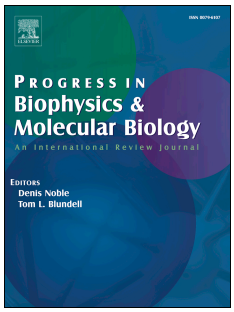
Accepted Manuscript



Structural–functional diversity of the natural oligopeptides

Alexander A. Zamyatnin

PII: S0079-6107(17)30186-4

DOI: 10.1016/j.pbiomolbio.2017.09.024

Reference: JPBM 1287

To appear in: *Progress in Biophysics and Molecular Biology*

Received Date: 29 July 2017

Revised Date: 27 September 2017

Accepted Date: 29 September 2017

Please cite this article as: Zamyatnin, A.A., Structural–functional diversity of the natural oligopeptides, *Progress in Biophysics and Molecular Biology* (2017), doi: 10.1016/j.pbiomolbio.2017.09.024.

Structural–Functional Diversity of the Natural Oligopeptides

Alexander A. Zamyatnin[a, b, *1]
[a]Bakh Institute of Biochemistry, Federal Research Centre "Fundamentals of Biotechnology", Russian Academy of Sciences, Moscow, Russia
[b]Departamento de Informática, Universidad Técnica Federico Santa Maria, Valparaiso, Chile

A B S T R A C T

Natural oligopeptides may regulate nearly all vital processes. To date, the chemical structures of many oligopeptides have been identified from >2000 organisms representing all the biological kingdoms. We have considered a number of mathematical (sequence length), chemical, physical, and biological features of an array of natural oligopeptides on the basis of the oligopeptide EROP-Moscow database (http://erop.inbi.ras.ru, 15,351 entries) data. There is the substantial difference of these substances from polypeptide molecules of proteins according to their physicochemical characteristics. These characteristics may be critical for understanding the molecular mechanisms of the action of oligopeptides that lead to the development of physiological effects.

Keywords: *oligopeptide; protein fragment; structure; functions; EROP-Moscow database; UniProt database*

Content

1. Introduction

Description of a large variety of natural nucleotide and amino acid sequences, which are the molecular basis of biological diversity of our planet, led to the creation of databases of these linear structures, which have been an object of numerous structural and functional studies. The first database was created in the United States by the initiative of Margaret Dayhoff; it was published as the "Atlas of Protein Sequence and Structure" in 1965 (Dayhoff, 1965; Strasser, 2010). This book contained approximately 100 pages with text that described approximately 70 protein amino acid sequences, predominantly cytochromes *c*, hemoglobins, and fibrinopeptides obtained from different sources. This database was printing until 1979. On the basis of its information, the integrated **P**rotein **I**dentification **R**esource (**PIR**) database was created in 1984 (current name is **P**rotein **I**nformation **R**esource).

Somewhat later (in 1986), in Switzerland, Amos Bairoch founded the European database SwissProt which, independently of the American database, was used for the same purposes (Bairoch and Boeckmann, 1991). Its basis was the data of deciphered primary structures of experimentally studied proteins. After 8 years, it was combined with data from the **TrEMBL**

*Corresponding author
E-mail address: aaz@inbi.ras.ru (A.Zamyatnin)