INTERNATIONAL JOURNAL OF ADVANCES IN PHARMACY, BIOLOGY AND CHEMISTRY

Research Article

In Silico Design of Inhibitors for *Staphylococcus epidermidis* Biofilm

Zahra M. Al-Khafaji *Aymen F. Al-Mulla*.

Institute of Genetic Engineering and Biotechnology for Postgraduate Studies,

Univ. of Baghdad, Iraq.

ABSTRACT

In Silico drug design represents a new approach for drug discovery and industry. Structure based computer aided drug design (CADD) was used in this study to find an antibiofilm agents to suppress *Staphylococcus epidermidis* biofilm production which is considered the main virulence factor of this bacterium. The sarA protein was chosen as the target for this process as it stimulates *ica*ADBC operon which is responsible for biofilm production. The first step was constructing a 3D structure of the protein which was obtained using the RaptorX homology modeling. Pharmacophore generation was performed using the Hip Hop generator from Discovery Studio package. One hundred seventy seven molecules were chosen by ligand based virtual screening using ZincPharmer. Thirty seven molecules were found suitable as having negative binding free energies with sarA protein in EADock engine from the SwissDock Website. These molecules can be tested for *in vitro* studies as antibiofilm agents.

Key words: drug design, bioinformatics, antibiofilm, *S. epidermidis*, Homology Modeling, Pharmacophore.

INTRODUCTION

Drug discovery process is a critical issue in the pharmaceutical industry since it is a very cost and time consuming process (Rao and Srinivas, 2011). Two different methods are widely used in the pharmaceutical industry for finding hits: high throughput screening and virtual screening. The former process is commonly used in all major pharmaceutical industries. However, the cost in synthesis of each compound, *in vitro* testing and low hit rate are posing huge problems for pharmaceutical industries. Current efforts within the industry are directed to reduce the timeline and costs (Böhm *et al.*, 2000). At present, hundreds of thousands to millions of molecules have to be tested within a short period for finding novel hits, therefore, highly effective screening methods are necessary for today's researchers. In view of the above problems in finding new drugs by HTS; cost effective, reliable *in Silico* screening procedures are in practice (Young, 2009).

In Silico drug design means rational design by which drugs are designed/discovered by using computational methods.It can be applied by either of two strategies of design depending on the knowledge of the target, presence of the primary sequence and 3D structure. The first approach, Structure-based drug design (SBDD) is one of the earliest techniques used in drug design. Drug targets are typically key molecules involved in a specific metabolic or cell signaling pathway that is known, or believed, to be related to a particular disease state. Drug targets are most often proteins and enzymes in these pathways.SBDD uses the known 3D geometrical shape or structure of proteins to assist in the development of new drug compounds,which is derived from x-ray crystallography or nuclear magnetic resonance (NMR) techniques that can resolve the structure of proteins to a resolution of a few angstroms (Rao and Srinivas, 2011). The other approach is ligand based drug design which is used when the target is unknown,for example, cell surface receptors make excellent drug targets, but are very difficult to crystallize. So if homology modeling was unreliable or low identity score for the homolog protein was observed, in this case the techniques used

for structure-based drug design cannot be used. Pharmacophore models and 3D-QSAR models can be used instead (Young, 2009).
The aim of this study was to predict inhibitors for *S. epidermidis* biofilm using structure based computer aided drug design strategies.

MATERIALS AND METHODS

A- NCBI: http://www.ncbi.nlm.nih.gov/used for the retrieval of protein sequence and information (Matthiesen, 2010).
B- Uniprot Database: http://www.uniprot.org/used for protein information (Apweiler *et al.*, 2004).
C- Mega5.1 software: used for alignment purposes (Tamura *et al.*, 2011).
D- RaptorX: http://raptorx.uchicago.edu/used for protein modeling purposes (Källberg *et al.*, 2012).
E-Qmeanserver: http://swissmodel.expasy.org/qmean/cgi/index.cgiused for protein model quality estimation (Benkert *et al.*, 2009).
F- SwissDock: http://www.swissdock.ch/used for molecular docking purposes (Hetal *et al.*, 2013).
G- Discovery Studio software v.2.5: used for multidrug design purposes (Tsai *et al.*, 2009).
H- ZINC Database: http://zinc.docking.org/used for cheminformatic purposes (Irwin and Shoichet, 2005).
I-ZINCPharmer: http://zincpharmer.csb.pitt.edu/used for pharmacophore screening purposes (Koes and Camacho, 2012).
J- T.E.S.T software v.4.1:used for toxicity and mutagenicity estimation (Sushko *et al.*, 2010).
K- Cello: http://cello.life.nctu.edu.tw/used for estimation of protein localization (Yu *et al.*, 2006).
L-BTXpred: http://www.imtech.res.in/raghava/btxpred/index.htmlused for the prediction of bacterial toxins (Saha and Raghava, 2007).
M-VaxiJen: http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.htmlused for the prediction of antigenicity and subunit vaccines (Doytchinova and Flower, 2007).
N- STRING database:http://string-db.org/used toestimatethe protein interactome (Franceschini*et al.*, 2013).

RESULTS

Structure based drug design strategies were used and as follow:
Target Determination: Some of the proteins and genes of *S.epidermidis* were studied to selectthe most effective and suitable target for biofilm formation and maturation. SarA protein was chosen as target for this study for its interactions with other proteins in different pathways (Franceschini *et al.*, 2013).SarA protein of 25 strains of *S. epidermidis* were aligned to assign the most conservedregion.The WebLogo website was used to observe differences between the 25 sequences of the protein. SarA protein of the VCU144 strain of *S. epidermidis* was chosen as the target protein sequence (MAISKINDCFELLAMVTYADRLKGIIKKEFSISF EEFAVLTYISENKEEEYYLKDIINHL NYKQPQVVKAVKNLSQENYFNKKRNEHDERT VLILVDSKQRKKIDDLLKRVNNRITEANNENEV).
Results revealed that sarA is cytoplasmic protein, non-toxin and non-antigen.

Homology Modeling: unfortunately the sarA protein of *S. epidermidis* has not been crystallized yet, and no 2D NMR studies have been found for it. So, the next choice was to model it from the most identical protein. RaptorX website (Källberg *et al.*, 2012) was used for modeling the protein.
The protein was modeled by using 2fnpA and 2frhA proteins (codes of sarA proteins of different *S. aureus* strains in Protein Data Bank) as templates. The identity score was 85% and 90% respectively. Figure (1) shows the modeled protein in a PDB format.

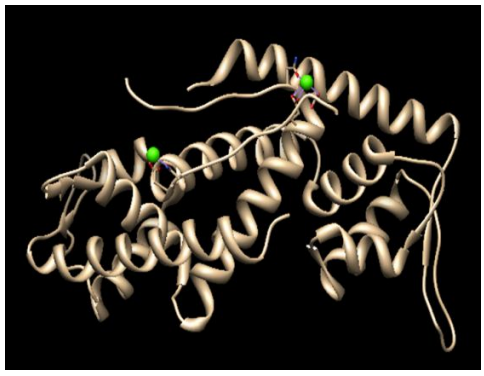
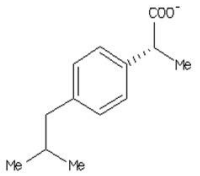


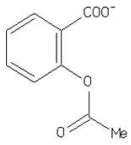
Figure (1) : SarA 3D Homolog of *S. epidermidis*
RaptorX website estimated two hypothetical structures. These results were analyzed by Qmean website for determining the reliability of the structure, where the first one gave 70% score and the other gave 66% score of reliability. The first model, then, was chosen as PDB format of the sarA protein .

Ligands Search: was carried out by survey the literatures and it has been found that 20 molecules from different groups can be used as initial hits (Acetaminophen, Acetic Acid, Albendazole, Acetylsalicylic Acid, Diacetyl, Eugenol, Piroxicam, Ibuprofen, Ferric ammonium citrate, Indomethacin,

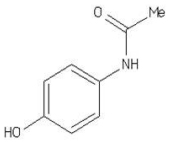
Levamisole, Methyldopa, Niclosamide, Pentazocine, Rifampicin, Thymol, Vancomycine, Diclofenac, (Z)-5-(bromomethylene)dihydrofuran-2(3H)-one, (E)-3-(bromomethylene)isobenzofuran-1(3H)-one). These molecules were tested for initial docking with the EADock DSS engine in the SwissDock Website. The result was positive (i.e., their minimum binding energy was negative) for only 8 molecules:



Acetaminophen

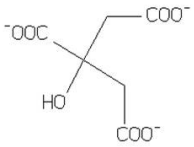


Acetylsalicylic acid

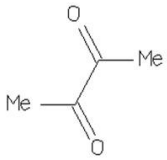


Ibuprofen

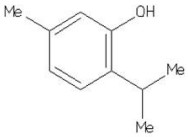
Acetic acid



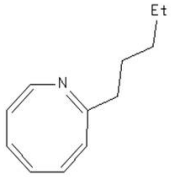
Ferric Ammonium Citrate



Diacetyl



Thymol



Pentazocine

Five molecules with the lowest binding free energy score were chosen for screening ligands (acetaminophen, ibuprofen, acetic acid, diacetyl, ferric ammonium citrate).

Pharmacophore Virtual Screening: the goal of virtual screening is to select, relatively rapidly and cheaply, small subset of compounds predicted to have activity against a given biological target out of a large database of compounds. The Hip Hop hypothesis in the Discovery Studio was used and gave 2 hypothysed pharmacophores (Figures 2 A and B).
Then these pharmacophores were used in the ZincPharmer to screen more than 180 million different conformations in Zinc database (Figure 3).

One hundred seventy seven (177) molecules were obtained from Zinc database after filtering through Lipinski rule of five and through analyzing them with the T.E.S.T program for their mutagenicity. All molecules wereused in the SwissDock to estimate their binding affinity to the protein. Many criteria from docking results can be used for estimating binding affinity including, binding free energy, full fitness, hydrogen bonding and total free energy. Binding free energy was used as the main criterion for ranking the best powerful ligands. The final result was 37 molecules having positive docking results after docking through SwissDock (i.e., having negative free binding energy with sarA protein) (Table 1).

DISCUSSION AND CONCLUSION

SarA protein was chosen at the beginningout ofmany targets as it acts as an *icaADBC* operon stimulator (regulator) (Tormo *et al.*, 2005) and hitting it will stop the most powerful operon in biofilm synthesis process. SarA is a 124-residues DNA binding protein encoded by the *sarA* locus, which consists of three overlapping transcripts, driven by three distinct promoters, P1, P3 and P2 (Bayer *et al.*, 1996).

DNA binding and profiling studies suggest that sarA protein may regulate target genes by directly binding to target gene promoters or indirectly via downstream effects on regulons (e.g. binding to the *agr* promoter) or by stabilizing mRNA during the log phase (Cheung *et al.*, 2004; Roberts *et al.*, 2006).

SarA, like its homolog sarR, is a dimeric winged helix structure with each monomer consisting of 5 α-helices, 3 β-strands and several loops (α1α2-β1α3α4-β2β3-α5). The sarA dimer possesses a central helical core and two winged helix motifs. Within each winged helix motif there is a helix-turn-helix motif (α3α4) and a β-hairpin turn wing (β2β3), both of

which are putative DNA binding domains (Liu *et al*., 2006).

The importance of this protein comes from its multifunctional regulatory activity. First of all, it acts at the initiation step of biofilm production by direct binding to *icaA* promoter enhancing transcription of *icaADBC* operon. Furthermore, sarA influences the regulation of biofilm formation via an agr-dependent pathway. It has also been found that sarA enhances the proteolytic enzymes activity, which has an important rule in the regulation of biofilm development (Tormo *et al*., 2005). So, blocking this protein will hit the biofilm development process at many stages.

Docking step was performed usingSwissDock which uses calculations performed in the CHARMM force field (Grosdidier*et al.*, 2011).

A pharmacophore model is an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response (Yang, 2010). In this study, Discovery Studio was used for pharmacophore modeling as it uses Hip Hop generator for this job.

Because sarA in *S.epidermidis* has no previous studies to be used as drug target, a new strategy was used to build a pharmacophore for this target. The first step was selecting twenty molecules acting as antibiofilm chosen from different antibiofilm categories. Then by initial docking with the target protein only eight molecules appear to bind with the protein. The five molecules with the best score selected from the initial docking step were used in Discovery Studio program to build a pharmacophore to be used in drugs like molecule screening. After building the pharmacophore it was entered to ZincPharmer website for virtual screening where more than 180 million conformations of Zinc database for small molecules were screened to give 177 molecules similar to the used pharmacophore. Then the resultedmolecules were tested with SwissDock at the final docking step to predict the most probable inhibitors for sarA protein.

The final result was 37 molecule with a negative free binding energy that means high affinity to bind to sarA protein and these molecules can be tested *in vitro* for their biofilm inhibition activity.

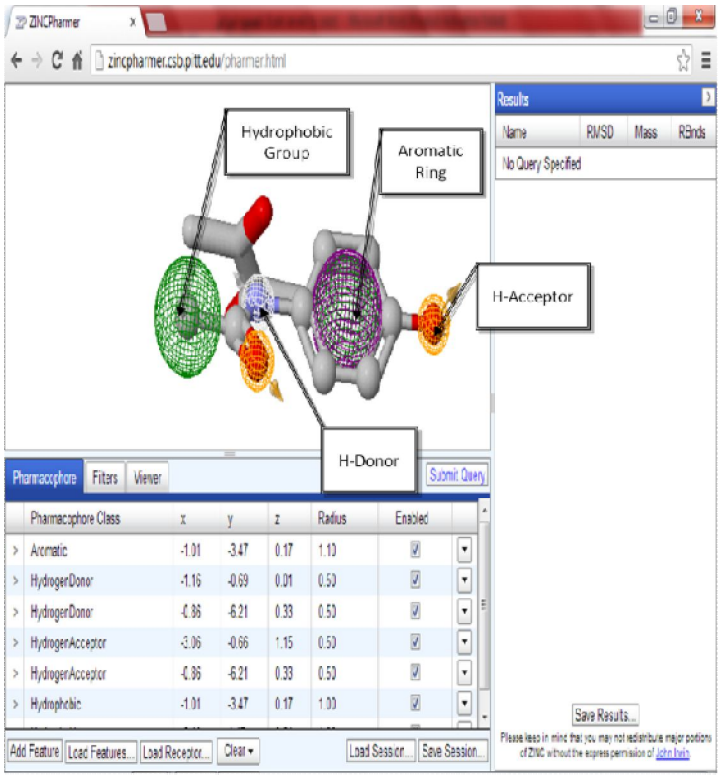


Figure (3) Pharmacophore features in ZincPharmer

Table (1) Molecules with positive docking result

	Molecules (chemical or zinc ID)	Minimum Free Binding Energy
1	Acetaminophene	-5.369
2	Acetic Acid	-7.239
3	Acetylsalicylic Acid	-0.0485
4	Diacetyl	-12.626
5	Ibuprofen	-11.998
6	Ferric ammonium citrate	-4.455
7	Pentazocine	-5.222
8	Thymol	-5.444
9	zinc_1768401	-4.065
10	zinc_11851832	-7.218
11	Zinc_254936	-6.810
12	Zinc_53792818	-19.662
13	Zinc_65054752	-2.869
14	Zinc_71763859	-13.979
15	Zinc_3074344	-11.67
16	Zinc_3077159	-8.5
17	Zinc_4675592	-6.29
18	Zinc_5957291	-8.17
19	Zinc_6702467	-7.61
20	Zinc_70735489	-7.83
21	Zinc_12496101	-5.393
22	Zinc_13545166	-14.183
23	Zinc_1454	-3.3
24	Zinc_1530959	-5.85
25	Zinc_1532514	-13.79
26	Zinc_1648334	-68.56
27	Zinc_1683666	-16.23
28	Zinc_2040136	-43.82
29	Zinc_4403799	-21.12
30	Zinc_5751050	-19.99
31	Zinc_1481956	-4.05
32	Zinc_1542916	-5.18
33	Zinc_1544545	-22.12
34	Zinc_1725270	-4.56
35	Zinc_18099446	-10
36	zinc_4097406	-22.3
37	zinc_4498304	-5.1

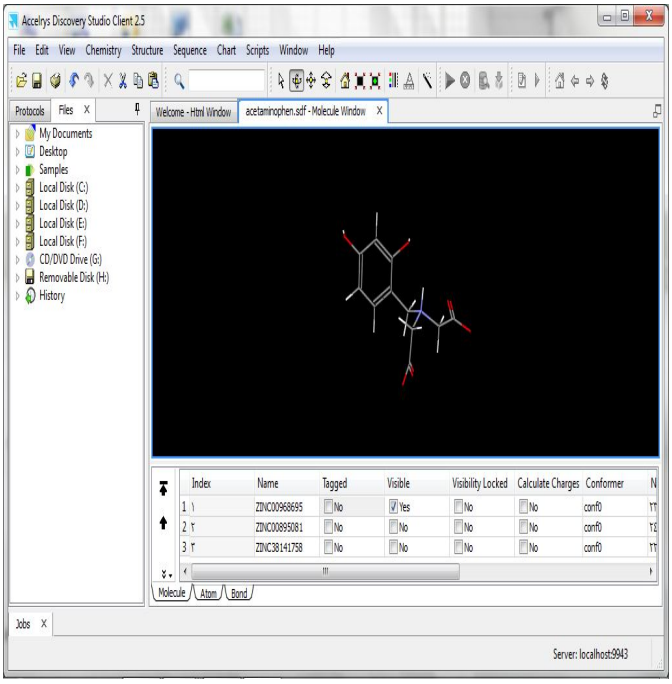


Fig A

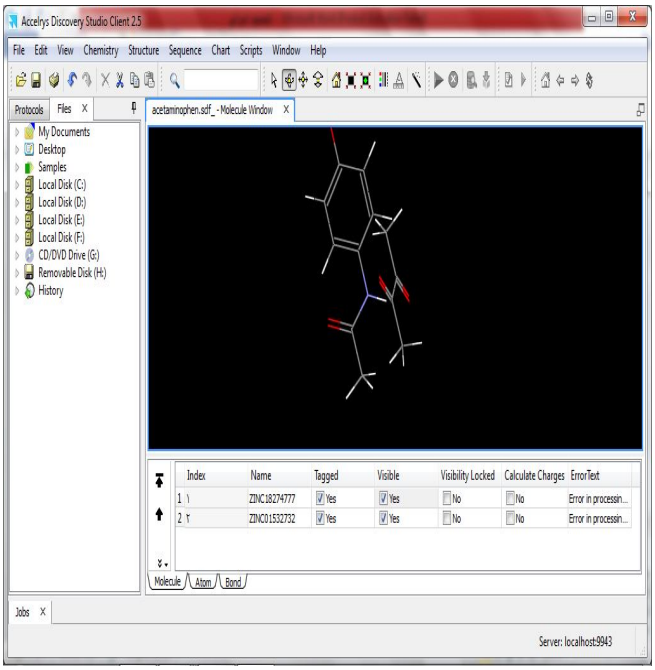


Fig B

Figure (2) A, B Hypothetical pharmacophores generated in Discovery Studio (redcolored groups represent H-acceptors while bluegroups represent H-donors)

References

1. Alem M and Douglas J. Effects of aspirin and other nonsteroidal anti-inflammatory drugs on biofilms and planktonic cells of Candida albicans. Antimicrobial Agents and Chemotherapy. 2004; 48:41-47.
2. Apweiler R, Bairoch A, Wu C, Barker W, Boeckmann B, Ferro S. UniProt: the Universal Protein knowledgebase. Nucleic Acids Research. 2004; 32:D115–D119.
3. Bayer M, Heinrichs J and Cheung A. The molecular architecture of the sar locus in Staphylococcus aureus. Journal of Bacteriology. 1996; 178:4563–4570.
4. Benkert P, Künzli M and Schwede T. QMEAN server for protein model quality estimation. Nucleic Acids Research. 2009; 37:W510-W514.
5. Bink A, Pellens K, Cammue B and Thevissen K. Anti-biofilm strategies: how to eradicate Candida biofilms? The Open Mycology Journal. 2011; 5:29-38.
6. Böhm H, Schneider G, Kubinyi H, Mannhold R and Timmerman H. Virtual Screening for Bioactive Molecules(Methods and Principles in Medicinal Chemistry). Wiley-VCH, Weinheim, Germany.2000.
7. Cheung A, Bayer A, Zhang G, Gresham H, Xiong Y. Regulation of virulence determinants in vitro and in vivo in Staphylococcus aureus. FEMS Microbiology Letters. 2004; 1649:1–9.
8. Doytchinova I and Flower D. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007; 8: 1-7.
9. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A. STRING v9.1: protein-protein interaction networks, with increased coverage and integration.Nucleic Acids Research. 2013; 41:D808-815.
10. Grosdidier A, Zoete V and Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Research. 2011; 39:270-277.
11. Hetal K, Ratna A and Pratibha B. Docking studies of components of tulsi and mamejavo against plasmodium lactate dehydrogenase. International Research Journal of Biological Sciences. 2013; 2:8–12.
12. Imperi F, Massai F, Pillai C, Longo F, Zennaro E, Rampioni G. New life for an old drug: the anthelmintic drug niclosamide inhibits Pseudomonas aeruginosa quorum sensing. Antimicrobial Agents and Chemotherapy. 2012; 57:996-1005.
13. Irwin J and Shoichet B. ZINC – a free database of commercially available compounds for virtual screening. Journal of Chemical Information and Modeling. 2005; 45:177–182.
14. Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H. Template-based protein structure modeling using the RaptorX web server. Nature Protocols. 2012; 7:1511–1522.
15. Koes D and Camacho C. ZINCPharmer: pharmacophore search of the ZINC database. Nucleic Acids Research. 2012; 40:W409-W414.
16. Liu Y, Manna A, Pan C, Kriksunov I, Thiel D, Cheung A. Structural and function analysis of the global regulatory protein SarA from Staphylococcus aureus. Proceedings of the National Academy of Sciences USA. 2006; 103:2392–2397.
17. Lönn-Stensrud J,Landin M, Benneche M, Petersen F and Scheie A. Furanones, potential agents for preventing Staphylococcus epidermidis biofilm infections? Journal of Antimicrobial Chemotherapy. 2009; 63:309– 316.
18. Matthiesen R. Bioinformatics Methods in Clinical Research. Humana Press, New York, USA. 2010.
19. Noori H and Spanagel R. In silico pharmacology: drug design and discovery's gate to the future. In Silico Pharmacology. 2013; 1:1-2.
20. Rao V and Srinivas K. Modern drug discovery process: An In silico approach. Journal of Bioinformatics and Sequence Analysis. 2011; 2:89-94.
21. Roberts C, Anderson K, Murphy E, Projan S, Mounts W, Hurlburt B. Characterizing the effect of the Staphylococcus aureus virulence factor regulator, SarA, on log-phase mRNA half-lives. Journal of Bacteriology. 2006; 188:2593–2603.
22. Saha S and Raghava G. BTXpred: prediction of bacterial toxins. In Silico Biology. 2007; 7:405-412.
23. Sushko I, Novotarskyi S, Körner R, Pandey A, Cherkasov A, Li J. Applicability domains for classification problems: benchmarking of distance to models for AMES mutagenicity set. Journal of Chemical Information and Modeling. 2010; 50:2094-2111.
24. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution. 2011; 28: 2731-2739.
25. Tormo M, Marti M, Valle J, Manna A, Cheung A, Lasa I. SarA is an essential positive regulator of

Staphylococcus epidermidis biofilm development. Journal of Bacteriology. 2005; 187:2348-2356.
26. Tsai K, Teng L, Shao Y, Chen Y, Lee Y, Li M. The first pharmacophore model for potent NF-kappaB inhibitors. Bioorganic & Medicinal Chemistry Letters. 2009; 19:5665-5669.
27. Yang S. Pharmacophore modeling and applications in drug discovery: challenges and recent advances. Drug Discovery Today. 2010; 15:444-450.
28. Young D. Computational Drug Design. John Wiley& Sons, New Jersey ,USA. 2009.
29. Yu C, Chen Y, Lu C and Hwang J. Prediction of protein subcellular localization. Proteins: Structure, Function and Bioinformatics. 2006; 64:643-651.