ISSN: 0975-766X
CODEN: IJPTFI
Available Online through
Research Article
www.ijptonline.com

PREDICTION OF THE THREE-DIMENSIONAL STRUCTURE OF THE PROTEIN SAHPF AND ANALYSIS OF ITS MOLECULAR DYNAMICS

R.Kh. Ayupov, N.I. Akberova
Kazan Federal University, Institute of Fundamental Medicine and Biology,Kazan, 18 Kremlyovskaya street.
Email: aurusta@mail.ru

Received on 05-06-2016 *Accepted on 27-06-2016*

Abstract:

SaHPF is the protein of gram-positive bacterium *Staphylococcus aureus*, which is causesa variety of diseases (pneumonia, meningitis, endocarditis, etc.), including nosocomial infections. SaHPF ishibernation-promoting factor, presumably interacts with the 30S subunit the ribosome and alters its conformation, and this results binding together of two ribosomes. Such ribosomal dimers do not perform protein synthesis, and this allows the cell to survive the unfavorable environmental conditions.The understanding of SaHPF-ribosome interaction mechanisms will allow to develop new antibacterial drugs.

Protein structure was predicted using Robetta, Quark, I-Tasser, SWISS-MODEl and Phyre2 software, which used methods: homology, threading and *ab initio*. In total 24 protein models were built. Quality assessment of the obtained protein structures models was performed by Qmean program. Among the predicted structures, the most qualitative assessment was obtained model of the protein built in Robetta program, which builds the models by combining the methods of homology and *ab initio*. This model of the protein was analyzed by equilibrium molecular dynamics method in NAMD program, using Charmmforce field. The analysis of molecular dynamics trajectories using principal component and normal mode methods revealed a special mobility of the loop and the C-terminal domain of the protein, which may complicate the resolution of SaHPF structure by experimental methods (crystallography and NMR).

Keywords: SaHPF, *Staphylococcus aureus*, prediction structure,Robetta, *in silico*, molecular dynamics.

1. Introduction

SaHPF - *S.aureus*hibernation promoting factor is the protein involved in the process of *S.aureus* ribosome dimerization. SaHPF consists of 190 amino acid residues [1].

In paper [2] ribosomal particles with 100S sedimentation constant were founded in Gram-positive bacteria *S.aureus*. Using 2D electrophoresis they analyzed proteins associated with a ribosome. As a result it was revealed that in ribosome dimers this protein is bound with70S in a molar ratio of 1:1. With anincrease of salt concentration the elution of protein from the complex with a ribosome occurred, leading to 100S decay, an inverse addition of protein resulted in dimer formation. These results confirm the key role of protein in the formation of *S.aureus* ribosome dimers. The formation of dimeric ribosome particle allow a cell to survive adverse environmental conditions, stopping energy-consuming processes polypeptide chain synthesis of translation. *S.aureus*is Gram-positive bacteria and the cause of many diseases [3,4]. Understanding the mechanisms of interaction of protein with the ribosome will allow to develop new drugs of antibacterial action against ribosomes [5].

Methods of protein structure prediction *in silico* enable to obtain the primary data on protein domains and to compare structural features of the of homologous proteins models [6]. Available programs of spatial proteins model prediction have their peculiarities. Most of them at the first stage of spatial model development use primary sequence of proteins homology method presented in the data base of PDB structures [7].I-Tasser builds the structures of proteins by homology method of primary and secondary protein structures [8]. Robettabuild three-dimensional structures byof the methods of homology and *ab initio* [9].

QUARK uses only *ab initio* methodfor the construction of protein models [10]. The construction of protein models by *ab initio* method is time and resource consuming process. Therefore, at the first stage Robettabuild the areas of proteins, which has homologous to known structures, and by *ab initio* method completionunknown part of the model. QUARK is limited by the length of protein primary sequence. The program build three-dimensional models for proteins at the length of no more than 200 amino acid residues.

The quality of protein model construction by different programs is estimated by CASP system [11].

NAMD is a popular program used for molecular dynamics performance [12]. The force fields Amber [13] and Charmm [14] are used for the physical and chemical description of molecules [14]. Molecular dynamics trajectories are analyzed by various scripts. Also the analysis may be performed in VMD program [15] and in the statistical package bio3D within R environment [16, 17]. Using molecular dynamics one may describe the structural features of protein, the mobility of its domains, and analyze its interaction with various ligands [18,19].

The aim of this study is the prediction of SaHPF protein structure 3d model and the analysis of its domains mobility by equilibrium molecular dynamics.

2. Methods

The 3D model of the protein was predicted using the following programs: I-Tasser [8], Robetta [9], QUARK [10], Phyre2 [20] and Swiss-Model [21]. The evaluation of built protein models was carried out using Qmean program [22]. The equilibrium molecular dynamics of predicted model was carried out in NAMD 2.8 program using force field Charmm 36.

The preparation of files for molecular dynamics was carried out in VMD program. Protein structure was placed in the water box, the system charge was neutralized by the addition of NaCl ions. The water box dimensions were 68 × 84 × 80 Å. For the first stage the energy of protein-water-ions system was minimized during 40000 steps. Then the system was heated to 300 °K with the increment of 1° per 100 steps of dynamics. After heating the equilibration of kinetic and potential energy of system was performedfor 200,000 steps. The equilibrium molecular dynamics of prepared system was performed for 20 ns (20,000,000 steps) at 300 °K. Molecular dynamics analysis was performed using VMD software and the statistical package bio3D in R environment, the mobility of protein domains was analyzed using on-line servers AD-ENM [23] and Flex Serv [24].

3. Results and Discussion

24 SaHPF models were developed using the programs of protein structure prediction (Table 1). The following proteins were used bysoftware as homologous structures: 1L4S [25], 1IMU [26], 3LYV [27], 1N3G [28], 3KA5 [29], 3V26 [30]. It was found that SaHPF protein consists of two domains, and significant differences in the predicted models have been associated with the location of these domains in relation to each other. Table 1 shows the evaluations of constructed model quality obtained during structure analysis in Qmean program. The results of the program operation are evaluated according to several criteria.

Figure 1 shows the position of the predicted 3D protein model relative to the evaluation of experimentally allowed structures from databases (PDB). The error probability of each amino acid residue position in the predicted structure is shown on Figure 2.

The protein consists of two domains and a hinge, which connects them. The first domain (N-terminal) is presented by 100 amino acid residues, a poorly structured hinge consists of of 40 amino acid residues, the second domain (C-terminal) is composed of 50 amino acid residues. The positions of the amino acid residue atoms presented in β-layers and α-helix are determined most clearly, while the hinge and loops first and second domensof the protein atoms have a large enough position uncertainty.

Table 1. The assessment of Qmean quality for SaHPF structure predicted by different programs.

№	Program	Model	Total QMEAN-score	Z-score	Completenessofbuildingstructures	Homologousmodel
1.	Swiss-model	1	0.613	-1.19	101 (3-103)	1L4S
2.		2	0.565	-1.62	102 (2-103)	1N3G
3.		3	0.608	-1.22	102 (2-103)	1IMU
4.	Robetta	1	0.561	-2.12	190	1IMU 3LYV
5.		2	0.553	-2.20		
6.		3	0.632	-1.38		
7.		4	0.594	-1.77		
8.		5	0.566	-2.06		
9.	Quark	1	0.513	-2.62	190	- (only method *ab initio*)
10.		2	0.535	-2.39		
11.		3	0.399	-3.80		
12.		4	0.545	-2.28		
13.		5	0.580	-1.92		
14.		6	0.537	-2.36		
15.		7	0.540	-2.34		
16.		8	0.585	-1.87		
17.		9	0.529	-2.45		
18.		10	0.544	-2.30		
19.	Phyre2	1	0.652	-0.30	57 (132-188)	3KA5
20.	I-Tasser	1	0.544	-2.29	190	1N3G 3KA5 3V26
21.		2	0.479	-2.97		
22.		3	0.561	-2.12		
23.		4	0.484	-2.91		
24.		5	0.497	-2.77		

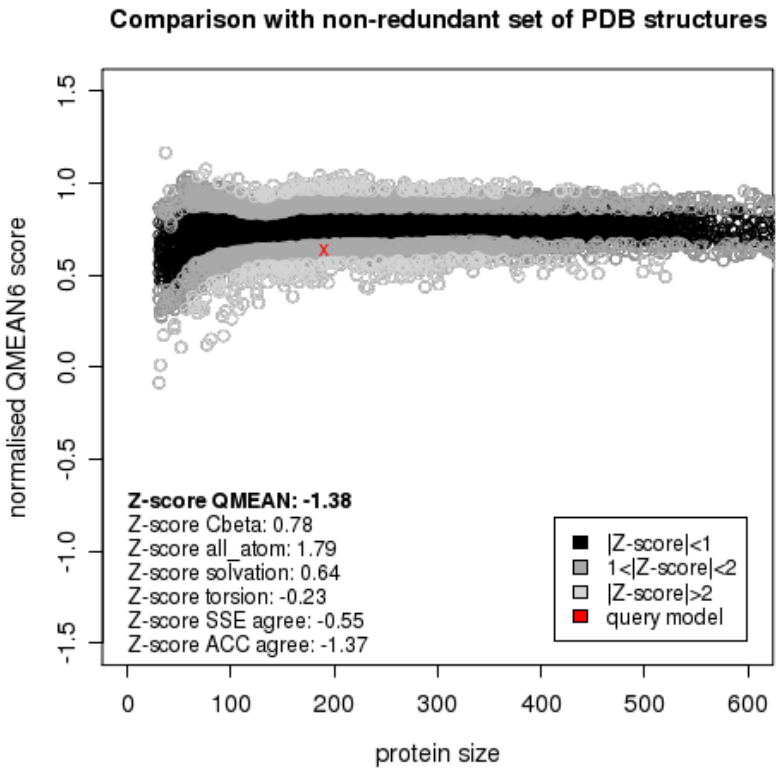

Fig.1 Qmeanquality evaluation for the predicted SaHPF protein model 3 constructed in Robetta program. The model estimation is denoted by «X».

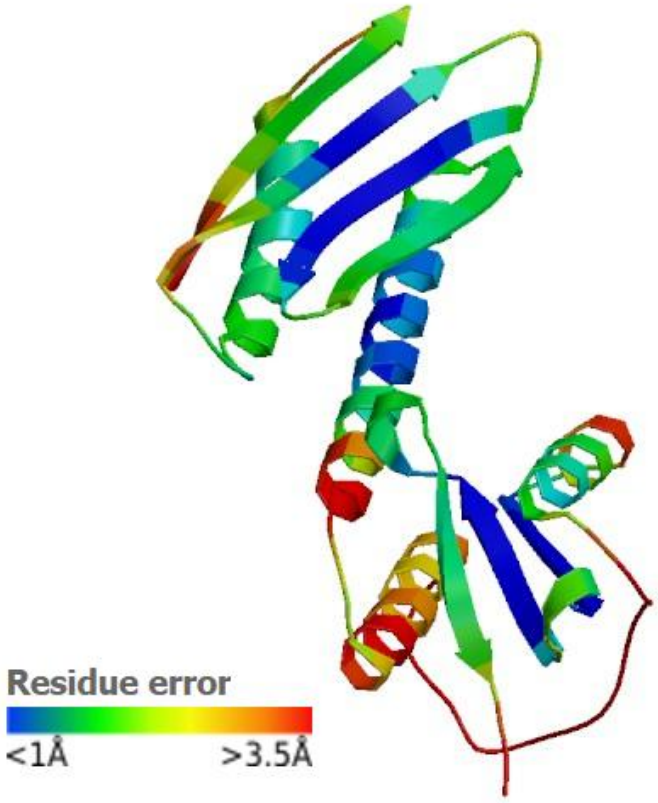

Figure 2. Error probability (in Å) of amino acid residues positions for the predicted SaHPF protein model № 3 built in Robetta program.

The analysis of RMSD (root mean square deviation from the initial structure) shows that the protein molecule is quite movable. For simulation steps with stable and minimum RMSD values, RMSF (deviation from average atomic coordinates value) was estimated (Fig.3), which showed that the amino acid residues of the hinge and the second domain are major contributors to the mobility of the protein. The RMSF average value for the first domain was 2.08 Å (fluctuation values for a.a. residues were in the interval from 1.06 to 4.51 Å), for a hinge the average was 3.15 Å (from 1.14 to 5.62 Å), for second domain the average was 2.87 Å (from 1.45 to 5.11 Å). RMSF for first domain was significantly less than the RMSF values for hinge and the second domain (p-value <0.001) (Fig.4). Principal component analysis and normal mode analysis molecular dynamics trajectories also confirm high mobility of hinge and second domain (Fig.5). It is most likely that the current problems of the SaHPF structure determination by experimental methods are caused by the mobility of the hinges and the second domain of this protein.

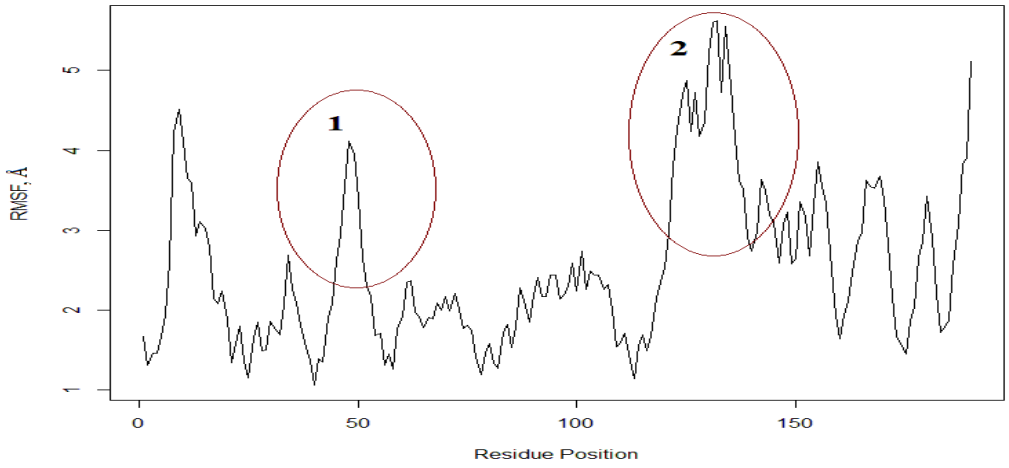

Figure 3. RMSF for MD frames with plato RMSD. "1" and "2" show SaHPF areas with the greatest mobility.

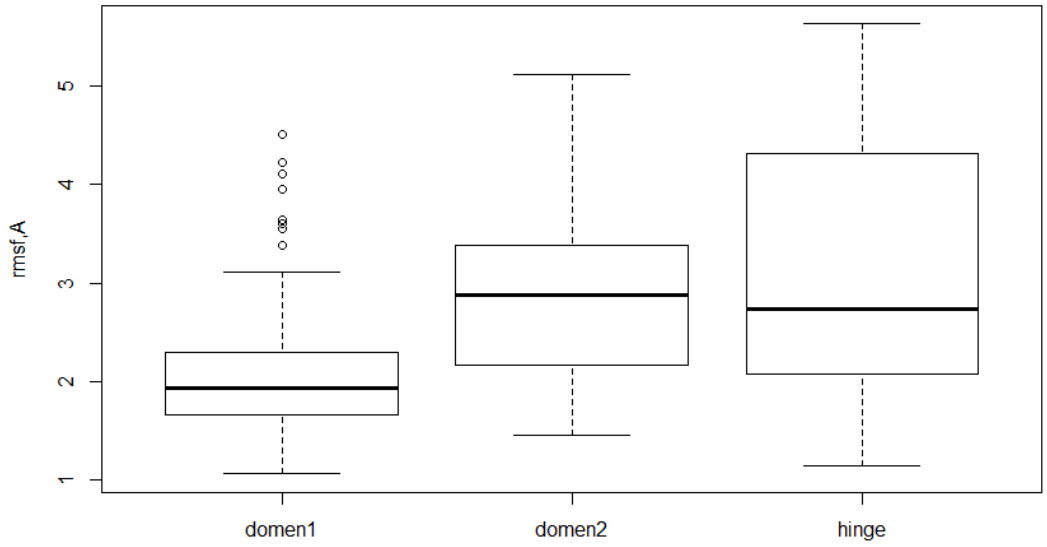

Figure4. 95% confidence interval of RMSF for MD frames with plato RMSD.

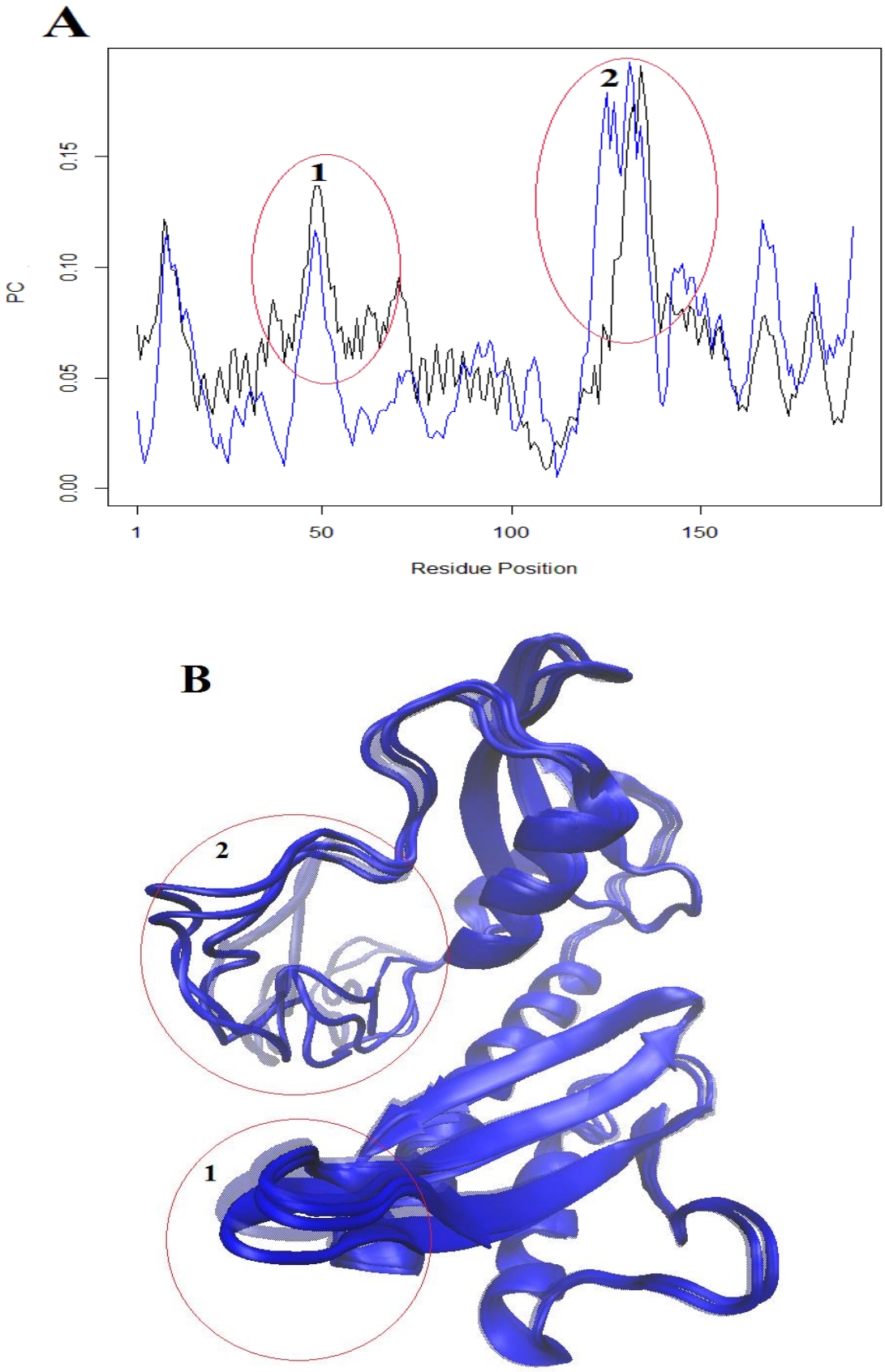

Figure 5.Intra-domain collective motions in the SaHPF structure determined by elastic normal modes analysis. A. The contribution of the amino acid residues in the first (pictured in black) and the second (pictured in blue) main components. B. Visualization of mobile parts in SaHPF structure obtained by normal mode method. "1" and "2" show SaHPF areas with the greatest mobility.

4. Summary

Computer methods predicted SaHPF protein structure model of sufficiently high quality according to Qmean evaluation. SaHPF comprises two domains and hinge connecting them. The first domain is a fully structured one, while the rest of the protein molecule does not have a clear secondary structure. The analysis of molecular dynamics trajectories for predicted SaHPF model confirms the high mobility of the loops first and second domains and especially for hinge. These structural characteristics of the protein may be a significant problem at the resolution of its three-dimensional structure by experimental methods.

Acknowledgement

The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University. The work was funded by RFBR, according to the research project No. 16-34-60001 mol_a_dk. This work was supported by the research grant of Kazan Federal University.We are grateful to Dr. Roland Stotefor the helpful comments while working on this article.

References

1. SaHPF. UniProt: http://www.uniprot.org/uniprot/D2Z097.
2. Ueta M, Wada, Ch, Wada A. Formation of 100S ribosomes in Staphylococcus aureus by the hibernation promoting factor homolog SaHPF // Genes to Cells, 2010. – v. 15(1). – p. 43-58.
3. Melzer M, Welch C. Thirty-day mortality in UK patients with community-onset and hospital-acquired meticillin-susceptible Staphylococcus aureus bacteraemia // J Hosp Infect, 2013. – v. 84(2). – p. 143-150.
4. Miller LG, Eells SJ, Taylor AR, David MZ, Ortiz N, Zychowski D, Kumar N, Cruz D, Boyle-Vavra S, Daum RS. Staphylococcus aureus colonization among household contacts of patients with skin infections: risk factors, strain discordance, and complex ecology // Clin Infect Dis, 2012. – v. 54(11). – p. 1523-1535.
5. Garreau de Loubresse, N., Prokhorova, I., Holtkamp, W., Rodnina, M., Yusupova, G., Yusupov M., Structural basis for the inhibition of the eukaryotic ribosome // Nature, 2014. – v. 513. – p. 517-522.
6. Ayupov, R.K., Andrianov, G.V., Akberova, N.I.. Comparative analysis of mice acetylcholinesterases by functional amino acid residues and molecular screening // Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2015. – v. 6(6) – P. 1781-1786.
7. Berman,H. M., Westbrook,J., Feng,Z., Gilliland,G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E..The Protein Data Bank // Nucleic Acids Research, 2010. – v. 28(1) – P. 235-242.

8. Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., Zhang,Y..The I-TASSER Suite: Protein structure and function prediction // Nature Methods, 2015. – v. 12 – P. 7-8.
9. Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., Kim,D., Kellogg, E.,DiMaio, F., Lange, O., Kinch, L.,Sheffler, W., Kim, B-H., Das, R., Grishin, N.V.,Baker, D..Structure prediction for CASP8 with all-atom refinement using Rosetta. // Proteins,2009. – v. 77 – Suppl. P. 9:89-99.
10. Xu, D., Zhang, Y.. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. // Proteins, 2012. – v. 80 – P. 1715-1735.
11. CAPS. Site: http://predictioncenter.org/
12. Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, Ch., Skeel, R.D., Kale, L., Schulten K.. Scalable Molecular Dynamics with NAMD // J ComputChem, 2005. – v. 26 – P. 1781–1802.
13. Salomon-Ferrer, R., Case, D.A.,Walker,R.C..An overview of the Amber biomolecular simulation package.// WIREs Comput. Mol. Sci., 2013. – v. 3 – P. 198-210.
14. Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., and MacKerell Jr., A.D. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi1 and chi2 dihedral angles. // Journal of Chemical Theory and Computation, 2012. – v. 8 – P. 3257-3273.
15. Humphrey, W.,Dalke, A.,Schulten,K..VMD - Visual Molecular Dynamics. // Journal of Molecular Graphics, 1996. – v. 14 – P. 33-38.
16. Skjærven, Yao, Scarabelli, Grant.Integrating protein structural dynamics and evolutionary analysis with Bio3D. // BMC Bioinformatics, 2014. – v. 15 – P. 399.
17. Grant, Rodrigues, ElSawy, McCammon, Caves. Bio3D: An R package for the comparative analysis of protein structures. // Bioinformatics, 2006.– v. 22 – P. 2695-2696.
18. Ayupov, R.K., Akberova, N.I.. Molecular dynamics of the pyridoxine derivative in the acetylcholinesterase active cavity // Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2015. – v. 6(6) – P. 1717-1722.
19. Ayupov, R.K., Akberova, N.I.. The ligand behavior on the surface of acetylcholinesterase // Research Journal of Pharmaceutical, Biological and Chemical Sciences, 2015. – v. 6(4) – P. 2202-2206.
20. Kelley, L.A.,Mezulis, S., Yates,Ch.M.,Wass, M.N., Sternberg, M.J.E.. The Phyre2 web portal for protein modeling, prediction and analysis // Nature Protocols, 2015. – v. 10 – P. 845–858.

21. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K.,Studer, G., Schmidt, T., Kiefer, F., Cassarino,T.G, Bertoni, M.,Bordoli, L.,Schwede,T.. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. // Nucleic Acids Research, 2014.– v. 42 – P. W252-W258.
22. Benkert, P., Tosatto, S.C.E. and Schomburg, D. QMEAN: A comprehensive scoring function for model quality assessment. // Proteins: Structure, Function, and Bioinformatics, 2008.– v. 71(1) – P. 261-277.
23. AD-ENM. Site: http://enm.lobos.nih.gov/
24. FlexServ. Site: http://mmb.pcb.ub.es/FlexServ/
25. Ye, K., Serganov, A., Hu, W., Garber, M., Patel,D.J..Ribosome-associated factor y adopts a fold resembling a double-stranded RNA binding domain scaffold // Eur.J.Biochem, 2002.– v. 269 – P.5182.
26. Parsons, L., Eisenstein, E., Orban, J. Solution structure of HI0257, a bacterial ribosome binding protein // Biochemistry, 2001. – v. 40. – P. 10979-10986.
27. 3LYV. Site: http://www.rcsb.org/pdb/explore/explore.do?pdbId=3LYV
28. Raka, A., Kalininb, A.,Shcherbakovb, D., Bayer, P.. Solution structure of the ribosome-associated cold shock response protein Yfia of Escherichia coli // Biochemical and Biophysical Research Communications, 2002.– v. 299(5) – P. 710–714.
29. 3KA5. Site: http://www.rcsb.org/pdb/explore/explore.do?pdbId=3KA5.
30. Polikanov Y, Blaha G, SteitzTh.. How Hibernation Factors RMF, HPF, and YfiA Turn Off Protein Synthesis // Science, 2012. – v. 336(6083). – P. 915-918.

Corresponding Author:
R.Kh. Ayupov*,
Email: *aurusta@mail.ru*