Development and application of the complex model of the cholinergic system

Volodymyr Vasylenko[1],

Mariia Martynova[2],

Saulius Mickevičius[1]

*[1] Vytautas Magnus University,
Faculty of Natural Sciences,
Universiteto St. 10-314,
Kaunas district, 53361, Lithuania*

*[2] University of West Bohemia,
Faculty of Applied Sciences,
Plzeň, Technická 8, 30614,
Czech Republic*

The vital activity of organisms is regulated by biochemical processes. Even minor disruptions or incorrect intervention in these processes leads to serious, often irreversible, consequences. The study of the cholinergic system and its modulators is of particular interest. Reasonable influence on the metabolism of acetylcholine can help in solving a wide range of veterinary, medical, and fundamental biochemical problems. Currently, not all mechanisms of biological effects on the cholinergic system have been studied. Limited laboratory methods are the main obstacle to further study of these processes. The results of *in vitro* and *in vivo* studies can most often be questioned due to the significant influence of external factors, high variability of variables, and difficulties in processing large volumes of biochemical data. In recent years, more and more researchers have emphasized the need to introduce new methods and research methods in the area of biological sciences. Without innovative approaches in applied research, it is no longer possible to obtain a reliable and satisfactory result. It is necessary to set a goal to find and develop safe modulators of the cholinergic system both in the direction of catalytic reactions and its inhibition by using the methods of computational biology. Existing software products give positive results in solving highly specific issues, but they do not offer a comprehensive solution to the existing problems of biomodelling. This article examines examples of the existing modelling techniques and considers the possibility of creating a universal software product for the integral construction and analysis of the cholinergic system.

Keywords: bioinformatics, cholinergic system, computational biology, intelligent data analysis, machine learning, modelling

INTRODUCTION

The cholinergic system (ChS) is one of the most important structural units of neurohumoral regulation and a disturbance of its work leads to a potentially lethal state of a living organism. ChS is a complex system of neural interconnection of the basal forebrain and mainly the limbic and near-limbic regions (Cummings, 2000). In addition, the processes involving the acetylcholine (ACh) transmitter are the basis for most of motor and sensory reactions in the peripheral nervous system. A disruption of the processes associated

* Corresponding author. Email: volodymyr.vasylenko@vdu.lt

with acetylcholine leads to the development of neurodegenerative diseases such as myasthenia gravis (Appel et al., 1979) or Alzheimer's disease (Ferreira-Vieira et al., 2016).

Biological regulation of the level and activity of acetylcholine is carried out by the enzyme cholinesterase (ChE), mainly acetylcholinesterase (AChE) and pseudoacetylcholinesterase (P-AChE). The role of buterylcholinesterase (BuChE) is usually neglected due to its insignificant amount, although this is most likely due to insufficient knowledge of the processes in which the latter takes part. In veterinary and human medicine, activators (agonists) of ACh receptors, such as carbachol and antagonists, such as trimetaphan or surugatoxin, are used as ChS modulators (Wiffen et al., 2017).

Unfortunately, the use of various kinds of ChS modulators is a reserve method due to their narrow therapeutic index and the lack of specific therapy methods for intoxication. Bronchospasm, paradoxical muscle weakness, including the diaphragm, as well as bradycardia, up to asystole, most often become potentially dangerous effects of such therapy (Neely et al., 2020; Pakala et al., 2020). An important fact is that today it is not possible to move away from the use of such compounds due to the lack of an alternative. It is also necessary to consider the toxicity in the case of poisoning with ACh modulators in the form of, for example, pesticides or potent poisonous gases, which are not the subject of this article.

It is important to pay more academic and applied attention to the search for biologically active molecules with high selectivity for cholinergic processes and with a minimum of pleiotropic effects. This can be achieved through the application of methods of computational biology. This requires software development that considers all aspects of the ChS model. Only *in silico* methods can create a sufficiently realistic model of cholinergic processes suitable for studying its modulation.

In 2019, a group of researchers at Stanford University concluded that the lack of the ability to create deep models that mimic experimental analysis limits research methods in the natural

sciences. Ramsundar et al. noted in 2019 that the existing means of bio- and chemoinformatics are able to fully simulate the process of an experiment in a very limited number of cases and use a narrow range of variable data parameters.

Modern methods of scientific research in the biological field are less and less consistent with applied problems for many reasons. As science advances, the main reason is the increasing complexity of the required experiments. In other words, if a century and a half ago it was enough to examine the internal structure of a cell in a microscope or to decompose the light of the flame of certain chemical compounds into a spectrum, today we set ourselves the task of understanding complex and ambiguous processes, such as changing the genome or predicting the chemical properties of non-existent compounds. Today, we cannot rely solely on observation due to the limited capabilities of the observer. Naturally, modern research methods and their development should be based on the available empirical data, but their further development is impossible without the active introduction of information technologies. Data hybridization, in turn, requires multidisciplinary knowledge from a specialist. The next-generation researcher should be comfortable both in the laboratory and at the computer.

If we consider the currently used classical research methods (*in vivo* and *in vitro*), several critical shortcomings can be noted. In the case of *in vitro* tests, this is the primitiveness of the model and an analytical error in calculations. Regarding *in vivo* testing, the sensitivity of the methods is of deep concern (high variance leads to unreliable results). It is equally important to consider the complexity of creating a sample (model) corresponding to the given parameters, and the high cost, which in some cases reaches millions of dollars.

It should be recognized that the existing *in silico* methods, which have significant objectivity and adaptability, are certainly capable of processing large amounts of data and are distinguished by a high level of globalization (as a result of bringing data from different sources into a single digital format).

The above approaches do not have such properties necessary for ChS studies as multi-vector and reproducibility of discrete processes and do not have sufficient interconnection to form conclusions based on a combination of their results. Biochemical studies require the modelling of dynamic processes considering the composition, homeostasis, side processes; they need multitasking and deep systematization of the data obtained. Further development of these sciences is possible only after the introduction of information technologies into the active use by researchers.

MATERIALS AND METHODS

To confirm the promise of computational methods in biology, an analysis of scientometric data such as Web of Science, Elsevier, and Google Scholar was carried out based on software solutions for simulating the software requirements of the cholinergic system without laboratory testing. The usable options of proposed software were determined.

RESULTS AND DISCUSSION

As a result of laboratory modelling of ChS, it is possible to obtain a result with high reliability and significant selectivity to the process under study. This imposes a significant limitation on the number of system components due to the need to maintain a narrow range of physicochemical properties of the system. There is no possibility of accurate modelling of the required order of reactions, taking into account the affinity of chemical compounds. In addition, it is impossible to reproduce biological responses differentially, in real scale, for example, within one receptor. Thus, an important and, possibly, the main problem is the impossibility of considering all, even the known parameters of a given system. First of all, this is due to the need to process large amounts of data. As a clarification, we can cite the constancy of the internal conditions of biological systems. Homeostasis is a complex dynamic equilibrium system for

biological processes. The dynamism of homeostasis speaks of initiating the processes of counteraction and maintaining a stable state of the system as a response to any change. It is not possible to achieve such control of the reaction environment in laboratory models, especially under the condition of microreactions, which is associated with the need to instantly respond to changes.

The tasks described in the paper constitute a simple thesis of the development of a digital laboratory with long-term prospects for implementation in virtual reality. Attempts to digitalize biological research have been under way since the mid-1970s. The results of such work are successful and are of great practical importance, but in the context of large-scale tasks such as modeling ChS, these solutions are mono-vector, and in the case of parallel processes, they are significantly noisy. This leads to the loss of a significant part of the data and systematic errors in the results.

According to the analysis of literature data, in particular the resource-aggregator Web of Science, an accelerating growth of interest (as a result of the number of publications) in bioinformatics is observed. The idea of informatization of chemistry was for the first time formulated back in 1926 (Lotka, 1926), and the first mention of the terms bioinformatics and chemoinformatics refers to 1997 (Lim, 1997).

The number of studies focused on chemical models is incomparable with the interest in general biological solutions (Willett, 2020). A significant increase in interest in bioinformatics has been observed since 2015 (Fig. 1), which confirms the importance of information methods for the global scientific community. Bioinformatics methods are used in many areas of the life sciences and are funded by many organizations (Tables 1, 2). Asia and Asian universities take the first place in terms of the number of studies in complex information spheres (Fig. 2).

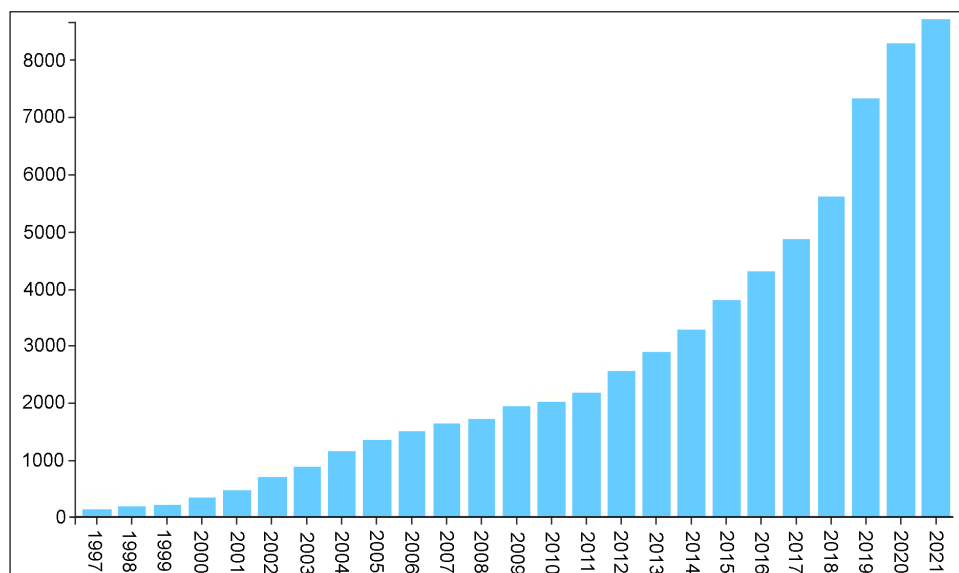Currently, there is a huge amount of information in bioinformation databases, such as

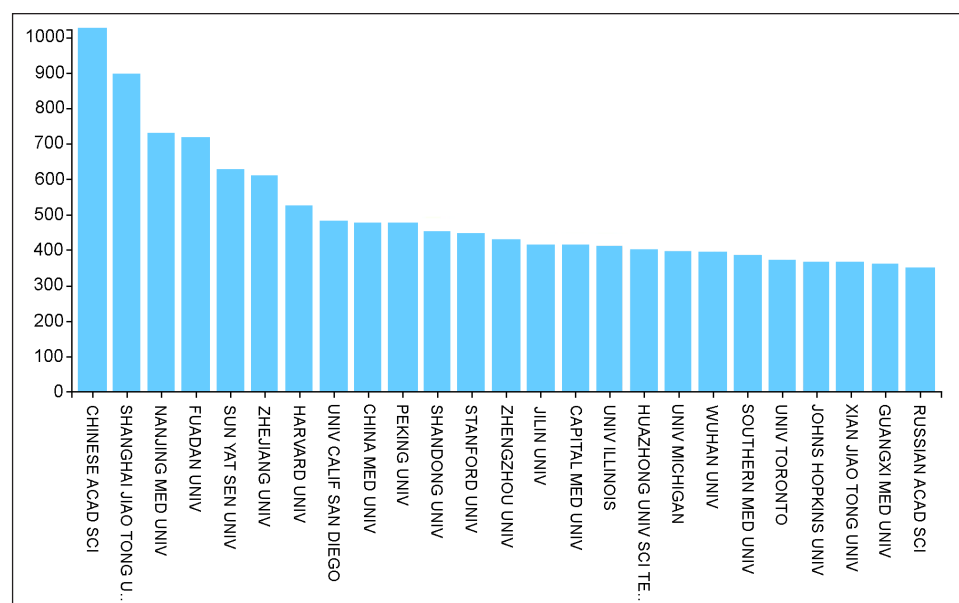**Fig. 1.** Number of bioinformatics publications by year

Fig. 2. Number of bioinformatics publications by university

Table 1. **Output statistics by directions**

Biochemistry molecular biology	10050
Biochemical research methods	7236
Biotechnology applied microbiology	7141
Mathematical computational biology	5984
Oncology	5807
Genetics heredity	5545
Cell biology	4160
Computer science interdisciplinary applications	4074
Multidisciplinary sciences	3876
Medicine research experimental	3774
Computer science theory methods	2602

Table 1. (**Continued**)

Pharmacology / Pharmacy	2506
Multidisciplinary science Artificial Intelligence	2399
Statistics probability	2270
Computer science Information Systems	2123
Microbiology	2074
Engineering electrical electronics	1814
Plant sciences	1676
Biophysics	1541
Immunology	1535
Biology	1455
Chemistry Multidisciplinary	1328
Neuroscience	1062
Engineering biomedical	937
Pathology	800

Table 2. **Output statistics by financing institutions**

National Natural Science Foundation of China NSFC	9837
United States Department of Health Human Services	7385
National Institutes of Health NIH USA	7352
European Commission	2898
NIH National Institute of General Medical Sciences NIGMS	2069
National Science Foundation NSF	1742
NIH National Cancer Institute NCI	1606
UK Research Innovation UKRI	1323
NIH National Institute of Allergy Infectious Diseases NIAID	1015
National Basic Research Program of China	890
Ministry of Education Culture Sports Science and Technology Japan MEXT	835
German Research Foundation DFG	718
NIH National Heart Lung Blood Institute NHLBI	699
Fundamental Research Funds for the Central Universities	680
Biotechnology and Biological Sciences Research Council BBSRC	672
Japan Society for the Promotion of Science	671
NIH National Human Genome Research Institute NHGRI	647
Medical Research Council UK MRC	621
National Council for Scientific and Technological Development CNPQ	612
China Postdoctoral Science Foundation	590
Grants in Aid for Scientific Research Kakenhi	583
NIH National Center for Research Resources NCRR	579
Wellcome Trust	553
NIH National Institute of Diabetes Digestive Kidney Diseases NIDDK	545
National High Technology Research and Development Program of China	507

the National Center for Biotechnology Information European Nucleotide Archive (Luscombe et al., 2001). It is worth noting that even such vast and popular databases have significant drawbacks and data gaps. If one pays attention to the structure of these bases, one will notice that they are all based on the structure of objects without considering their physical and chemical properties. Information about such exists, but it is not a key for classification, because it is used only as an addition. Thus, the organization and understanding of biological data in bioinformatics are generalized only in two dimensions, depth and width (Luscombe et al., 2001). Diniz and Canduri (2017) describe the striving for the maximum understanding of biochemical properties in terms of structure. In the general case, biological information is structured according to the quantitative and geometric principle, almost completely ignoring the chemical properties and the influence of accompanying factors. Significant progress can be achieved only through the search for new analytical strategies. The leap in the evolution of deep learning networks in the last ten years opens great prospects for learning systems to solve more and more complex problems (Chen et al., 2018). The introduction of new methods will make it possible to predict a wide range of chemical and biological phenomena more accurately (Lo et al., 2018).

Currently, there are many functional highly specialized implemented solutions in the fields of bioinformatics (Walsh et al., 2017; Chojnacki et al., 2017). At the same time, the term "solutions" refers to products of various profiles: from the unified bioinformatics bases mentioned above to software implementations of individual methods or models (Lo et al., 2018; Tangadpalliwar et al., 2019). Nevertheless, the field of problems in this young area is still vast, and further pace of development depends on our ability to solve these problems.

First of all, specialists are faced with the problem of describing biological data and translating them into the digital format (Ramsundar et al., 2019). In other words, disagreements and discrepancies arise as early as the stage of computer coding. A good illustration of such problems is the digitization of the records and judgments of specialists (which also raises the problem of subjectivization). The majority of the problematic aspects are embedded in the data translation techniques themselves, for example, the well-known and widely used formats Simplified Molecular-Input Line-Entry System (SMILES) and Extended Connectivity Fingerprints (ESFP). The method of describing molecules using SMILES text lines requires, in most cases, translation into other formats for employing the data using machine learning methods (and sometimes even simple analytical programs), which entails the need to introduce a process of molecular characterization. Currently, there is an RDKit package, which is a collection of functions for working with this format, which allows partial automation of this process.

In turn, ESFP is a more universal format: its lines are the same size, which is a critical property for many modern bioinformatics methods. The properties of molecules are described on two levels, starting with the characteristics of the properties of each present atom and ending with the descriptions of the bonds of each particular atom with other atoms in the molecule. This format is more convenient for the informatics side, but it carries with it the initial loss of some specific information, for example, two different molecules may have the same fingerprints.

Also, one cannot ignore the general problems of this type of data: large volumes, incompleteness, noisiness, inaccuracy (Cook et al., 2019). All these preprocessing errors then introduce a systematic nature of errors into the data processing itself.

The trend towards the interdisciplinary interaction of life sciences and informatics has led to the formation of a list of classical problems of bioinformatics, which are characteristic only for this area, which, in turn, led to the emergence of numerous specialized solution methods (Guo et al., 2018).

Currently, there are several software products for solving these problems, but their

main common feature is narrow specialization (mono-vector) and the lack of the possibility of combining methods and modelling their dynamic interaction for complex analysis or modelling of the task (Saldívar-González et al., 2020; Chojnacki et al., 2017), which is critically important when the question of assessing the impact and changes is raised not in a limited field, but in a complex manner within the whole system. These solutions give usable results; however, to draw truly complete and reliable conclusions, a lot of additional actions and resources are needed. This can be either random laboratory tests for additional confirmation of a hypothesis or a secondary assessment of research results using other software products. In addition, dynamic processes are not considered at all (Youjun, Jianfeng, 2017; Sharma, Sharma, 2018), and currently system modelling affinity or chirality have a number of significant shortcomings.

Thus, considering all the properties of homeostasis and system components, full-fledged modelling of such a complex system as ChS is impossible.

It is proposed to create a software product in the form of a complex of modules that comprise a single working environment capable of simulating the cholinergic system, considering dynamic processes taking into account all direct and indirect factors, predicting possible interactions, assessing the effectiveness of the process and the quality of the result, and modelling processes at the micro-level and macrosystems (in other words, capable of changing the discreteness).

The interface should be intuitive and user-friendly, be comfortable in working with specific information in the field (presentation of information and data in text, tabular, and graphical versions), have some universal base of implemented techniques and a wide range of settings for typical procedures, as well as the ability to create custom processes and algorithms for solving highly specialized tasks.

An important aspect of the work of the proposed product is a powerful mechanism for preprocessing the incoming data, which will be able to increase the quality of raw data to the required quality and level of trust (it is possible to consider other characteristics, such as anonymization, randomization, systematization, etc.); if necessary, to generate additional data samples based on a small sample or increase/decrease the amount of noise. Also, an urgent task is to solve the problem of a full-fledged, unified, and complete coding format for specific information in the fields of bio- and chemoinformatics, which would be supported by all methods and algorithms included in the solution, exclude the loss or the possibility of ignoring important special information, and at the same time would imply the possibility of convenient visualization for the researcher.

The possibility of observing and diversified monitoring of dynamic processes in real time can be an important and necessary property for solving a whole range of problems in the region.

The results of the ChS simulation will be automatically generated reports that can be easily interpreted by specialists. It is understood that such results should include tables with the data necessary for the formation of conclusions, graphical interactive objects, and statistical analysis and visualization of its results, if necessary (for example, finding an unobvious correlation).

The development process of the proposed software solution can be divided into three stages according to the number of functional parts of the product:

Preprocessing. Development of a modular complex product for data preprocessing – coding of specific area data, their features, improving their quality, creating the necessary initial samples. It does not exclude the emergence of a new format for coding biological and chemical data.

Processing. The most important and long-term stage of development consisting of the selection of existing implementations that meet the requirements for the final product, revision and unification of the set of existing solutions, and the creation of author's modules for yet unsolved problems.

Analytics. The final module, which includes analytical, statistical, and visual methods for producing results. The output of this module will be an automatic or almost automatically-generated report on the complex experiment carried out.

In the future, this type of a software product can become a full-fledged virtual analogue of a chemical (biological) laboratory.

CONCLUSIONS

The use of existing modulators of the cholinergic system has significant limitations. Modern methods of the search for innovative biologically active molecules and study into the pleotropic effects of the existing ones are either unable to provide the required level of preparatory and primary data or are not optimal, that is, costly and ineffective in practical application.

Digital products require wider implementation in the practice of academic and applied research as they can simplify testing, significantly reduce costs, and significantly increase their efficiency. The disadvantages of the existing software tools can be eliminated by increasing the number of inputs and using self-learning neural networks. This article lists the necessary requirements for a fundamentally new software product that uses unsupervised teaching methods. Criteria for input parameters and simulation results are indicated. Based on these data, it is assumed that a test version of such a system will be developed. After the implementation of the proposed solution for modelling the cholinergic system, it will become possible to study dynamic multiphase and multicomponent processes in real time. During the simulation, it will be possible to adjust the set parameters and assess the influence of the external environment (or homeostasis) and side processes on the investigated person and the system in general.

The development of such a software solution leads to the following obvious and indisputable advantages in the study of the cholinergic system and the search for its modulators:

a generalized fully-featured solution covering all the methodological base necessary for a researcher; partial or full automation of routine processes; an increase in the productive volume of process modelling; increasing the quality of analytics and objectifying conclusions; unification of the complete life cycle of process modelling; optimization of the assessment of the prospects of a hypothesis; reducing the cost of research and minimizing financial losses in cases of failure to prove a hypothesis.

It is necessary to pay much more attention to information methods in the life sciences. The cyberization of scientific experiments and the assessment of hypotheses will lead to an increase in the effectiveness of research and a decrease in potential risks and financial losses.

References

1. Appel SH, Elias SB, Chauvin P. The role of acetylcholine receptor antibodies in myasthenia gravis. Fed Proc. 1979; 38(10): 2381–5.

2. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. Drug Discov Today. 2018; 23(6): 1241–50.

3. Chojnacki S, Cowley A, Lee L, Foix A, Lopez R. Programmatic access to bioinformatics tools from EMBL-EBI update: 2017. Nucleic Acids Res. 2017; 45: W550–3.

4. Cook C, Lopez R, Stroe O, Cochrane G, Brooksbank C, Birney E, Apweiler R. The European Bioinformatics Institute in 2018: tools, infrastructure and training. Nucleic Acids Res. 2019; 47: D15–22.

5. Cummings JL. Cholinesterase inhibitors: A new class of psychotropic compounds. Am J Psychiatry. 2000; 157(1): 4–15.

6. Diniz WJ, Canduri F. Bioinformatics: an overview and its applications. Genet Mol Res. 2017; 16(1): 10.4238/gmr16019645.

7. Ferreira-Vieira TH, Guimaraes IM, Silva FR, Ribeiro FM. Alzheimer's disease: targeting the cholinergic system. Curr Neuropharmacol. 2016; 14(1): 101–15.

8. Guo R, Zhao Y, Zou Q, Xiaodong F, Shaoliang P. Bioinformatics applications on Apache Spark. GigaScience. 2018; 7(8); giy098.

9. Lim HA. Bioinformatics and cheminformatics in the drug discovery cycle. In: Hofestädt R, Lengauer T, Löffler M, Schomburg D, editors. Bioinformatics. GCB 1996. Lecture Notes in Computer Science. Berlin: Springer; 1997. p. 30–43.

10. Lo YC, Rensi SE, Torng W, Altman RB. Machine learning in chemoinformatics and drug discovery. Drug Discov Today. 2018; 23: 1538–46.

11. Lotka AJ. The frequency distribution of scientific productivity. J Wash Acad Sci. 1926;16: 317–23.

12. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. Methods Inf Med. 2001; 40: 346–58.

13. Neely GA, Sabir S, Kohli A. Neostigmine. Treasure Island (FL): StatPearls Publishing; 2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470596/

14. Pakala RS, Brown KN, Preuss CV. Cholinergic Medications. Treasure Island (FL): StatPearls Publishing; 2020. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538163/

15. Ramsundar B, Eastman P, Walters P, Pande V. Deep learning for the life sciences: applying deep learning to genomics, microscopy, drug discovery, and more. Sebastopol: O'Reilly Media, Inc.; 2019.

16. Sharma S, Sharma D. Intelligently applying artificial intelligence in chemoinformatics. Curr Top Med Chem. 2018; 18: 1804–26.

17. Saldívar-González FI, Huerta-García CS, Medina-Franco JS. Chemoinformatics-based enumeration of chemical libraries: a tutorial. J Cheminform. 2020; 12: 64.

18. Tangadpalliwar SR, Vishwakarma S, Nimbalkar N, Garg P. ChemSuite: A package for chemoinformatics calculations and machine learning. Chem Biol Drug. 2019; 93(5): 960–4.

19. Walsh I, O'Flaherty R, Rudd PM. Bioinformatics applications to aid high-throughput glycan profiling. Perspectiv Sci. 2017; l11: 31–9.

20. Wiffen P, Mitchell M, Snelling M, Stoner N, editors. Oxford handbook of clinical pharmacy. 3rd ed. Oxford: Oxford University Press; 2017.

21. Willett P. The literature of chemoinformatics: 1978–2018. Int J Mol Sci. 2020; 21(15): 5576.

22. Youjun XU, Jianfeng PEI. Deep learning for chemoinformatics. Big Data Res. 2017; 3(2): 2017019.

Volodymyr Vasylenko, Mariia Martynova, Saulius Mickevičius

CHOLINERGINĖS SISTEMOS KOMPLEKSINIO MODELIO RAIDA IR TAIKYMAS

Santrauka

Organizmų gyvybinę veiklą reguliuoja biocheminiai procesai. Net nedideli pažeidimai ar netinkamas įsikišimas į šiuos procesus sukelia rimtų, dažnai negrįžtamų, pasekmių. Ypač įdomus cholinerginės sistemos ir jos moduliatorių tyrimas. Pagrįsta įtaka acetilcholino metabolizmui gali padėti išspręsti įvairias veterinarijos, medicinos ir pagrindines biochemijos problemas. Šiuo metu nėra ištirti visi biologinio poveikio cholinerginei sistemai mechanizmai. Pagrindinė kliūtis tiriant šiuos procesus yra riboti laboratoriniai metodai. Tyrimų *in vitro* ir *in vivo* rezultatais dažnai verčia suabejoti reikšminga išorinių veiksnių įtaka, didelis kintamųjų kintamumas ir sunkumai apdorojant didelį biocheminių duomenų kiekį. Pastaraisiais metais vis daugiau tyrinėtojų pabrėžė būtinybę biologijos mokslų srityje diegti naujus tyrimų metodus. Be novatoriškų taikomųjų tyrimų metodų neįmanoma gauti patikimo ir patenkinamo rezultato. Naudojant skaičiavimo biologijos metodus, reikia sukurti saugius cholinerginės sistemos moduliatorius tiek katalizinių reakcijų, tiek jų slopinimo kryptimi. Esami programinės įrangos produktai duoda teigiamus rezultatus sprendžiant labai specifinius klausimus, tačiau jie nepasiūlo išsamaus biomodeliavimo problemų sprendimo. Šiame straipsnyje nagrinėjami esami modeliavimo būdai ir svarstoma galimybė sukurti universalų programinės įrangos produktą, skirtą integraliai cholinerginės sistemos statybai ir analizei.

Raktažodžiai: bioinformatika, cholinerginė sistema, skaičiavimo biologija, intelektuali duomenų analizė, mašininis mokymasis, modeliavimas