International Journal of Engineering & Technology, 7 (2.7) (2018) 84-87

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper



Identify virtual ligand hits using consensus scoring approach for drug target S. Aureus

P. Bharath Siva Varma[1]*, Adimulam Yesubabu[2], K. Subrahmanyam[3]

[1]Research Scholar, KLEF, India
[2]Head & Professor, Department of CSE, Sir C R Reddy College of Engineering, India
[3]Professor, Department of CSE, KLEF, India
**Corresponding author E-mail:pbsvarma@gmail.com*

Abstract

The associations amongst protein and ligand can be assessed for scoring functions based on the binding modes of ligand obtained from search algorithms. Several scoring functions have been proposed and every method has their own strengths and weaknesses, hence a multiple scoring analysis referred as consensus scoring increases the overall signal noise ratio. Therefore, top ten ligands obtained and were subjected to re-scoring using mcule, AutoDock and X-score functions. A computational consensus scoring analysis was taken up for a dataset of top ten PubChem compounds which are identified to exhibit better inhibitory properties against phosphotransacetylase, a putative drug target for S. aureus.

Keywords: *Consensus scoring; Search algorithm; Scoring function; AutoDock; X-score.*

1. Introduction

The complete human genome has many genes for which function can be ascertained. However, there are certain genes for which the associated function is not known. Such genes are being expressed in many normal as well as disease conditions suggesting the fact that they are primarily important in maintaining some useful activity within the biological process [1]. These proteins of unknown function have been reported in database and few structures of these sequences are also deposited in Protein Data Bank. This makes computational scientists to annotate either gene or protein structure to perform computer-aided in silico analysis to ascertain possible function based on bioinformatics analysis [2]. Many scoring functions may be either force field, empirical and knowledge-based [3]. There are two ways in which the scoring functions are utilized by Current docking methods. These methods use one of them. For ranking a protein-ligand conformation the full scoring function is used in the first approach. The search algorithm is used to modify the system and the same procedure is applied again to rank the new structure. A minimized function is utilized in guiding the search and the diligent one is then used for ranking the obtained structures in another approach [4] [5].

1.1 S. aureus

S. aureus is identified as one of the most salient bacteria which engender diseases in humans from a long time. It is one of the most significant generator of soft skin tissue illness such as ab scesses (boils), furuncles and cellulitis. Even though staphylococcus diseases aren't critical, S. aureuscan lead to critical diseases
like bloodstream infections, pneumonia, or bone and joint infections [6].
S. aureusis is established as a commensal related with skin, skin glands and mucous membranes, espicially in the nose of healthy persons. S. aureuscan is an engender of central venous catheter-associated bacteremia and ventilator-assisted pneumonia. There is a scope to be effected with serious deep-rooted infections, such as endocarditis and osteomyelitis. S. aureusis is answerable for toxin-mediated diseases, such as toxic shock syndrome, scalded skin syndrome and staphylococcal foodborne diseases (SFD). The significance of this human virus, apart from its capacity to engender dangerous diseases, is its extraordinary ability to arise antimicrobial resistance [7].

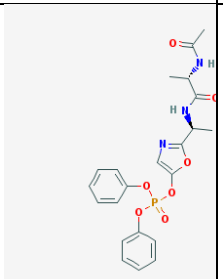
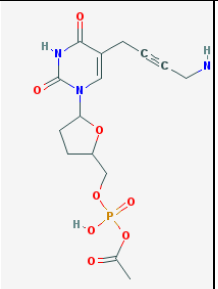
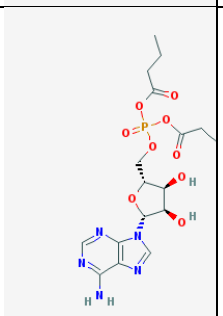
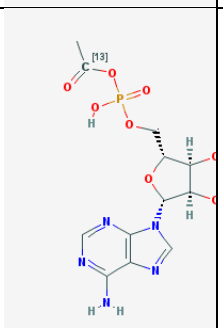
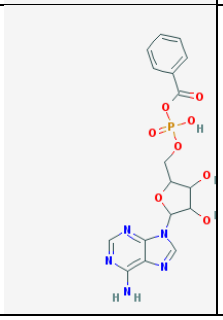
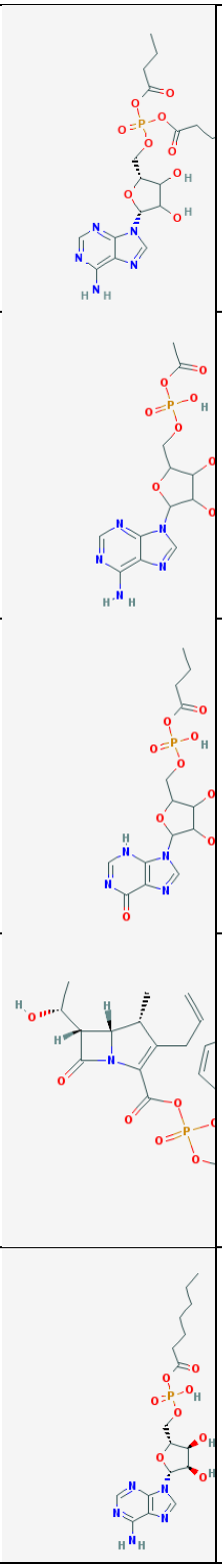
2. Materials and Methods

2.1. Data Set

A dataset of top 10 PubChem compounds that showed better binding potential with 4E4R protein from our earlier work [8] was considered as dataset. The structures of all ligands are given in table 1.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Table 1: 2-dimensional structures and IUPAC names of all ligands used for consensus scoring analysis.

No.	PubChem ID	2-D Structure	IUPAC Name
1	10096390		[2-[(1S)-1-[[(2S)-2-acetamidopropanoyl]amino]ethyl]-1,3-oxazol-5-yl] diphenyl phosphate
2	20655059		[[5-[2,4-dioxo-5-[4-[(2,2,2-trifluoroacetyl)amino]but-2-ynyl]pyrimidin-1-yl]oxolan-2-yl]methoxy-hydroxyphosphoryl] acetate
3	10074176		[[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-butanoyloxyphosphoryl] butanoate
4	10693910		[[(3aR,4R,6R,6aR)-4-(6-aminopurin-9-yl)-2,2-dimethyl-3a,4,6,6a-tetrahydrofuro[3,4-d][1,3]dioxol-6-yl]methoxy-hydroxyphosphoryl] acetate
5	5229061		[[5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl] benzoate
6	69981483		[[(2R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-butanoyloxyphosphoryl] butanoate
7	5241368		[[5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl] acetate
8	5229062		[[3,4-dihydroxy-5-(6-oxo-3H-purin-9-yl)oxolan-2-yl]methoxy-hydroxyphosphoryl] butanoate
9	57155509		diphenoxyphosphoryl (4R,5S,6S)-6-[(1R)-1-hydroxyethyl]-4-methyl-7-oxo-3-prop-2-enyl-1-azabicyclo[3.2.0]hept-2-ene-2-carboxylate
10	10813788		[[(2R,3S,4R,5R)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methoxy-hydroxyphosphoryl] heptanoate

2.2 Molegro virtual docker (MVD) software

Molegro Virtual Docker (MVD) is a desegregated stage for anticipating protein-ligand interactions. MVD identifies the target protein and determines the potential binding sites. MVD software relies upon another search algorithm which merges with a cavity expectation algorithm. Piecewise Linear Potential (PLP) is derived from dock scoring function of MVD. For the supplementary improvement of docking result, scoring function is re-ranked and instigated, which perceives the most optimistic docking arrangement by the docking algorithm.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2.3 Scoring Function

MVD uses scoring function to upgrade bonding with hydrogen term and new charge schemes. The dock score function is Escore, which is described in terms of sum of energies as following

$$E_{score} = E_{inter} + E_{intra}$$

where E_{inter} is the ligand−protein interaction energy:

$$E_{inter} = \sum_{i\in \text{ligand}} \sum_{j\in \text{protein}} \left[E_{PLP}(r_{ij}) + 332.0\frac{q_i q_j}{4r_{ij}^2} \right]$$

The piecewise linear potential is represented by Eplp. The charged atoms electrostatic interactions are in between charged atoms were described by second term. It is given by distance dependent dielectric constant D(r))4r. A unit of electrostatic energy is fixed by the numerical value 332.0.

E_{intra} is the internal energy of the ligand:

$$E_{intra} = \sum_{i\in \text{ligand}} \sum_{j\in \text{ligand}} E_{PLP}(r_{ij}) + \sum_{\text{flexible bonds}} A[1-\cos(m\cdot\theta-\theta_0)] + E_{clash}$$

The ligand atom pairs described by the double summation, excludes the pair of atoms which are associated by two bonds or less. The second one is a torsional energy term which is specified as indicated to the bonded atoms of hybridization types.

2.4 Consensus scoring

Consensus scoring was achieved by implementing three programs such as mcule-1click dock (https://mcule.com) program, AutoDock [9] and X-score [10] scoring functions, respectively. mcule 1-click docking is an online server for drug discovery platform performs docking.

3. Results and discussion

In order to understand which compounds exhibited better interactions with receptor can be known by implementing all three scoring functions. However, although docking performed on the dataset, the problem of identifying true hits with better affinity has not been possible because of the fact that few compounds displayed better scores for one scoring function whereas few other compounds reported better interaction energies for the other. The hypothetical protein [Staphylococcus aureus] is given below.
NCBI Reference Sequence: WP_000894032.1 LOCUS WP_000894032 310 aalinear BCT

DEFINITIONhypothetical protein [Staphylococcus aureus].

ACCESSION WP_000894032

VERSION WP_000894032.1GI:446816776

KEYWORDSRefSeq.

SOURCEStaphylococcus aureus

ORGANISMStaphylococcus aureus

Bacteria; Firmicutes; Bacilli; Bacillales; Staphylococcus.

source1..310

/organism="Staphylococcus aureus"

/db_xref="taxon:1280"

Protein 1..310

/product="hypothetical protein"

/calculated_mol_wt=34493

Region1..>309

/region_name="YbbR"

/note="Uncharacterized protein, YbbR domain [Function unknown]; COG4856"

/db_xref="CDD:227193"

Region141..220

/region_name="YbbR"

/note="YbbR-like protein; pfam07949"

/db_xref="CDD:254535"

Region237..307

/region_name="YbbR"

/note="YbbR-like protein; pfam07949"

/db_xref="CDD:254535"

ORIGIN

1 mleskwglrfiafllalfffflsvnnvfgnifntgnlgqks
sktiqdvpveilyntkdlhl 61
tkapetvnvtisgpqskiikienpedlrvvidlsnakag
kyqekyqvkgladdihysvkp

121 klanitlenkvtkkmtvqpdvsqsdidplykitkqevspqtvkvtggeeqlndia
ylkat

181 fktnkkingdtkdvaevtafdkklnklnvsiqpnevnlqvkvepfskkvkvnv
kqkgsla

241 ddkelssidledkeieifgsrddlqnisevdaevdldgisestektvkinlpehvtk
aqp

301 setkayinvk
//

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The scores of all dataset compounds for all scoring functions were given in Table-1.

Table 1: Ten ligand dataset binding scores obtained from mcule, AutoDock and X-score

No.	PubChem ID	mcule	AutoDock	HPScore	HMScore	HSScore
1	10096390	-8.8	-9.1	7.48	7.93	7.49
2	20655059	-8.7	-8.5	7.01	6.86	6.94
3	10074176	-7	-8.3	6.58	6.41	6.36
4	10693910	-7.3	-8.0	6.62	6.74	6.8
5	5229061	-7.7	-7.9	6.83	7.04	6.45
6	69981483	-6.2	-7.1	6.49	6.74	6.51
7	5241368	-7.4	-7.9	6.09	6.02	5.95
8	5229062	-5.9	-6.5	6.24	6.14	6.01
9	57155509	-6.0	-7.4	7.65	8.13	7.71
10	10813788	-6.2	-6.9	6.42	6.61	6.34

Table 2, it is evidenced that high rank-sum 20 was obtained for PubChem compound with ID 10096390 followed by ID 20655059 and third rank was reported by ID 57155509 compound, respectively. Therefore, from this table and analysis, it can be suggested that compound 10096390 is more pronounced to interact with phosphotransacetylase and hence represents best hit. DST and Zhang rule data given in Table 4 suggested PubChem ID 10096390, an oxazolyl derivative, as best compound from the dataset which showed better affinity of binding with phosphotransacetylase.

Table 2: Generated classes and rank-sum for all compounds in the study.

No.	PubChem ID	mcule	AutoDock	HPScore	HMScore	HSScore	Rank Sum
1	10096390	4	4	4	4	4	20
2	20655059	4	4	3	2	3	16
3	10074176	2	3	2	1	1	9
4	10693910	2	3	2	2	2	11
5	5229061	3	3	2	2	2	12
6	69981483	1	1	2	2	2	8
7	5241368	3	3	1	1	1	9
8	5229062	1	1	1	1	1	5
9	57155509	1	2	4	4	4	15
10	10813788	1	1	1	2	1	6

4. Conclusion

A computational consensus scoring analysis was taken up for a dataset of top ten PubChem compounds which are identified to exhibit better inhibitory properties against phosphotransacetylase, a putative drug target for S. aureus. Submitting the compounds to re-score the docked poses resulted in varied scores for mcule, AutoDock and X-score functions. Hence, application of rank-sum technique, DST unweighted, weighted and Zhang rule suggested high rank for PubChem ID 10096390 and hence reported as best hit and hence experimental testing would reveal other related features of this compound towards inhibiting infectious agent S. aureus.

References

[1] Jensen, L. J., Gupta, R., Staerfeldt, H. H., & Brunak, S. (2003). Prediction of human protein functions according to Gene Ontology categories. Bioinformatics, 19(5), 635-642.

[2] Marianne A. Grant. INTEGRATING COMPUTATIONAL PROTEIN FUNCTION PREDICTION INTO DRUG DISCOVERY INITIATIVES. Drug Dev Res. 2011 Feb; 72(1): 4–16.

[3] Ballester, P. J., & Mitchell, J. B. O. (2010). A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. Bioinformatics (Oxford, England), 26(9), 1169–1175.

[4] Kitchen DB, Decornez H, Furr JR, Bajorath J (2004). "Docking and scoring in virtual screening for drug discovery: methods and applications". Nature reviews. Drug discovery 3 (11): 935–49.

[5] Luty BA, Wasserman ZR, Stouten PFW, Hodges CN, Zacharias M, McCammon JA. A molecular mechanics/grid

[6] http://www.health.state.mn.us/divs/idepc/diseases/staph/basics.html

[7] Konrad Plata, Adriana E. Rosato and GrzegorzWęgrzyn. Staphylococcus aureus as an infectious agent: overview of biochemistry and molecular genetics of its pathogenicity. Vol. 56 No. 4/2009, 597–612.

[8] P. Bharath Siva Varma, Yesubabu Adimulam and K. Subrahma nyam, “In Silico virtual screening of PubChem compounds against phosphotransacetylase, a putative drug target for Staphylococcus aureus”, International Journal of Computational Biology and Drug Design, Vol. 10, No. 1, 2017.

[9] Morris, G.M., Goodsell, D.S., Halliday, S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., J. Comput. Chem., 19 (1998) 1639.

[10] Wang, R.; Lai, L.; Wang, S.Further Development and Validation of Empirical Scoring Functions for Structure-Based Binding Affinity Prediction. J. Comput.-Aided Mol. Des. 2002, 16, 11-26.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.