ISSN 1993-6842 (on-line); ISSN 0233-7657 (print)
Biopolymers and Cell. 2022. Vol. 38. N 3. P 186–194
doi: http://dx.doi.org/10.7124/bc.000A7A

UDC 579.61

Identification of potential novel membrane drug targets of *Acinetobacter baumannii* ATCC 19606 using subtractive proteomics approach

M. V. Shmatkov[1], G. P. Volynets[2,3], V. G. Bdzhola[2], V. I. Matiushok[2,3], O. M. Chubukov[2], S.M. Yarmoluk[2]

[1] Institute of High Technologies, Taras Shevchenko National University of Kyiv
2, korp.5, Pr. Akademika Hlushkova, Kyiv, Ukraine, 03022

[2] Institute of Molecular Biology and Genetics, NAS of Ukraine
150, Akademika Zabolotnoho Str., Kyiv, Ukraine, 03143

[3] LLC "Scientific and service firm "Otava"
150, Akademika Zabolotnoho Str., Kyiv, Ukraine, 03143
g.p.volynets@gmail.com

Aim. To identify the potential novel membrane drug targets of *Acinetobacter baumannii* ATCC 19606. **Methods.** Clustering of paralogues was performed by USEARCH software, the identification of essential non-homologous proteins to the human proteome was done with BLASTp and Database of essential genes, determination of proteins from unique metabolic pathways was carried out using KAAS server at KEGG. The drug target novelty was estimated with DrugBank. The sub-cellular localization of the proteins was predicted with PSORTb v. 3.0.3, CELLO v. 2.5 and BUSCA. Tertiary structures of proteins were built with trRosetta and 3D models quality was analyzed using MolProbity server. The potential binding sites were predicted with PrankWeb, BIOVIA Discovery studio 2021 visualizer and Caver analyst 2.0. **Results.** Six potential novel membrane drug targets were identified within the *Acinetobacter baumannii* ATCC 19606 proteome such as rod shape-determining protein RodA, DedA family protein, undecaprenyl-diphosphate phosphatase, putative lipid II flippase FtsW, prolipoprotein diacylglyceryl transferase, apolipoprotein N-acyltransferase. Tertiary structures of the proteins were built and ligand-binding sites were predicted. **Conclusions.** The identified potential novel membrane-associated drug targets of *Acinetobacter baumannii* ATCC 19606 can be useful for further drug development in order to find novel treatments of the infectious diseases caused by *Acinetobacter baumannii*.

Keywords: *Acinetobacter baumannii*, subtractive proteomics, drug targets, membrane proteins.

© 2022 M. V. Shmatkov *et al.*; Published by the Institute of Molecular Biology and Genetics, NAS of Ukraine on behalf of Biopolymers and Cell. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited

Introduction

Acinetobacter baumannii is a gram-negative bacterium that is an opportunistic pathogen and affects people with suppressed immunity, causing pneumonia, meningitis, infections of urinary systems and other diseases. Due to its ability to acquire rapidly multidrug resistance, this pathogen is one of the main causes of hospital-acquired infections and can infect patients through surgical wounds, catheters, mechanical ventilators, *etc.* [1]. The multi-drug resistance of *A. baumannii* can be caused by a number of molecular mechanisms such as expression of enzymes that can destroy antibiotics, increased density of efflux pumps on the bacterial surface which can remove the drugs from the cells, modulation of membrane permeability, *etc.* [2, 3]. Therefore, the discovery of novel types of antibiotics toward *A. baumannii* is of great importance.

The development of new antimicrobial drugs for the treatment of infectious diseases is an expensive and time-consuming process. There are several approaches for the antibiotics discovery. One of them is a phenotypic high throughput screening of the compound collection of the antibacterial activity *in vitro*. However, due to its high cost, slowness and low productivity, this method is impractical for use in its pure form. On the other hand, there are *in silico* approaches for the prediction of interaction of small molecules with macromolecules. These methods are cheap and fast, but they can give only approximate data and are not accurate. Therefore, a rational way to find new antibiotics is to combine both methods: *in silico* screening of a large collection of compounds to predict potentially active compounds against a specific target and testing these focused libraries *in vitro*. Therefore, it is important to find appropriate molecular targets for the antibiotic development. The subtractive proteomics approach, which was used in our study, allows quick identification of the novel potential drug targets with given properties [4].

Materials and Methods

Proteome retrieving

Complete proteome of *A. baumannii* ATCC 19606 was downloaded from NCBI genomes database [5].

Clustering paralogues

To get [the] paralogues from *A. baumannii* ATCC 19606 proteome we used USEARCH software [6]. UCLUST algorithm allows quick grouping of the proteins with high identity into clusters. Sequence identity cut-off was set to 0.6. We kept only one sequence from each cluster and other proteins were removed.

Identification of the sequences non-homologous to the human proteome

We performed homology analysis of *A. baumannii* proteins in order to remove the sequences similar to the human host proteins. This step allows reduction of the risk of unwanted interactions between the drug and human proteins, and therefore preventing of a number of side effects. For this analysis we performed BLASTp with E-value = 10^{-3} against human proteome [7, 8]. The sequences possessing homology to human proteins were removed.

Search for essential genes

For the drug development, the essentiality of target protein for the bacterial cell living and functioning is very crucial. To detect such proteins, we submitted our sequences from BLASTp against DEG (Database of Essential Genes) [9]. This database contains the essential genes and proteins of different species, including *A. baumannii* ATCC 17978, which has 85.535 % genome symmetric identity to analyzed strain *A. baumannii* ATCC 19606. There are two sets of proteins for *A. baumannii* ATCC 17978 in the database: the first one is for rich medium, another — for murine model of pneumonia, excluding the proteins from rich medium. We submitted BLASTp sequences against both sets and added the results. The E-value was set to 10^{-5}.

Metabolic pathways analysis

The essential non-homologous sequences were submitted to KAAS server at Kyoto Encyclopedia of Genes and Genomes [10]. We used the single-directional best hit assignment method. Each protein, that was predicted to take part in any metabolic pathway, was given a KO-code and put into BRITE hierarchy. We compared the received metabolic pathways with the human metabolic pathways and kept the sequences, which had KO-number and were involved into the unique metabolic pathways only. This step is important for further reduction of the potential side effects.

Detection of already known drug targets

The sequences obtained in the previous step were analyzed for the presence of already known drug targets. We submitted the sequences to the DrugBank database with BLASTp built-in [11]. E-value was set to 10^{-5}. The proteins that showed significant similarity to any known target were removed.

Prediction of the subcellular localization

To predict subcellular localization of obtained sequences we used web services PSORTb v3.0.3 [12], CELLO v.2.5 [13] and BUSCA [14]. The specific localization for the sequence was proposed only in a case if all three web-services obtained the same result. We kept only the proteins with the predicted membrane localization.

3D-structure modelling and models validation

To obtain three-dimensional structures of the proteins we used trRosetta web-server, which combines neural network prediction and homology modeling approaches [15–17]. The obtained models were validated by MolProbity web-server [18].

Binding sites detection and analysis

We used BIOVIA Discovery studio 2021 visualizer, Caver analyst 2.0 software and the PrankWeb service to find and analyze potential ligand-binding sites [19, 20]. The proteins with the binding sites predicted simultaneously by three programs were proposed as the potential drug targets.

Results and Discussion

To identify novel membrane-associated potential drug targets in *Acinetobacter baumannii* ATCC 19606 we used the subtractive proteomics approach. At the first step we have downloaded full proteome of *A. baumannii* ATCC 19606 in FASTA format from the NCBI

genome database. This proteome was submitted by Sydney University and consists of 3680 proteins. We performed the clustering of paralogues using USEARCH software with sequence identity cut-off set to 0.6 and obtained 3586 clusters. The most of them had only one sequence in each cluster. At the next step we performed BLASTp of 3586 sequences against the human proteome and determined that 2542 proteins are non-homologous to human proteins with the E-value set to 10^{-3}. The set of the non-homologous non-redundant sequences was submitted to BLASTp against two lists of essential proteins of *Acinetobacter baumannii* ATCC 17978. 311 proteins were determined as homologous to essential proteins for the rich medium sample and 147 were determined as homologous to the essential proteins for the murine model of pneumonia [21].

Since some of the sequences were duplicated, the total number of unique sequences was 438. The set of the essential non-homologous non-redundant sequences was submitted to KAAS server for the prediction of their involvement in metabolic pathways. KAAS did not predict the metabolic pathways for 105 sequences, which were removed. It was found that 148 proteins belong only to bacterial unique metabolic pathways. Most of these proteins belong to lipopolysaccharide biosynthesis pathway, peptidoglycan biosynthesis, two-component system, quorum sensing, biofilm formation and beta-lactam resistance. Among 148 essential non-homologous non-redundant proteins from unique metabolic pathways, using DrugBank database with BLASTp built-in, we determined 77 sequences that did not possess similarities to any of the known targets with reported drugs.

For these proteins we predicted subcellular localization using the PSORTbv3.0.3, CELLOv.2.5 and BUSCA web-services. It was found that among them 29 were cytoplasmic, 21 — membrane-associated, 1 — periplasmic, and for 26 proteins the localization in the cell was not clearly defined.

Because the aim of this work was to identify novel membrane-associated proteins, we selected 21 membrane proteins for further analysis. We built three-dimensional structures of these proteins using trRosetta server. TM-score was higher than 0.5 for all models, which means that the prediction on the proteins structure topology was correct. In order to ensure the quality of the models, we validated them with MolProbity server. According to the MolProbity score requirements for good quality model, the score has to be ≥ 66 %. In our case, all models had the score above 85 %, which confirms their high quality.

Then, we used PrankWeb server to predict the potential ligand binding sites for these proteins. Each cavity was analyzed for the amount of amino acid residues, evolutionary conservation and the probability that this cavity is a real binding site. At this stage, we discarded ten proteins that according to the data of PrankWeb server do not possess binding sites. The remaining models we analyzed with BIOVIA Discovery studio 2021 visualizer and Caver analyst 2.0 programs; according to the results of these programs we removed five other proteins. The results of subtractive proteomics analysis are summarized in the Table 1. Therefore, we obtained six proteins that can be potential novel membrane-associated drug targets, such as apolipoprotein N-acyltransferase, DedA family protein, puta-

Table 1. **Number of the proteins at each step of subtractive proteomics analysis**

Step of subtractive proteomics analysis	Number of proteins
Full proteome	3680
The proteins after clustering paralogues	3586
Non-homologous to human proteome sequences	2542
Essential proteins	438
Proteins involved in unique metabolic pathways	148
Proteins without known drugs	77
Membrane proteins	21
Potential novel drug targets	6

tive lipid II flippase FtsW, prolipoprotein diacylglyceryl transferase, rod shape-determining protein RodA and undecaprenyl-diphosphate phosphatase (Table 2).

We have built three-dimensional structures of the proteins using web-server trRosetta, which combines neural network prediction and homology modeling approaches. 3D structures of these proteins with highlighted ligand-binding pockets are presented in Fig. 1.

Apolipoprotein N-acyltransferase takes part in the third step of the post-translational lipid modification. This enzyme catalyzes the addition of palmitate to the N terminus of diacylated apolipoproteins [22].

The DedA family proteins are highly conserved membrane proteins, but they are poorly characterized and their functions still remain unknown. However, there are some hypotheses that the DedA family proteins can function as the proton-dependent transporters, provide colistin resistance and take part in the normal cell division [23, 24].

FtsW is responsible for the transport of Lipid II across the bacterial inner membrane [25].

Prolipoprotein diacylglyceryl transferase catalyzes the first reaction of the three-step post-translational lipid modification for lipoprotein biosynthesis. It has two entrances to

Table 2. **Potential novel membrane drug targets of *Acinetobacter baumannii* ATCC 19606, their functions and NCBI protein accession numbers**

Protein	Function	NCBI protein accession number
Apolipoprotein N-acyltransferase	Lipoprotein biosynthesis	WP_001203085.1
DedA family protein	Unknown	WP_000421341.1
Putative lipid II flippase FtsW	Lipid II transport	WP_000907680.1
Prolipoprotein diacylglyceryl transferase	Cell wall synthesis	WP_000959084.1
Rod shape-determining protein RodA	Cell wall synthesis	WP_000075224.1
Undecaprenyl-diphosphate phosphatase	Lipoprotein biosynthesis	WP_000426931.1

the giant central cavity: front entrance at the periplasmic space and side entrance that opens to the cytoplasm [26].

The Rod shape-determining protein RodA is a peptidoglycan polymerase that plays crucial role in the cell growth and determines its shape [27]. The binding site of the protein is located at the extracellular side and is well conserved between the species [28].

Undecaprenyl-diphosphate phosphatase converts undecaprenyl-diphosphate, which is a product of peptidoglycan synthesis, to undecaprenyl-phosphate, that can be used again for the bacterial cell wall construction. The binding site of the protein is accessed from the periplasm [29].

All identified proteins are crucial to the membrane and cell wall synthesis and functioning. Since it is easier for drug to interact with the proteins, the active sites of which are open into the periplasmic space, these targets can be considered as promising. Additionally, the effect of drugs that act on the membrane proteins does not depend on the presence or absence of efflux pumps in the bacterium, because the drug does not need to penetrate into the cytoplasm.

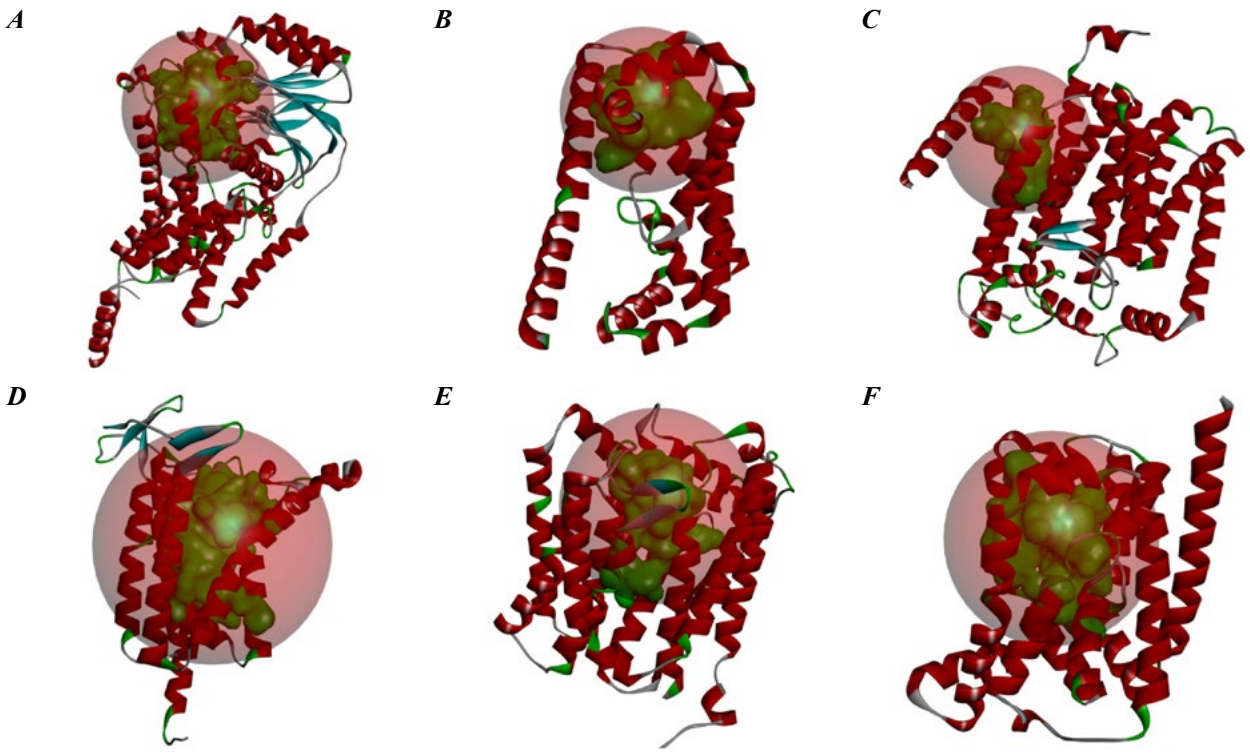

Fig. 1. Three-dimensional (3D) structures of *Acinetobacter baumannii* ATCC 19606 apolipoprotein N-acyltransferase (a), DedA family protein (b), putative lipid II flippase FtsW (c), prolipoprotein diacylglyceryl transferase (d), rod shape-determining protein RodA (e) and undecaprenyl-diphosphate phosphatase (f) with indicated ligand-binding pockets, predicted with BIOVIA Discovery studio 2021 visualizer.

Conclusion

Using the subtractive proteomics approach we identified six novel potential drug targets of *Acinetobacter baumannii* ATCC 19606, which are essential, non-homologous, non-redundant membrane proteins from unique metabolic pathways. Their three-dimensional structures were predicted and analyzed for the small-molecule ligand binding sites. We hope that the results of this study will help in the development of new antimicrobial drugs toward *Acinetobacter baumannii.*

REFERENCES

1. *Antunes LCS, Visca P, Towner KJ.* Acinetobacter baumannii: evolution of a global pathogen. *Pathog Dis.*, 2014; **71**(3):292–301.
2. *Long Q, Huang C, Liao P, Xie J.* Proteomic insights into Acinetobacter baumannii drug resistance and pathogenesis. *Crit Rev Eukaryot Gene Expr.*, 2013; **23**(3):227–55.
3. *Garnacho-Montero J, Amaya-Villar R.* Multiresistant Acinetobacter baumannii infections: epidemiology and management. *Curr Opin Infect Dis.*, 2010; **23**(4):332–39.
4. *Barh D, Tiwari S, Jain N, Ali A, Santon AR, Misra AN, Azevedo V, Kumar A.* In silico subtractive genomics for target identification in human bacterial pathogens. *Drug Dev Res.* 2011; **72**(2):162–77.
5. *Hamidian M, Blasco L, Tillman LN, To J, Tomas M, Myers GSA.* Analysis of complete genome sequence of Acinetobacter baumannii strain ATCC 19606 reveals novel mobile genetic elements and novel prophage. *Microorganisms.* 2020; **8**(12):1851.
6. *Edgar RC.* Search and clustering orders of magnitude faster than BLAST. *Bioinformatics.* 2010; **26**(19):2460–1.
7. *Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.* BLAST+: architecture and applications. *BMC Bioinformatics.* 2009; **10**:421.
8. *Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.* Basic local alignment search tool. *J Mol Biol.* 1990; **215**(3):403–10.
9. *Luo H, Lin Y, Gao F, Zhang CT, Zhang R.* DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. *Nucleic Acids Res.* 2013; **42**:D574–80.
10. *Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M.* KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 2007; **35**:W182–5.
11. *Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M.* DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* 2018; **46**:D1074–82.
12. *Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FSL.* PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics.* 2010; **26**(13):1608–15.
13. *Yu CS, Chen YC, Lu CH, Hwang JK.* Prediction of protein subcellular localization. *Proteins.* 2006; **64**(3):643–51.
14. *Savojardo C, Martelli PL, Fariselli P, Profiti G, Casadio R.* BUSCA: an integrative web server to predict subcellular localization of proteins. *Nucleic Acids Res.* 2018; **46**:W459–66.
15. *Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S, Baker D.* Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci.* 2020; **117**(3):1496–503.
16. *Su H, Wang W, Du Z, Peng Z, Gao SH, Cheng MM, Yang J.* Improved protein structure prediction using a new multi-scale network and homologous templates. *Adv Sci (Weihn).* 2021; **8**(24):e2102592.
17. *Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, Anishchenko I, Baker D, Yang J.* The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc.* 2021; **16**(12):5634–51.

18. *Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Videau LL, Deis LN, Verma V, Keedy DA, Hintze BJ, Chen VB, Jain S, Lewis SM, Arendall WB 3rd, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* 2018; **27**(1):293–315.
19. *Jendele L, Krivak R, Skoda P, Novotny M, Hoksza D.* PrankWeb: a web server for ligand binding site prediction and visualization. *Nucleic Acids Res.* 2019; **47**:W345–49.
20. *Jurcik A, Bednar D, Byska J, Marques SM, Furmanova K, Daiel L, Kokkonen P, Brezovsky J, Strnad O, Stourac J, Pavelka A, Manak M, Damborsky J, Kozlikova B.* CAVER Analyst 2.*0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories.* Bioinformatics. 2018; **34**(20):3586–88.
21. *Wang N, Ozer EA, Mandel MJ, Hauser AR.* Genome-wide identification of Acinetobacter baumannii genes necessary for persistence in the lung. *mBio.* 2014; **5**(3):e01163–14.
22. *Noland CL, Kattke MD, Diao J, Gloor SL, Pantua H, Reichelt M, Katakam AK, Yan D, Kang J, Zilberleyb I, Xu M, Kapadia SB, Murray JM.* Structural insights into lipoprotein N-acylation by Escherichia coli apolipoprotein N-acyltransferase. *Proc Natl Acad Sci U.*S.A. 2017;**114**(30):E6044–53.
23. *Tiwari V, Panta PR, Billiot CE, Douglass MV, Herrera CM, Trent MS, Doerrler WT.* A Klebsiella pneumoniae DedA family membrane protein is required for colistin resistance and for virulence in wax moth larvae. *Sci Rep.* 2021; **11**(1):24365.
24. *Panta PR, Kumar S, Stafford CF, Billiot CE, Douglass MV, Trent MS, Doerrler WT.* A DedA family membrane protein is required for Burkholderia thailandensis colistin resistance. *Front Microbiol.* 2019; **10**:2532.
25. *Mohammadi T, van Dam V, Sijbrandi R, Vernet T, Zapun A, Bouhss A, Diepeveen-de Bruin M, Nguyen-Distèche M, de Kruijff B, Breukink E.* Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J.* 2011; **30**(8):1425–32.
26. *Mao G, Zhao Y, Kang X, Li Z, Zhang Y, Wang X, Sun F, Sankaran K, Zhang XC.* Crystal structure of *E. coli* lipo-protein diacylglyceryl transferase. *Nat Commun.* 2016; **7**:10198.
27. *Emami K, Guyet A, Kawai Y, Devi J, Wu LJ, Allenby N, Daniel RA, Errington J.* RodA as the missing glycosyl-transferase in Bacillus subtilis and antibiotic discovery for the peptidoglycan polymerase pathway. *Nat Microbiol.* 2017; **2**:16253.
28. *Sjodt M, Brock K, Dobihal G, Rohs PDA, Green AG, Hopf TA, Meeske AJ, Srisuknimit V, Kahne D, Walker S, Marks DS, Bernhardt TG, Rudner DZ, Kruse AC.* Structure of the peptidoglycan polymerase RodA resolved by evolutionary coupling analysis. *Nature.* 2018; **556**(7699):118–21.
29. *El Ghachi M, Howe N, Huang CY, Olieric V, Warshamanage R, Touzé T, Weichert D, Stansfeld PJ, Wang M, Kerff F, Caffrey M.* Crystal structure of undecaprenyl-pyrophosphate phosphatase and its role in peptidoglycan biosynthesis. *Nat Commun.* 2018; **9**(1):1078.

Ідентифікація нових потенційних мембранних мішеней для лікарських препаратів проти *Acinetobacter baumannii* ATCC 19606 за допомогою підходу субтрактивної протеоміки

М. В. Шматков, Г. П. Волинець, В. Г. Бджола, В. І. Матюшок, О. М. Чубуков, С. М. Ярмолюк

Мета. Визначити нові потенційні мішені для лікарських препаратів проти *Acinetobacter baumannii* ATCC 19606. **Методи.** Кластеризацію для визначення паралогів було здійснено за допомогою програмного забезпечення USEARCH, визначення протеїнів, що не гомологічні до людських було проведено з використанням веб-серверів BLASTp та Database of essential genes, ідентифікацію білків, що залучені в унікальні для бактерії метаболічні шляхи здійснювали з використанням серверу KAAS від KEGG, новизну мішеней визначали по відношенню до бази даних DrugBank, субклітинну локалізацію білків передбачали за допомогою веб-ресурсів PSORTb v. 3.0.3, CELLO v. 2.5 та BUSCA. Просторову структуру білків будували з використанням веб-серверу trRosetta, якість моделей перевіряли за допомогою веб-серверу MolProbity.

Ідентифікацію ліганд-зв'язувальних сайтів проводили за допомогою інтернет ресурсу PrankWeb, програмного забезпечення BIOVIA Discovery studio 2021 visualizer та Caver analyst 2.0. **Результати.** Серед протеїнів *Acinetobacter baumannii* ATCC 19606 було запропоновано шість нових потенційних мембранних мішеней для розробки потенційних лікарських препаратів, такі як аполіпопротеїн-N-ацилтрансфераза, протеїн з родини DedA, передбачувана фліппаза ліпід II FtsW, проліпопротеїн диацилгліцерил трансфераза, пептидоглікан полімераза RodA, ундекапреніл-дифосфат фосфатаза. Третинні структури цих протеїнів були побудовані та передбачені ліганд-зв'язуючі сайти. **Висновки.** За допомогою підходу субтрактивної протеоміки було ідентифікувано шість нових потенційних мембранних мішеней для розробки потенційних лікарських препаратів щодо *Acinetobacter baumannii* ATCC 19606.

Ключові слова: *Acinetobacter baumannii*; субтрактивна протеоміка; мішені для лікарських засобів; мембранні білки.

Received 08.03.2022