# BIOINFORMATIC ANALYSIS OF LIMITED PROTEOLYSIS ENZYMES OF HUMANS AND PULMONARY FRESHWATER MELLUSCS

**Semenov I. O, Chirkin A. A.**
Vitebsk State University named after P.M. Masherov,
Vitebsk, Belarus

Limited proteolysis is one of the options for post-translational modification of the protein, providing its functional activity. Although it is known that proteolytic enzymes are conservative in structure, it is of great scientific and practical interest to search for proteases closest to humans in different and accessible animal species. Currently, the main source of proteolytic enzymes for biopharmacy and the food industry are expensive and inaccessible marine aquatic organisms.

The aim of the work was to identify the homology of enzymes of limited proteolysis in humans and pulmonary freshwater mollusks.

Materials and methods. The nucleotide sequences of the following proteolytic enzymes were used as objects for homology detection of humans (Homo sapiens) and the mollusk Biomphalaria glabrata: Calpain 1 (EC 3.4.22.52), Calpain 2 (EC 3.4.22.53), Proteasome subunit beta type-6 (EC 3.4.25.1), Neprilysin 2 (EC 3.4.24.11). Hepsin (EC 3.4.21.106), Caspase 1 (EC 3.4.22.36).

Search for proteases was carried out on the server https://www.ebi.ac.uk/merops. The selection of the nucleotide sequences of human proteins was carried out in the database https://www.ensembl.org/index.html. The search for homologous sequences for mollusks was implemented on the server https: // www. ncbi.nlm.nih.gov using the BLAST resource. Using the resource https://www.uniprot.org, the reading frames were clarified when assessing the amino acid sequence during the alignment process. Pair alignment and comparison of sequences of proteolytic enzymes of humans and mollusks was performed in the MEGA 5.2 program. The search for 3D enzyme structures was carried out in the database http://www.rcsb.org. To build 3D models of enzymes, the resource https://swissmodel.expasy.org was used.

The work is based on the following algorithm: search for a suitable human enzyme in the MEROPS database → go to the ENSMBLE database to select a human nucleotide sequence → search for a homologous nucleotide sequence in an organism Biomphalaria glabrata using the BLAST resource → construction of amino acid sequences from nucleotide in the MEGA 5.2 program → verification of the reading frame by the amino acid sequence of a human enzyme taken from the UNIPROT database → pair alignment and assessment of the degree of homology of primary strains Tour → Search 3D structure of human proteases in the database ProteinDataBank → Building a 3D structure of the enzyme pattern of the human enzyme and its analysis using the life SWISS-MODEL [1].

Results and discussion. Table 1 presents a comparative bioinformatic analysis of six human proteolytic enzymes and the mollusk Biomphalaria glabrata, which is a related organism with the mollusk coil (Planorbarius corneus) widely represented in the lakes and rivers of Belarus. The first three enzymes presented in the table

(Calpain1, Calpain2, Caspase1) belong to the cysteine protease family; Neprilysin 2 - representative metalloprotease; Hepsin is a serine protease; Proteasome subunit beta type-6 belongs to the threonine protease family. Quantitative data of the bioinformatic analysis are presented in table 1.

***Table 1.*** – Evaluation of the homology of the primary structures of the molecules of proteolytic enzymes of man and mollusk

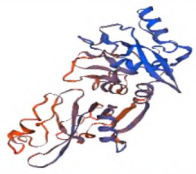
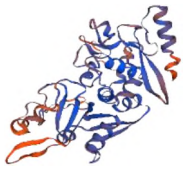
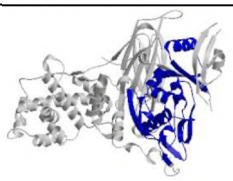
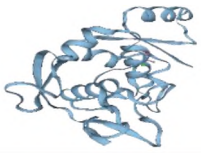
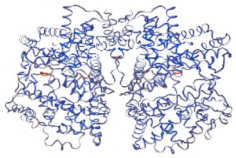
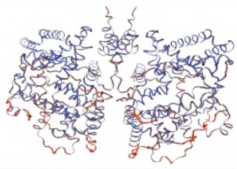
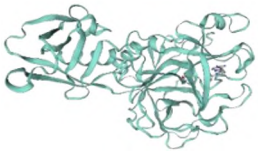
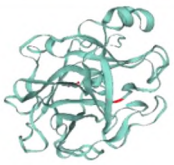
| Enzyme | Type of sequence | Expected value | Query coverage | Identities |
|---|---|---|---|---|
| Calpain 1 | NS | 4e-08 | 9% | 65.29% |
| EC 3.4.22.52 | AAS | 0.0 | 94% | 45.02% |
| Calpain 2 | NS | 7e-18 | 15% | 64.52% |
| EC 3.4.22.53 | AAS | 0.0 | 96% | 44.99% |
| Caspase 1 | NS | 9.1 | 4% | 70.42% |
| EC 3.4.22.36 | AAS | 8e-16 | 45% | 28.65% |
| Neprilysin 2 | NS | 1e-147 | 74% | 40.41% |
| EC 3.4.24.11 | AAS | 1e-147 | 74% | 40.41% |
| Hepsin | NS | 4.2 | 3% | 73.97% |
| EC 3.4.21.106 | AAS | 2e-39 | 57% | 37.35% |
| Proteasome subunit | NS | 2e-55 | 51% | 62.10% |
| beta type-6 | AAS | 6e-47 | 51% | 54.84% |
| EC 3.4.25.1 | | | | |

When pairing the nucleotide sequences of these human and mollusk enzymes, it was found that the active sites (Activesite), binding sites (Bindingsite) and metal bonds (Metalbinding) for 4 enzymes were completely homologous in humans and mollusks (Calpain 1, Caspase 1, Hepsin, Proteasome subunit beta type-6); 2 enzymes are partially homologous (Calpain 2, Neprilysin 2). The homology percentage for Caspase 1 (70.42%) and Hepsin (73.97%) cannot be considered reliable due to the low coverage of the sequence (Query coverage: 4% and 3%, respectively) and the rather high value of random alignment (Expected value: 9.1 and 4.2, respectively).

Table 2 presents a comparative bioinformatic analysis of spatial models of human proteolytic enzymes and the mollusk Biomphalaria glabrata. Due to the lack of sequences of suitable length, reliable spatial structures of the Caspase1 and Proteasome subunit beta type-6 enzymes could not be constructed; the QMEAN factor for these sequences was less -5.0. To solve the problem, sequencing of mollusk DNA is necessary and the inclusion of more complex analytical apparatuses in the construction of spatial structures.

Conclusion. As a result of the studies, it was shown that the enzymes of limited proteolysis of the mollusk have a high degree of homology with human enzymes. This allows us to consider it appropriate to further consider this model organism for the use of its proteolytic enzymes in the food industry and biopharmaceuticals.

***Table 2.*** – 3D structures of cellular proteolytic enzymes of humans and mollusks

| Фермент | Характеристика | | *Homo sapiens* | *Biomphalaria glabrata* |
|---|---|---|---|---|
| Calpain1 | GMQE<br>QMEAN<br>Identity | 0,34<br>-2,86<br>51,91 |  |  |
| Calpain 2 (каталитическая субъединица) | GMQE<br>QMEAN<br>Identity | 0,79<br>-1,76<br>62,13 |  |  |
| Neprilysin 2 | GMQE<br>QMEAN<br>Identity | 0,75<br>-2,44<br>38,66 |  |  |
| Hepsin (каталитический домен) | GMQE<br>QMEAN<br>Identity | 0,72<br>-1,44<br>40,63 |  |  |

Note: GMQE – global model quality assessment; QMEAN is a composite assessment based on various geometric properties and provides both global and local absolute quality estimates based on one model; Identity - homology, identity.

## References


**1.**Chirkin A.A., Dolmatova V.V. Comparative analysis of proteolytic enzymes of human and pulmonary freshwater molluscs // Agr. bio. div. Impr. Nut., Health Life Qual. - 2018. - P. 234-242. https://doi.org/10.15414/agrobiodiversity. 2018.2585-8246.234-242


## DISTANCE AND ON-LINE LEARNING AS CHALLENGE TO THE MODERN WORLD


**Siusiuka V., Sergienko M., Kolokot N., Shevchenko A., Roslik O.**
State Medical University of Zaporizhzhia,
Department of Obstetrics and Gynaecology,
Zaporizhzhia, Ukraine


Distance forms of work were used in State Medical University of Zaporizhzhia for many years; mainly it was related to upgrading of teacher's skills. Owing to the quarantine within the territory of Ukraine in order to prevent distribution of acute