of the number of cytochrome b molecules, according to data from the Swiss-Prot database on the number of amino acid residues in various biological kingdoms. The data for animals are presented at two scales of the ordinate—a truncated one, revealing small peaks throughout the entire region where lengths of amino acid residues exist, and a full one, where the entire maximum peak of the distribution is visible.

TABLE 3: Content of the number of cytochrome b amino acid sequences in the Swiss-Prot database in various biological phyla of animal (*Metazoa*) kingdom.

Biological phyla and other taxonomic groups of animals	Number of molecules	Number of amino acid residues
<i>Chordata</i> (chordate)	1 568	367–398
<i>Arthropoda</i> (arthropods)	16	358, 377, 378, <b>379</b> , 381, 382, 383
<i>Mollusca</i> (mollusks)	4	367, <b>379</b> , 380, 397
<i>Cnidaria</i> (cnidarians)	3	384, 386, 393
<i>Echinodermata</i> (echinoderms)	3	<b>379</b> , 380
<i>Nematoda</i> (nematodes)	3	365, 370
<i>Annelida</i> (annelids)	2	<b>377</b> , <b>379</b>
<i>Acanthocephalan</i> (thorny-headed worms)	1	372
<i>Bryozoa</i> (moss animals)	1	368
<i>Chaetognatha</i> (arrow worms)	1	378
<i>Entoprocta</i> (goblet worms)	1	378
<i>Onychophora</i> (velvet worms)	1	<b>379</b>
<i>Placozoa</i> (tricoplaxes)	1	383
<i>Priapulida</i> (penis worms)	1	377
<i>Rotifer</i> (rotifers)	1	<b>379</b>
<i>Xenacoelomorpha</i> (xenoturbellids)	1	380
<b>All Metazoa</b>	<b>1 608</b>	<b>358–398</b>

characteristic of representatives of the biological phylum of chordates (97.5%). In addition, it was found that the number of molecules containing 379 amino acid residues (1042) in chordates is 66%. At the same time, single molecules with

this number of residues were also found in animals of other biological phyla (highlighted in bold in Table 3). These are represented by arthropods (migratory locust, *Locusta migratoria*), mollusks (spear squid, *Heterololigo bleekeri*),

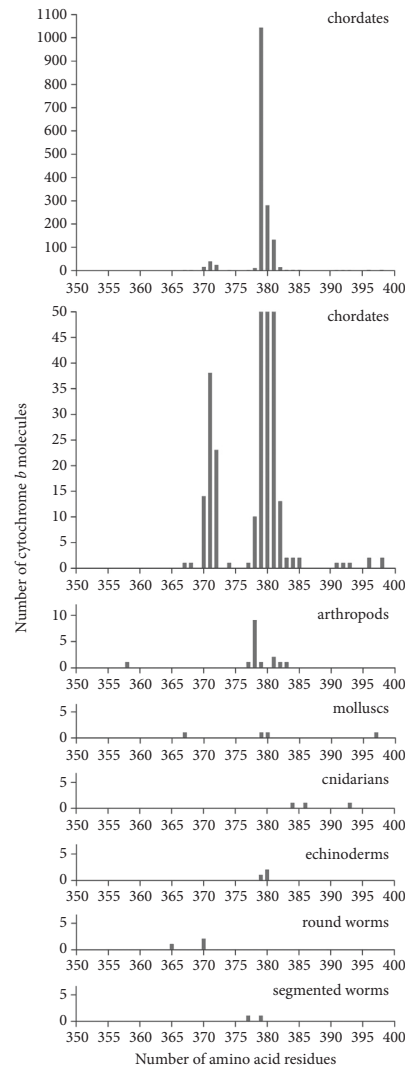


FIGURE 4: Distribution of the number of cytochrome b molecules, according to data from the Swiss-Prot database on the number of amino acid residues in various biological phyla. The data for chordates are presented at two scales of the ordinate—a truncated one, revealing small peaks throughout the entire region where lengths of amino acid residues exist, and a full one, where the entire maximum peak of the distribution is visible.

TABLE 4: Content of the number of cytochrome b amino acid sequences in the Swiss-Prot database in various taxonomic groups<sup>a</sup> of chordate (*Chordata*) phylum.

Biological phyla and other taxonomic groups of animals	Number of molecules	Number of amino acid residues
<i>Mammalia</i> (mammals)	1 249	377–384
<i>Aves</i> (birds)	134	379–381
<i>Serpentes</i> (snakes)	78	367–374
<i>Teleostei</i> (teleosts)	51	378–398
<i>Elasmobranchii</i> (elasmobranchs)	18	380, 381
<i>Testudines</i> (turtles)	11	379–381
<i>Amphibia</i> (amphibians)	9	380
<i>Actinopterygii</i> (ray-finned fishes)	8	379, 380
<i>Cyclostomata</i> (jawless)	4	385, 396
<i>Lepidosauria</i> (reptiles)	2	379, 380
<i>Cephalochordata</i> (lancelets)	2	367, 398
<i>Archosauria</i> (reptiles)	1	383
<i>Coelacanthiformes</i> (lobefin fishes)	1	380
<b>All Chordata</b>	<b>1 568</b>	<b>367–398</b>

<sup>a</sup>Due to the incomparability of different data on the division of chordates into biological classes [13–15], we took arbitrary taxonomic groups used in the Swiss-Prot database.

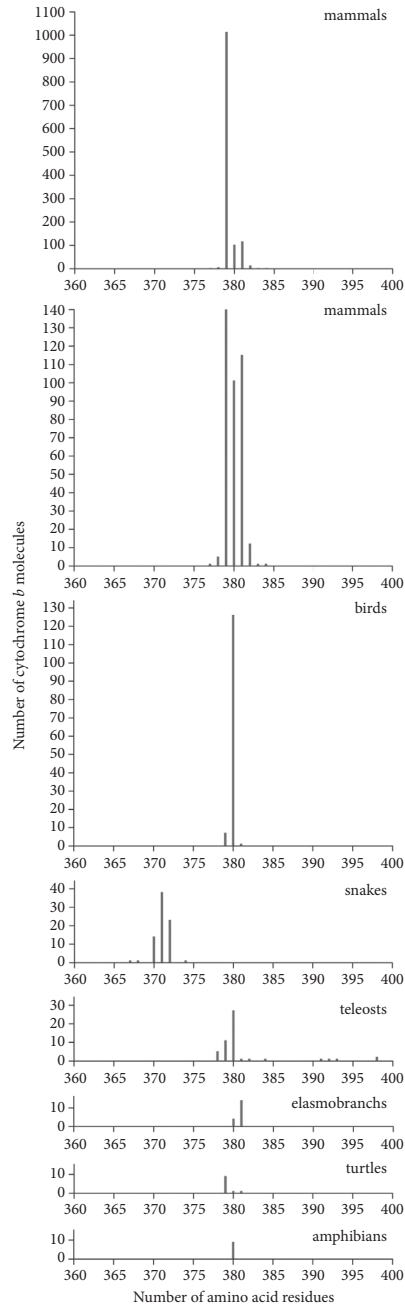


FIGURE 5: Distribution of the number of cytochrome b molecules, according to data from the Swiss-Prot database on the number of amino acid residues in various biological classes. The data for mammals are presented at two scales of the ordinate—a truncated one, revealing small peaks throughout the entire region where lengths of amino acid residues exist, and a full one, where the entire maximum peak of the distribution is visible.

echinoderms (starfish, *Patiria pectinifera*), annelids (common earthworm, *Lumbricus terrestris*), and primary tracheal worms (velvet worm, *Epiperipatus biolleyi*).

We found b cytochromes to be the most abundant in mammals (79.7%) among biological classes of animals such as chordates, and the number of molecules containing 379 amino acid residues (1013) in mammals was 81.1% (Table 4 and Figure 5). The interval of the observed numbers of amino acid residues for these is less than 40, and the boundary values of this interval coincide with the number of

residues in two lancelets: the common lancelet, *Branchiostoma lanceolatum*, and the Florida lancelet, *Branchiostoma florida*.

Analysis of the occurrence of b cytochromes containing 379 amino acid residues showed that most of them were found in representatives of rodents and artiodactyls (Table 5 and Figure 6). At the same time, representatives of bats, insectivores, and carnivores, as well as many orders represented by single molecules (Table 5), were always found to have only 379 residues.

TABLE 5: Content of the number of cytochrome b amino acid sequences in the Swiss-Prot database in various biological orders of mammal (*Mammalia*) order.

Biological orders of mammals	Number of molecules	Number of amino acid residues	Number of known species <sup>a</sup>
<i>Rodentia</i> (rodents)	381	377, 378, 379, 380, 381, 383, 384	2 552
<i>Artiodactyla</i> (aetiodactyls)	228	379, 380	551
<i>Chiroptera</i> (bats)	183	379	1 386
<i>Carnivora</i> (carnivorous)	118	379	305
<i>Eulipotyphla</i> (insectivorous)	97	379, 380	527
<i>Primates</i> (primates)	78	378, 379, 380	518
<i>Dasyuromorphia</i> (carnivorous marsupials)	59	381	78
<i>Lagomorpha</i> (lagomorphs)	44	379	98
<i>Diprotodontia</i> (two forward teeth marsupials)	12	379, 380, 381	155
<i>Didelphidae</i> (opossums)	12	382	111
<i>Perissodactyla</i> (odd-toed ungulates)	10	379	21
<i>Peramelemorphia</i> (marsupial omnivores)	5	381	5
<i>Caenolestidae</i> (rat opossum)	3	381	10
<i>Pilosa</i> (sloths and anteaters)	3	379	24
<i>Proboscidea</i> (elephants)	3	378	23
<i>Scandentia</i> (treeshrews)	3	379	3
<i>Monotremata</i> (monotremes)	2	379	7
<i>Afrosoricida</i> (tenrecs and golden moles)	1	379	55
<i>Cingulata</i> (armadillos)	1	379	20
<i>Dermoptera</i> (colugos)	1	379	2
<i>Hyracoidea</i> (hyraxes)	1	379	5
<i>Macroscelidea</i> (elephant shrews)	1	379	20
<i>Notoryctemorphia</i> (marsupial moles)	1	381	2
<i>Sirenia</i> (manatees and dugongs)	1	379	5
<i>Tubulidentata</i> (aardvarks)	1	379	1
<i>Microbiotheria</i> (monito del monte)	0	—	3
<i>Pholidota</i> (pangolins)	0	—	8
<b>All Mammalia</b>	<b>1 249</b>	<b>377–384</b>	<b>6 495</b>

<sup>a</sup>According to [14].

#### 4. Discussion

The data obtained indicate that mitochondrial cytochrome b, containing 379 amino acid residues and characterized by many different amino acid sequences, is found only in eukaryotes (100%), only in animals (100%) and mainly in mammals (79.7%). At the same time, the result of the analysis of different orders of mammals (Figure 6) does not allow us to assert that the representatives of one of the orders are characterized by the most frequent presence of this protein with this number of residues. According to the data given in Table 5, in most cases, the number of species with identified b cytochromes is less than the number of known species of this order. The only exceptions are the small orders *Peramelemorphia* (marsupial omnivores), *Scandentia* (tree shrews), and *Tubulidentata* (aardvarks). The ratio of the number of detected b cytochromes to the number of known biological species in this biological class (mammals) is also different in different orders. Therefore, the number of described b cytochromes may change significantly with further study of known, but not yet considered, mammals. Moreover, today, representatives of new biological species are constantly being discovered, including mammals, both fossils [16] and living mammals [17]. Therefore, we should expect continuing identification of more and more amino acid sequences of b cytochromes in newly discovered mammals and other living organisms. These data can change

both the ratio of the number of all cytochrome b molecules in different biological orders and the ratio of molecules containing 379 amino acid residues.

However, the data currently available indicate that representatives of mammals have b cytochromes containing only 379 amino acid residues in half of all mammalian orders (Table 5, Figure 6), and the total number of such representatives is more than 80% (1013 out of 1249) of all animal b cytochromes. Note also that human cytochrome b is composed not of 379 residues, but of 380 [18]. At the same time, the number of residues is 379 [19] and 381 [20] in such widely studied animal species as the bovine and the mouse, respectively.

It is well known that peptide molecules of the same length, but different amino acid sequences, can have the same functions [21–24]. For example, the pentapeptides met-enkephalin YGGFM and leu-enkephalin YGGFL are natural ligands of opioid receptors [25]. Therefore, the existence of a large number of b cytochromes with 379 amino acid residues is apparently not surprising. However, the length of the protein alone cannot help us understand the special property that determines its effective functioning. It is known that the functional property of cytochrome b is its ability to participate in the process of electron transfer [26], and the structural feature (amino acid sequence) is its special properties that allow it to be a transmembrane protein [27]. Consequently, it is obvious that the functional property of

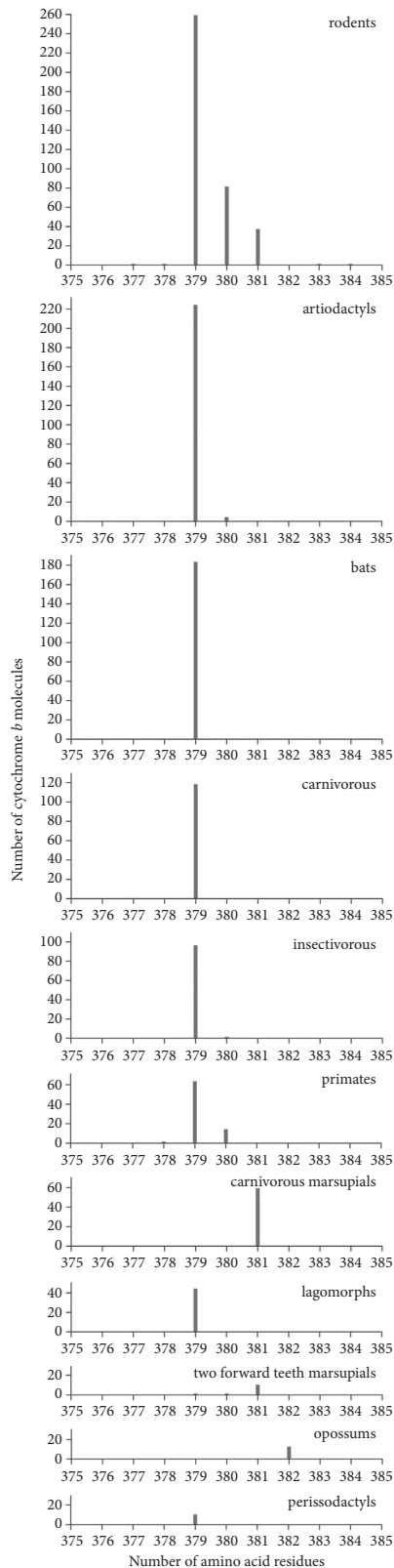


FIGURE 6: Distribution of the number of cytochrome b molecules, according to data from the Swiss-Prot database on the number of amino acid residues in various biological orders.

cytochrome b depends not only on the length of the protein but also on the features of the structure of its extra- and intramembrane fragments. Therefore, a detailed study of the physicochemical features of the cytochrome b polypeptide chain along its entire length is required, along with analysis of the features of the primary structure.

## 5. Conclusions

In conclusion, it should be noted that the b cytochromes of representatives of biological species of archaea and prokaryotes were studied significantly fewer in number than representatives of eukaryotes (Table 1); representatives of plants and fungi were fewer in number than representatives of animals (Table 2); representatives of chordates were significantly greater in number than representatives of other biological types (Table 3); and there are more representatives of mammals than representatives of other biological classes (Table 4). Therefore, further investigation, adding to the knowledge gleaned by this study in terms of the b cytochromes of yet-to-be-studied biological species, may change the general understanding of the occurrence of this protein in different taxonomic groups. It can also be assumed that a coupled study of the biological and physicochemical characteristics of this protein will explain the reason for the anomalous value of 379 in terms of amino acid residues contained in most of the already known b cytochromes, while the relatively simple methodological approach described in this study can be useful for further search and identification of novel patterns of proteins similarities.

## Data Availability

Data can be made available on request.

## Conflicts of Interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors thank Dr. Y. Chaikovskiy for fruitful discussions and comments on the manuscript and A.S. Borchikov and M.G. Vladimirov for technical assistance. Professor Alexander Zamyatin has passed away on December 11, 2020.

## References

- [1] S. Ramakumar, "Stochastic dynamics modeling of the protein sequence length distribution in genomes: implications for microbial evolution," *Physica A: Statistical Mechanics and its Application*, vol. 273, pp. 476–485, 1999.
- [2] G. Jhang, "Protein-length distributions for the three domains of life," *Trends in Genetics*, vol. 16, pp. 107–109, 2000.
- [3] A. Tiessen, P. Pérez-Rodríguez, and L. J. Delaye-Arredondo, "Mathematical modeling and comparison of protein size distribution in different plant, animal, fungal and microbial species reveals a negative correlation between protein size and



- protein number, thus providing insight into the evolution of proteomes," *BMC Research Notes*, vol. 5, no. 1, p. 85, 2012.
- [4] S. Eroglu, "Language-like behavior of protein length distribution in proteomes," *Complexity*, vol. 20, no. 2, pp. 12–21, 2014.
- [5] A. A. Zamyatnin and T. A. Belozerskaya, "Size of the natural linear peptide structures," *Russian Journal of Biology Physics Chemistry*, vol. 4, pp. 315–321, 2019.
- [6] A. A. Zamyatnin and T. A. Belozerskaya, "Multiple diversity of mitochondrial cytochrome b amino acid sequences of the same length in animals," *Frontiers in Molecular Biosciences*, vol. 7, p. 102, 2020.
- [7] A. Bairoch and B. Boeckmann, "The SWISS-PROT protein sequence data bank," *Nucleic Acids Research*, vol. 19, no. suppl, pp. 2247–2249, 1991.
- [8] A. Bairoch and R. Apweiler, "The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000," *Nucleic Acids Research*, vol. 28, no. 1, pp. 45–48, 2000.
- [9] R. Heilig, R. Eckenberg, J. L. Petit et al., "The DNA sequence and analysis of human chromosome 14," *Nature*, vol. 421, pp. 601–607, 1997.
- [10] J. Henry, P. Favrel, and E. Boucaud-Camou, "Isolation and identification of a novel Ala-Pro-Gly-Trp-amide-related peptide inhibiting the motility of the mature oviduct in the cuttlefish," *Sepia Officinalis Peptides*, vol. 18, no. 10, pp. 1469–1474, 1997.
- [11] D. M. Church, L. Goodstadt, L. W. Hillier, M. C. Zody, S. Goldstein, and X. She, "Lineage-specific biology revealed by a finished genome assembly of the mouse," *PLoS Biology*, vol. 7, Article ID e1000112, 2009.
- [12] C. R. Woese, O. Kandler, and M. L. Wheelis, "Towards a natural system of organisms: proposal for the domains archaea, bacteria, and eucarya," *Proceedings of the National Academy of Science USA*, vol. 87, pp. 4576–4579, 1990.
- [13] M. J. Benton, *Vertebrate Paleontology*, John Wiley & Sons, New York, NY, USA, 2014.
- [14] C. J. Burgin, J. P. Colella, P. L. Kahn, and H. S. Upham, "How many species of mammals are there?" *Journal of Mammalogy*, vol. 99, pp. 1–14, 2018.
- [15] P. R. Harding, "Scheme of general classification of the phylum chordata," *Journal of the Arkansas Academy of Science*, vol. 9, pp. 20–30, 1956.
- [16] S. N. Davis, C. R. Torres, G. M. Musser et al., "New mammalian and avian records from the late eocene la meseta and submeseta formations of Seymour Island, Antarctica," *PeerJ*, vol. 8, Article ID e8268, 2020.
- [17] K. M. Helgen, C. M. Pinto, R. Kays et al., "Taxonomic revision of the olingos (*Bassaricyon*), with description of a new species, the olinguito," *ZooKeys*, vol. 324, pp. 1–83, 2013.
- [18] S. Anderson, A. T. Bankier, B. G. Barrell et al., "Sequence and organization of the human mitochondrial genome," *Nature*, vol. 290, pp. 456–465, 1981.
- [19] S. Anderson, M. H. de Bruijn, A. R. Coulson, I. C. Eperon, F. Sanger, and I. G. Young, "Complete sequence of bovine mitochondrial DNA conserved features of the mammalian mitochondrial genome," *Journal of Molecular Biology*, vol. 156, pp. 683–717, 1982.
- [20] M. J. Bibb, R. A. van Etten, C. T. Wright, M. W. Walberg, and D. A. Clayton, "Sequence and gene organization of mouse mitochondrial DNA," *Cell*, vol. 26, pp. 167–180, 1981.
- [21] A. A. Zamyatnin, "Specificity of the amino acid residue content in endogenous regulatory oligopeptides," *Protein Sequence Data Analysis*, vol. 4, pp. 56–70, 1991.
- [22] A. A. Zamyatnin, "Structural classification of endogenous regulatory oligopeptides," *Protein Sequence Data Analysis*, vol. 4, pp. 53–56, 1991.
- [23] A. A. Zamyatnin, "Structural-functional diversity of the natural oligopeptides," *Progress in Biophysics and Molecular Biology*, vol. 133, pp. 1–8, 2018.
- [24] A. A. Zamyatnin, "Physicochemical and functional characteristics of complete system of natural oligopeptides," *Russian Journal of Biology Physics Chemistry*, vol. 3, pp. 225–235, 2018.
- [25] J. Hughes, T. W. Smith, H. W. Kosterlitz, L. A. Fothergill, B. A. Morgan, and H. R. Morris, "Identification of two related pentapeptides from the brain with potent opiate agonist activity," *Nature*, vol. 258, pp. 577–580, 1975.
- [26] W. J. Nawrocki, B. Balleul, B. Picot et al., "The mechanism of cyclic electron flow," *Biochimica et Biophysica Acta*, vol. 1860, pp. 433–438, 2019.
- [27] M. D. Esposti, S. de Vries, M. Crimi, A. Ghelli, T. Patarnello, and A. Meyer, "Mitochondrial cytochrome b: evolution and structure of the protein," *Biochimica et Biophysica Acta*, vol. 1143, pp. 243–271, 1993.