Journal of Bioresources 6(2): 56-60 (2019)
Print: ISSN 2394-4315
Online: ISSN 2582-2276

Original Research article

Comparison of Efflux Pump Protein Structures Predicted by different ComputationalTools

Nakul Neog, Minakshi Puzari and Pankaj Chetia*
Department of Life Sciences, Dibrugarh University, Assam- 786004, India
***Corresponding Author:** chetiapankaj@dibru.ac.in
Received: August 28, 2019; revised: October 30, 2019; accepted: November 10, 2019

Abstract: Three-dimensional protein structures are essential in bioinformatics for performing docking operations and molecular dynamics simulation. The Protein Data Bank (PDB) stores three-dimensional structure of proteins. However not all protein structures are available in PDB. To solve this problem, various web servers and software packages are available for prediction of protein structures. Yet before using them, it is important to test which one of them is more accurate. In our study, we compared the efflux pump protein structures of *Shigella flexneri* predicted by SWISS-MODEL, EasyModeller, Phyre2 and RaptorX.The protein sequence of AcrB of *Shigella flexneri* was retrieved from UniProt and the sequence was aligned using NCBI- BLAST to get similar structures present in PDB database. After structure prediction in respective web server/software, structures were visualized in RasMol and Ramachandran Plots were generated and analysed using RAMPAGE. The results depicted that EasyModeller and Phyre2 showed poor results than RaptorX and SWISS-MODEL. Ramachandran plots of structure predicted in RaptorX and SWISS-MODEL has 97.3% and 96.8% residues, respectively in favoured region. However, less no of residues were found in outlier region of structure predicted by SWISS-MODEL than RaptorX. The qualities of protein structure predicted using RaptorX and SWISS-MODEL can be said to be superior to those of structures predicted by EasyModeller and Phyre2. So, in order to find the most reliable tool for protein structure prediction, accuracy tests and validations are need to be done for better results with all the computational tools available.
Key words: EasyModeller, Phyre2, Protein Structure, RaptorX, SWISS-MODEL

Introduction

The three-dimensional structure of protein is essential for judicious designof many biological experiments (Arnold *et al.*, 2006) as it helps to understand function of protein at molecular level (Keifer *et al.*, 2009). PDB is a protein structure database maintained by Research Collaboratory for Structural Bioinformatics (RCSB), where experimentally determined protein structures are deposited.However, numbers of known protein structures in PDB are less compared to the number of known protein sequences in UniProt (Arnold *et al.*, 2006; Kelley *et al.*, 2015). When structures are not found in PDB, the best approach is to perform homology modelling (Youssef *et al.*, 2014). Three-dimensional structures of proteins are more conserved than protein sequences (Kelley *et al.*, 2015). In homology modelling, homologous proteins with known structure (template)for target sequence are identified; then target protein structures are predicted from alignment of the known target protein sequence with the template structure.

Various web servers and software packages are available that help in prediction of protein structure. The computational tools have been refined over the ages and these techniques aim to shrink the knowledge gap (Kelley *et al.*, 2015). Confusion and difficulty can arise due to lack of

comparison of these tools (Youssef *et al.*, 2014). It is necessary to check the accuracy of results obtained from different softwares.

In this paper, we have predicted structures of efflux pump protein present in *Shigella flexneri* using different tools and compared the results to check the accuracy of predictions.*Shigella flexneri* belongs to genus *Shigella* of family Enterobacteriaceae and is the causative organism of shigellosis(Taneja and Mewara, 2016). Due to the emergence of antibiotic resistance in *Shigella* species, it has been declared as a global threat by the World Health Organization (WHO report, 2014). Efflux pumps are one of the mechanisms by which they have developed resistance against drugs (Fernandez-Recio *et al.*, 2004). Efflux pump activity can be inhibited by inhibitors(Piddock *et al.*, 2010) and to design them accurate structure of efflux pump protein has to be known. So we have predicted and compared the structures of AcrB, an efflux pump protein, which plays important role in substrate specificity(Vargiu and Nikaido, 2012)

Materials and methods

Softwares used for prediction and comparison

The amino-acid sequence for AcrB protein of *Shigella flexneri* was searched in UniProt. The sequence was downloaded in FASTA format andcompared with protein structures present in PDB using Basic Local Alignment Search Tool (BLAST) by National Center for Biotechnology Information (NCBI). The structure showing maximum identity to the query sequences were saved in .pdb format. Then homology modelling was done using SWISS-MODEL server, web portal phyre2, web portal RaptorX and EasyModeller, which is a graphical interface to Modeller.

SWISS-MODEL is the pioneer of fully automated web-based workspace for homology modelling (Arnold *et al.*, 2005) and uses visualization tool Swiss-PdbViewer (Guex N., 2009). It removes the need for complex software packages and is user-friendly (Biasini *et al.*, 2014). Phyre2 was launched in January 2011 and has been widely used for protein structure prediction as it is easy to use (Kelley *et al.*, 2015). RaptorX is a web-based server developed by Xu group and officially released in August 2011 for prediction of protein structure (Källberg *et al.*, 2012). MODELLER is a widely used program for comparative modelling of three-dimensional structures of protein. It is hard to use as it requires knowledge of python scripting and is command line based. EasyModeller tool is a frontend graphical interface for Modeller, which can be used in Windows platform having pre-installed MODELLER and Python (Kuntal *et al.*, 2010).

After prediction of protein structure, RasMol visualization tool was used to view the predicted structures, which is a molecular graphics program (Sayle and Miller-White, 1995). Ramachandran plot was generated and analysed for each predicted model using RAMPAGE. Based on the Ramachandran plot assessment, accuracy was determined for the predicted models.

Results

UniProt Search

The search for sequence of AcrB in *Shigella flexneri* showed several results. The sequence selected for structure prediction had entry name Q0T7C2_SHIF8 and it belonged to *Shigella flexneri* serotype 5b (strain 8401). It has length of 1049 residues. (Fig. 1)

BLAST Search

The sequence was aligned with protein structures present in PDB database using BLAST. The structure with PDB ID 1OY6 showed maximum identity with 0.0 E value and 100% query cover (Fig. 2). It belonged to chain A of AcrB protein in *Escherichia coli.*

Ramachandran Plot Analysis

The analysis of Ramachandran plot for model from SWISS-MODEL as seen in Fig.3 showed that the number of residues in favoured region were(~98.0% expected): 1005 (96.8%), number of residues in allowed region were (~2.0% expected): 31 (3.0%) and number of residues in outlier region were: 2 (0.2%).

For model from Phyre2, evaluation of residues in Ramachandran plot as seen in Fig.4 showed that the number

>tr|Q0T7C2|Q0T7C2_SHIF8 Efflux pump membrane transporter OS=Shigella flexneri serotype 5b (strain 8401) OX=373384 GN=acrB PE=3 SV=1
MPNFFIDRPIFAWVIAIIIMLAGGLAILKLPVAQYPTIAPPAVTISASYPGADAKTVQDT
VTQVIEQNMNGIDNLMYMSSNSDSTGTVQITLTFESGTDADIAQVQVQNKLQLAMPLLPQ
EVQQQGVSVEKSSSSFLMVVGVINTDGTMTQEDISDYVAANMKDAISRTSGVGDVQLFGS
QYAMRIWMNPNELNKFQLTPVDVITAIKAQNAQVAAGQLGGTPPVKGQQLNASIIAQTRL
TSTEEFGKILLKVNQDGSRVLLRDVAKIELGGENYDIIAEFNGQPASGLGIKLATGANAL
DTAAAIRAELAKMEPFFPSGLKIVYPYKDTTPFVISIHEVVKTLVEAIILVFLVMYLFLQ
NFRATLIPTIAVPVVLLGTFAVLAAFGFSINTLTMFGMVLAIGLLVDDAIVVVENVERVM
AEEGLPPKEATRKSMGQIQGALVGIAMVLSAVFVPMAFFGGSTGAIYRQFSITIVSAMAL
SVLVALILTPALCATMLKPIAKGDHGEGKKGFFGWFNRMFEKSTHHYTDSVGGILRSTGR
YLVLYLIIVVGMAYLFVRLPSSFLPDEDQGVFMTMVQLPAGATQERTQKVLNEVTHYYLT
KEKNNVESVFAVNGFGFAGRGQNTGIAFVSLKDWADRPGEENKVEAITMRATRAFSQIKD
AMVFAFNLPAIVELGTATGFDFELIDQAGLGHEKLTQARNQLLAEAAKHPDMLTSVRPNG
LEDTPQFKIDIDQEKAQALGVSINDINTTLGAAWGGSYVNDFIDRGRVKKVYVMSEAKYR
MLPDDIGDWYVRAADGQMVPFSAFSSSRWEYGSPRLERYNGLPSMEILGQAAPGKSTGEA
MELMEQLASKLPTGVGYDWTGMSYQERLSGNQAPSLYAISLIVVFLCLAALYESWSTPFS
VMLVVPLGVIGALLAATFRGLTNDVYFQVGLLTTIGLSAKNAILIVEFAKDLMDKEGKGL
IEATLDAVRMRLRPILMTSLAFILGVMPLVISTGAGSGAQNAVGTGVMGGMVTATVLAIF
FVPVFFVVVRRRFSRKNEDIEHSHTVDHH

Fig. 1. Sequence of AcrB from UniProtKb.

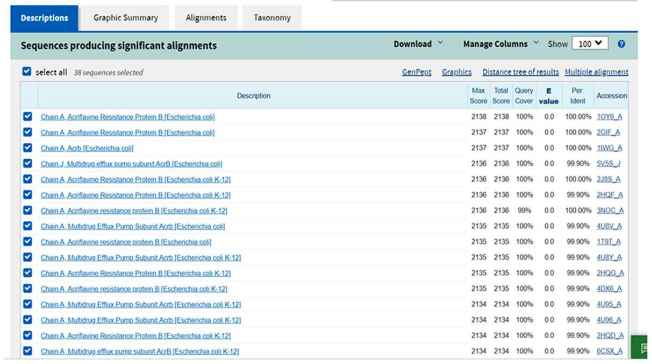

Fig. 2. BLAST results for the sequence from NCBI-BLAST web page.

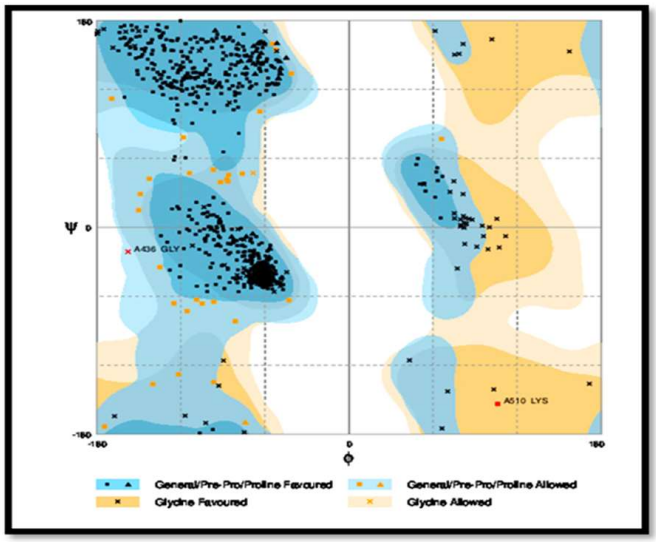

Fig. 3. Ramachandran plot for model from SWISS-MODEL

of residues in favoured region were (~98.0% expected): 868 (85.2%), number of residues in allowed region were (~2.0% expected): 122 (12.0%) and number of residues in outlier region : 29 (2.8%).

Analysis of Ramachandran plot for model as seen in Fig. 5 from RaptorX showed number of residues in favoured

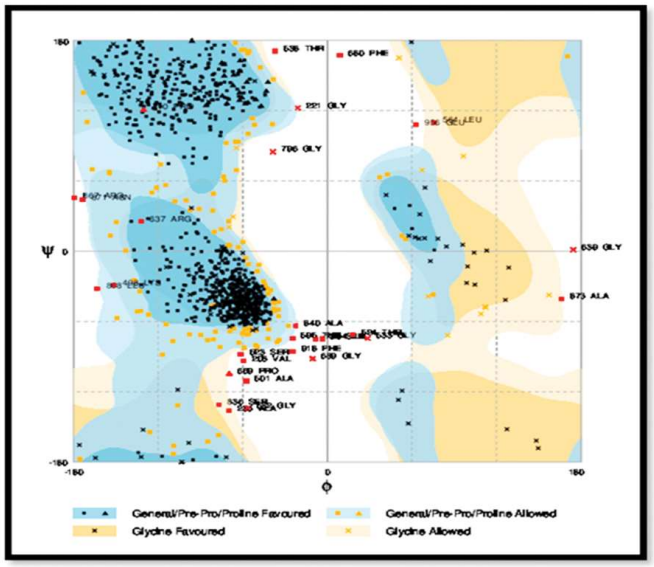

Fig. 4. Ramachandran plot for model from Phyre2.

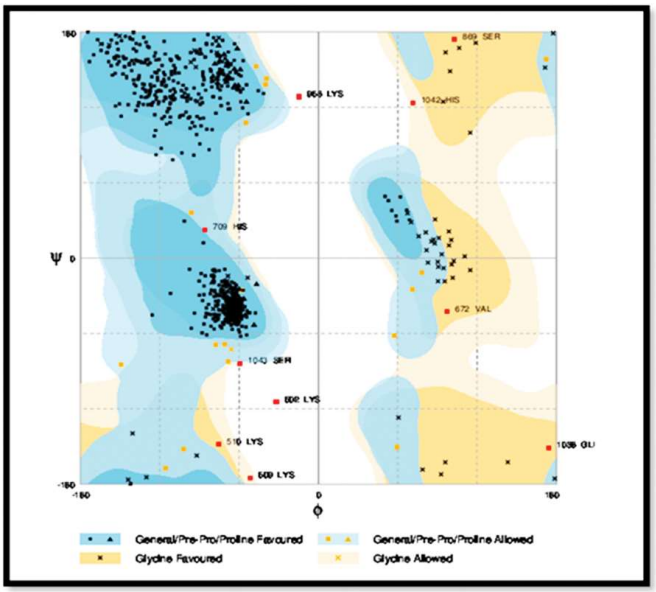

Fig. 5. Ramachandran plot for model from RaptorX.

region were (~98.0% expected): 1019 (97.3%), number of residues in allowed regionwere (~2.0% expected): 18 (1.7%) and number of residues in outlier region were:10 (1.0%).

Evaluation of Ramachandran plot for model from EasyModeller as seen in Fig. 6 showed number of residues in favoured region were (~98.0% expected): 574 (54.8%), number of residues in allowed region (~2.0% expected): 200 (19.1%) and number of residues in outlier region were: 273 (26.1%).

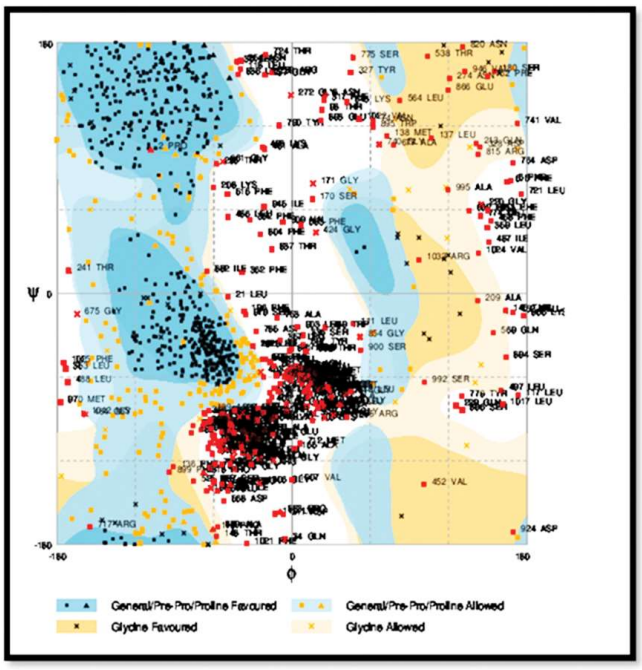

Fig. 6. Ramachandran plot for model from EasyModeller.

Discussion

From evaluation of residues of Ramachandran Plot for each predicted model, it is clear that in this particular study, Phyre2 and EasyModeller have shown poor results compared to RaptorX and SWISS-MODEL. EasyModeller model has only 54.8% residues in favoured region and 26.1% in outlier region making it highly inaccurate. Phyre2 is better than EasyModeller as it has 85.2% residues in favoured region and 2.8% in outlier region.Yet compared to SWISS-MODEL and RaptorX, it can be described as not so accurate. SWISS-MODEL model has 96.8% residues in favoured region and 0.2% in outlier region, making it highly accurate. RaptorX model has 97.3% residues in favoured region, which are more than SWISS-MODEL. RaptorX can be said as more accurate if we compare the number of residues in favoured region. However if we compare the residues in outlier region then RaptorX model has 1% in outlier region,

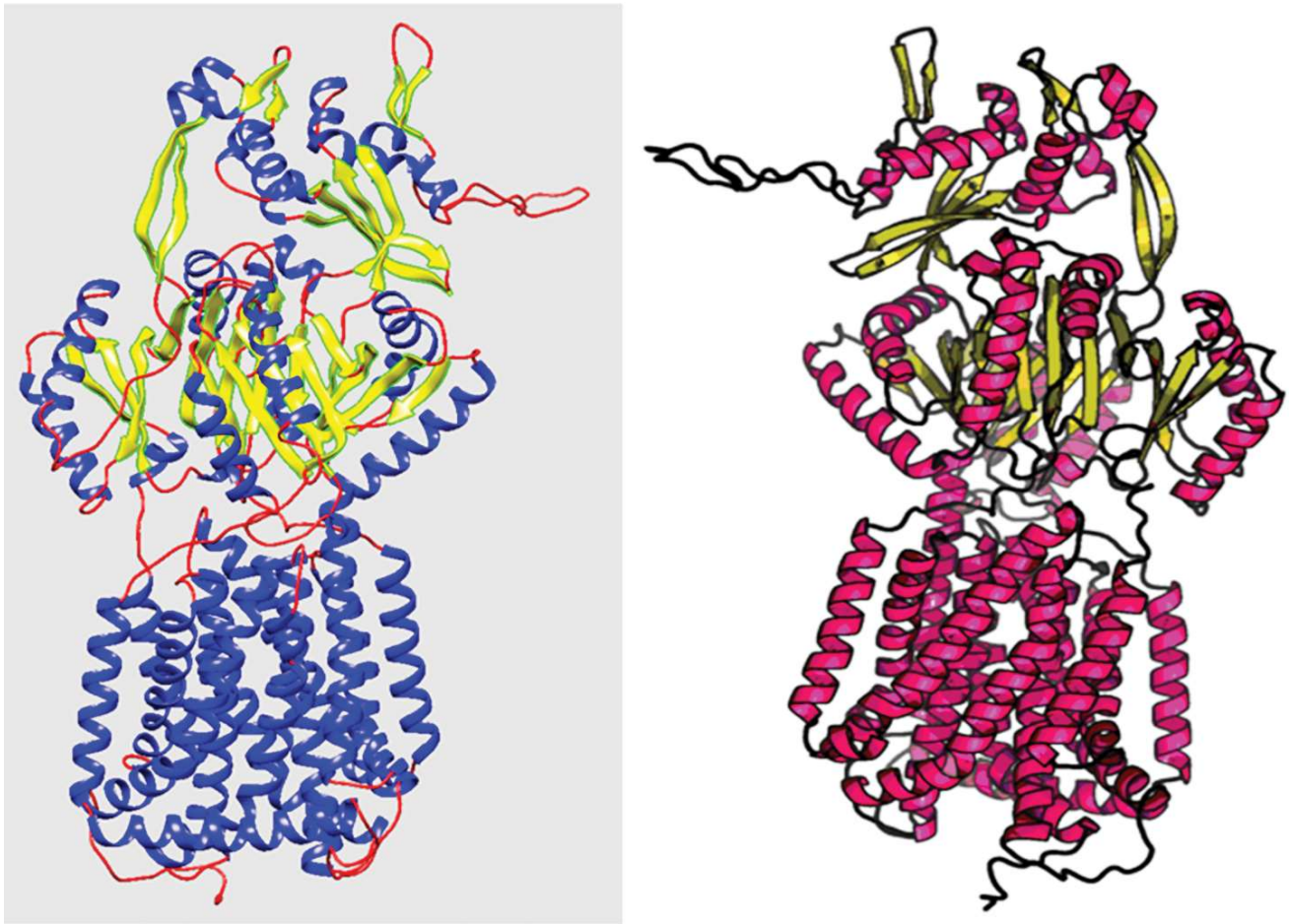

Fig. 7. 3D Structures of Modelled Protein AcrB (RaptorX).

which is more than the percentage showed by SWISS-MODEL. Again if we compare the number of residue in allowed region, then SWISS-MODEL model has 3% while RaptorX model has 1.7%. So SWISS-MODEL and RaptorX both are accurate, but to describe which is more accurate, depends on the parameter we consider for judging them. If we focus on favoured region residue number then RaptorX is better and if focus is on less residues in outlier region than SWISS-MODEL is more accurate. The qualities of proteins predicted in RaptorX and SWISS-MODEL were found to be superiorto those predicted in EasyModeller and Phyre2. The Ramachandran plot analysis of the modelled proteins validates the same.

From the study we can conclude that out of the four computational tools we have used for protein structure prediction, EasyModeller and Phyre2 are highly inaccurate as per evaluation of residues in Ramachandran plot and the models generated by them can't be considered for study. SWISS-MODEL and RaptorX are reliable and accurate for generation of models for further study in structural bioinformatics. There are several tools for protein structure prediction and in our study we have focused on only four of them. Hence we can't claim that they are the best tools available unless we have compared them with other available tools. So there is a need to test the accuracy of the other available software packages and web-servers as well, in order to find the most reliable tool that generates highly accurate result.

Acknowledgment

The authors are thankful to Department of Life Sciences, Dibrugarh University, Assam for providing the necessary facilities for the study.

References

Arnold K, Bordoli L, Kopp J and Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 22:195-201.

Biasini M, Bienert S, Waterhouse A, *et al.* 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucl. Acids Res. 42: W252-W258.

Fernandez-Recio J, Walas F,Federici L, Pratap JV, Bavro VN, Miguel RN, Mizuguchi K, Luisi B.2004. A model of a transmembrane drug-efflux pump from Gram-negative bacteria. FEBS Letters. 578:5-9.

Kelley LA, Mezulis S, Yates CM, Wass MN and Sternberg MJE. 2015. The Phyre2 web portal for protein modelling, prediction and analysis. Nature Protocols. 10: 845-858.

Piddock LJV, Garvey MI, Rahman MM and Gibbons S. 2010. Natural and synthetic compounds such as trimethoprim behave as inhibitors of efflux in Gram-negative bacteria. J. Antimicrob. Chemother. 65:1215-1223.

Guex N, Peitsch MC and Schwede T. 2009. Automated comparative protein structure modelling with SWISS-MODEL and SWISS-PdbViewer: a historical perspective. Electrophoresis. 30: S162-S173.

Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H and Xu J. 2012. Template-based protein structure modelling using the RaptorX web server. Nature Protocols. 7: 1511-1522

Keifer F, Arnold K, Künzil M, Bordoli L and Schwede T. 2009. The SWISS-MODEL Repository and associated resources. Nucl. Acids Res. 37: D387-D392.

Kuntal BK, Aparoy P and Reddanna P. 2010. EasyModeller: A graphical interface to MODELLER. BMC Research Notes. 3: 226.

Sayle RA and Milner-White EJ.1995. RASMOL: biomolecular graphics for all. TIBS. 20: 374-376.

Taneja N and Mewara A. 2016. Shigellosis epidemiology in India. Ind. J. Med. Res. 143: 565-576.

Vargiu AV and Nikaido H. 2012. Multidrug binding properties of the AcrB efflux pump characterized by molecular dynamics simulations. PNAS. 109: 20637-20642.

Youssef BA, Rashwan S and Redwan EM. 2014. Comparison of two academic software packages for protein structure prediction. Int. J. Biosci. Biotechnol. 6: 49-54.

WHO (2014): Antimicrobial Resistance: Global Report on Surveillance. WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland.