Article id: jbg.2017.2.4.1
Journal of Bioinformatics and Genomics, 2017, 2-4
doi: 10.18454/jbg.2017.2.4.1
Advance Access Publication Date: 15 August 2017

RESEARCH IN BIOLOGY USING COMPUTATION

AN IN-SILICO PERSPECTIVE TOWARDS TARGET ABILITY OF AVAILABLE DRUGS IN INFECTIOUS DISEASE TREATMENT: A POSSIBLE STRATEGY

Acknowledgement

This work was supported by Council of Scientific and Industrial Research, XII Five Year Plan Project "GENESIS" [Grant No. BSC0121]

Conflict of Interest

None declared.

Rupa Bhowmick [1,2], Noopur Sinha [1,2], Ram Rup Sarkar [1,2,*]

[1]CSIR-National Chemical Laboratory, Pune, Maharashtra, India-411008

[2]Academy of Scientific & Innovative Research (AcSIR), CSIR-NCL Campus, Pune, India-411008

*To whom correspondence should be addressed

Associate editor: Giancarlo Castellano

Received on 01 June 2017, revised on 26 June 2017, accepted on 13 July 2017.

Abstract

Motivation: *In early stage of therapeutics, several structure and ligand-based in-silico approaches have aided the modern drug discovery and design. However, such techniques are limited by availability of resolved 3D structures of targets and ligands. At the same time the growing concern of drug resistivity not only demands for new drugs but also the judicious use of presently available drugs. In such a scenario, the utilization of the already available drugs of a target molecule over the different homologous target of wider range of organisms is the better perspective for treatment. This requires confirmation of structural similarity of the targets (enzyme and protein targets) in those organisms.*

Results: *In the present study, based on the structural similarity of the target enzymes shared by different pathogenic micro-organisms, we have reviewed to gain an in-silico perspective of their efficacy in targeting a wider subset of organisms with few available drugs marketed for those. The results suggest efficient binding affinity of such drugs for the enzymes of organisms belonging to the cluster formed on the basis of structurally similarity.*

Implementation: *Such an approach can be adopted to utilize the presently available drugs for a wider range of pathogenic micro-organisms.*

Supplementary Information: *http://journal-biogen.org/article/view/58/25*

Keywords: *Infectious disease, Shikimic acid pathway, Drug target, Structural similarity, NiRMSD.*

Contact: *rr.sarkar@ncl.res.in*

1. Introduction

With the changing ways of human life, new diseases have emerged as a malady over the past few centuries. Some of the widespread epidemics, having a significant contemporary disease burden and huge impact on the global socio-economic structure, are tuberculosis, typhoid, cholera, diphtheria, whooping cough, pneumonia, etc. (Corbett *et al.,* 2003; Crump *et al.,* 2004; Lönnroth *et al.,* 2009; Crump and Mintz, 2010; Ryan, 2011, Jakinovich and Sood, 2014; Campbell and Nair, 2015). Estimated 13 million death cases are reported annually due to such infectious diseases (Cohen, 2000). On one hand, the use of modern antibiotics and drugs have been helpful, but on the other hand, increasing drug resistance amongst the pathogens have rendered many antibiotics and drugs ineffective in treating and controlling the disease progression. Epidemic antibiotic resistance has been reported in case of *Mycobacterium tuberculosis* (Strambu, 1999; Nachega and Chaisson, 2003; Nettleman, 2005; Zignol *et al.,* 2006), methicillin-resistant *Staphylococcus aureus* (MRSA) (Sattler *et al.,* 2002; Eady and Cove, 2003; Fridkin *et al.,* 2005; Miller *et al.,* 2005; Moran *et al.,* 2005), *Streptococcus pneumoniae* (Kays *et al.,* 2002; File, 2004; File, 2006), etc. Despite intensive efforts taken by the various global health organizations, multi-drug and antibiotic resistance has continued to be a global threat. The major concern although is that a majority of the common population remains unaware of

this crisis (Jha *et al.,* 2006). Efficacy of available drugs, associated side effects and the ability of selective targeting are few other factors, which limit effective treatment strategies. For example, 2.3 million adverse cases against about 6000 marketed drugs have been reported between 1969 till 2003. However, the trends state a slower rate of withdrawal of ~10 drugs per year, which instills the continued usage of harmful and ineffective drugs (Wysowski and Swartz, 2005).

Extensive research on pathways and mechanisms unique to the pathogens, have been carried out to develop drugs against such targets. In recent years, the study of pathways involved in folate synthesis in lower organism, specifically the Shikimic acid pathway, has gained attention as these pathways are exclusive to plants and lower organisms and are absent in the hosts. With the growing use of this and other metabolic pathways as a therapeutic target, increasing cases of drug resistance have also been reported. Insufficient research developments made in the field of drug and antibiotic development to combat the increasing strains of resistant species along with the enormous time taken for the lead molecules to clear different levels of clinical trials limit the ability to combat these resistant species. The statistics presented by DrugBank currently shows 8261 drugs available through various studies of which 2254 drugs are FDA approved (Law *et al.,* 2014). Conventional *in-silico* structure based drug discovery techniques can possibly reduce the time and effort required to screen down possible lead molecules from the huge pool of available drug and drug-like molecules. Present situation demands the most skillful utilization of the available knowledge about drugs to ensure effective targeting and to discover new therapeutics against the ever-increasing number of resistant species. A very appropriate statement in this context is by the Nobel laureate James Black who stated, "The most fruitful basis for the discovery of a new drug is to start with an old drug" (Chong and Sullivan, 2007).

2. Scope of the Review

In the present review, we have briefly discussed about the presently available and emerging computational approaches used in drug discovery. While briefly discussing about the benefits and drawbacks of these computational approaches, we have looked into the unique targets for the lower organisms, which have been exploited as therapeutic targets particularly focusing on the Shikimic acid pathway. We have corroborated a few analyses to the review, in order to check the structural relatedness of the enzymes of this pathway across different lower organisms. The choice of organisms has been made with *Mycobacterium tuberculosis* (MYCTU) as the reference organism and 23 orthologs of MYCTU have been selected, which have disease relevance (see Supplementary Table S1). The choice of *M. tuberculosis* as reference organisms is of particular importance to the analysis pertaining to its highly evolving multi-drug resistant nature, still claiming for an estimated 10.4 million new TB cases worldwide, as reported by WHO in 2015. An additional benefit is the availability of sequence and structural information at genomic and proteomic levels, as the organism remains thoroughly studied. Using different alignment tools, sequence and structural analysis of the seven core enzyme targets of Shikimic acid pathway, namely, 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DAH7PS), Dehydroquinate synthase (DHQS), 3-Dedydroquinate Dehydratase-Shikimate Dehydrogenase (DHQD), Shikimate dehydrogenase (SDH), Shikimate Kinase (SK), 5-enolpyruvate shikimate-3-phosphate synthase (EPSPS) and Chorismate Synthase (CS) have been performed for all the selected organisms (see Materials and Methods section, Supplementary file). The analysis has been further corroborated with *in-silico* drug binding efficiency of these targets with available drugs, to gain a perspective towards their usability for a wider range of organisms. Such an approach could be helpful in widening the therapeutic usage of already available drugs and also in understanding the mechanism of ligand binding of the targets helping in the development of new therapeutics.

3. Computational approaches available for preliminary drug discovery

In recent drug designing and discovery, computer-aided drug discovery/designing (CADD) methods have not only improved our understanding of drug-receptor interaction but have also assisted in identification of new therapeutic agents. The currently employed techniques are broadly classified as either structure-based or ligand-based techniques (Sliwoski *et al.,* 2014). The structure-based algorithms rely on the ability to determine and analyze the available 3D structure of the target biomolecules. The *in-silico* structure based approach comprises of various stages including structure identification or modeling of the target biomolecule, drug binding site identification/prediction, screening of suitable molecule (inhibitor/activator) for the receptor molecules, detailed study of the ligand-receptor binding using docking analysis and the final analyses of the dynamics behavior of the ligand-receptor interaction is observed by the Molecular dynamics study (Sinha *et al.,* 2016). Various popular tools and algorithms are available for the fulfillment of the desire objective of *in-silico* structure based method for example, comparative model building by satisfaction of spatial restraints obtained from the alignment of the target sequence with the template structures using Modeller software (Šali and Blundell, 1993), *in-silico* screening of inhibitor molecule from a large library of database followed by docking studies based on search based algorithms of Autodock (Morris *et al.,* 1998; Trott and Olson, 2010). A few more popular techniques are, Monte Carlo Metropolis minimization (Misura and Baker, 2005), ROSETTA (DiMaio *et al.,* 2009), LIGSITE (Hendlich *et al.,* 1997), I-TASSER (Zhang, 2008), Q-SITEFINDER (Laurie and Jackson, 2005), etc. Detailed atomic level dynamic behavior of the receptor-ligand interaction can be analyzed by force-field based methods of molecular dynamics (Raval *et al.,* 2012) using GROMACS (Hess *et al.,* 2008), NAMD (Phillips *et al.,* 2005), etc. Ligand-based algorithms, on the other hand, make use the knowledge of known agonist/antagonist molecules through chemical similarity searches or construction of predictive, Quantitative Structure-Activity Relation (QSAR). Ligand-based techniques are often used for identifying membrane protein targets, where little or no prior information of the target structures are available. Some of the popular ligand-based algorithms are AMANDA (Durán *et al.,* 2008), SVM-based QSAR models (Zhao *et al.,* 2006), CoMFA and CoMSIA 3D-QSAR methods (Bordás *et al.,* 2003). On the other hand, several other *in-silico* drug discovery and designing algorithms/approaches are also employed in pharmaceutical research. Multi-scale agent-based model (ABM) using Markov Chain Monte Carlo approaches is one of the model based drug discovery and development

approaches (Su *et al.,* 2014) where, drugs, metabolites, proteins and cells are considered as 'agents'. These agents are allowed to interact with each other according to the predefined set of rules. Another approach is proposed, which consider the undesired off-target effects of the drug in combinatorial therapies. In this approach the optimal combinatorial therapy problem is formulated into two complementary mathematical algorithms i.e. Balanced Target Set Cover (BTSC) and Minimum Off-Target Set Cover (MOTSC) and is solved by mixed integer linear programming (MILP) (Pang *et al.,* 2014). Moreover, the utility of structure and ligand based approaches are limited by the availability of the structural details of the targets and the ligands. Developing resolved 3D structures is a challenge for many proteins and is often difficult to obtain. In such scenarios, utilizing sequence information to obtain structural information can be helpful, as sequence information is easily attainable. In the next few sections, we will demonstrate the use of sequence information to obtain structural information with the help of a few analyses conducted on the seven core enzymes of the Shikimic Acid Pathway.

4. Shikimic Acid Pathway

Selective targeting of drugs is necessary to minimize the consequences of side effects on the host organism. As such, the use of targets those are exclusively available in the pathogen and are missing or non-functional in the host, have been emphasized. Presently available antibiotics act either on processes that are unique to bacteria, like the synthesis of cell walls or folic acid (Debono and Gordee, 1994; Kohanski *et al.,* 2007), or on bacterium-specific targets within processes that are common to both bacterium and human cells, including protein or DNA replication (Dahl and Rosenthal, 2008). Extensive research in this field has led to the identification of the enzymes of Shikimic acid pathway as potential drug targets. Shikimic acid pathway is a common route leading to production of the aromatic amino acids phenylalanine, tyrosine, and tryptophan (Figure 1) and links metabolism of carbohydrates to the biosynthesis of these aromatic compounds (Herrmann, 1995). A series of seven metabolic reactions convert phosphoenol pyruvate and erythrose-4-phosphate in chorismate, the precursor of the aromatic amino acids and many aromatic secondary metabolites. All the pathway intermediates can be considered as branch points to many other metabolic pathways, to which these intermediates act as a substrate (Bentley and Haslam, 1990). However, the pathway is found only in microorganisms and plants, and not in animals (Coggins *et al.,* 2003). Plants and bacteria exclusively depend on the Shikimic acid pathway for the synthesis of chorismate, which acts as a substrate for the production of p-amino benzoate and folates in plants and aromatic amino acids in bacteria. Mammals and many protists, on the other hand, rely exclusively on exogenous supply of folates and do not have a Shikimic acid pathway, which designates this as an important target for therapeutic interventions.

Fig 1 – Schematic diagram of Shikimic Acid Pathway

5. Structural and sequence analysis of Shikimic acid pathway enzymes

Structural similarity analysis was performed on different enzymes of the pathway amongst the selected set of organisms. For details of the technique see the subsection Selection of Organisms under Materials and Methods section in the Supplementary File. A hierarchical clustering has been shown on the basis of NiRMSD scores (Figure 2). The results have been discussed in next few sections and the description of the abbreviated form of the organisms has been provided in Table S1 of Supplementary file. For a particular enzyme, the set of organisms that tend to cluster together, are considered to have more structural similarity as compared to the other organisms.

5.1. NiRMSD Alignment Analysis

Table 1 shows the mean, median and standard deviation of the NiRMSD scores obtained for each pair of organism for each enzyme. With the least of mean, median and standard deviation for NiRMSD scores, the analysis predicts CS to be the most structurally conserved enzyme amongst all the seven enzymes across different species, followed by SDH and DHQS. Boxplots of the median values with the standard error bar has been given in Supplementary Figure S1. DHQD shows a maximum divergence in structure across different species.

The frequency distribution of the NiRMSD scores is shown in Supplementary Figure S2. As the data appears to be skewed, median value of each data set is chosen as the maximum (cut-off) NiRMSD score within which the clusters formed are considered to be highly structurally similar.

Table 1 – Mean, Median and Standard Deviation of the NiRMSD scores obtained for each enzyme.

Enzyme Name	Mean	Median	Standard Deviation
3-deoxy-D-arabino-heptulosonate 7-phosphate synthase (DAHPS)	3.1336	2.9	2.4701
*Dehydroquinate synthase (DHQS)	1.7747	1.55	1.0963
3-Dedydroquinate Dehydratase-Shikimate Dehydrogenase (DHQD)	7.9256	6.05	8.3769
*Shikimate dehydrogenase (SDH)	1.7869	1.69	0.9448
Shikimate Kinase (SK)	2.6222	1.97	2.1208
5-enolpyruvate shikimate-3-phosphate synthase (EPSPS)	2.5904	2.34	1.6431
*Chorismate Synthase (CS)	1.2163	1.17	0.9042

*These enzymes have relatively lower mean, median and standard deviation for NiRMSD score, and hence considered to have higher structural similarity across different organisms, than the rest of the enzymes.

5.1.1. 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DAH7PS)

Total three clusters are formed within the cut-off NiRMSD score of 2.9. DAHPS from *M. tuberculosis* strains: MYCTO and MYCTU and CORDI cluster together at a NiRMSD score of 2.42, showing close structural resemblance. Another cluster is formed by the two species of *Helicobacter*: HELPH and HELHP at a NiRMSD score of 2.206. The third cluster with a very low NiRMSD difference of 1.635 is formed by BORPE and HAEIN suggesting a very high structural similarity of the enzyme in these two organisms (Figure 2A).

5.1.2. Dehydroquinate synthase (DHQS)

Hierarchical clustering for the enzyme DHQS shows that all the organisms tend to cluster within a very low NiRMSD score of about 8. This suggests that the enzyme remains highly conserved across the species. Within a cut-off of 1.55, three clusters are formed. The two strains of *Listeria monocytogenes*: LISMO and LISMF group together at NiRMSD score of 1.478. ECOLI, SHIDS, SALTI, HAEIN and VIBCH are clustered together at NiRMSD score of 1.184. The third set is formed by the two strains of *Helicobacter pylori*: HELPH and strain HELPY, which cluster together at a NiRMSD score of 0, suggesting the enzyme present in both the strains is structurally the same. However, another species of *Helicobacter* genus: *H. hepaticus* (HELHP) considered in the analysis, showed a significant difference in the NiRMSD score from the rest of the two strains of the same genus. It can be observed from the analysis that enzymes from the organisms belonging to same genus might have structural differences. At the same time, enzymes from distant genus and species might have a greater structural similarity than enzyme from organisms belonging to the same genus (Figure 2B).

5.1.3. 3-Dedydroquinate Dehydratase-Shikimate Dehydrogenase (DHQD)

Pairwise NiRMSD score shows a huge variation for the enzyme DHQD implying that the enzyme structure varies a lot across different species. The median for NiRMSD score for the enzyme lies at 6.05. Three clusters are formed within this NiRMSD score, although this cut-off is much higher than that formed for any other enzyme. ACIBA, HELPY, HELPH, MYCTU, HAEIN, CORDI and VIBCH are clustered together at NiRMSD score of 4.513. Two strains of *H. pylori*: HELPY, HELPH and along with *H. hepaticus* strain HELHP were considered for this analysis. However, the clustering pattern shown by these three strains differs significantly. HELPH and HELPY fall into the same cluster, but HELHP shows a significant difference in structure which branches out of the other two strains at a NiRMSD score of 10.38. ECOLI and SHIDY group together at NiRMSD score of 4.352. Another cluster is formed by the two strains of *L. monocytogenes*: LISMO and LISMF, at NiRMSD score of 0.1034 (Figure 2C).

5.1.4. Shikimate dehydrogenase (SDH)

With a very low mean, median and standard deviation for the NiRMSD scores generated, SDH is predicted to be one of the most structurally conserved enzymes among the seven enzymes of Shikimic acid pathway, across different species. Within the median cut-off NiRMSD score of 1.69, three clusters are formed. The first cluster is formed at NiRMSD score of 1.657, which includes ECOLI, SALTI, HAEIN and VIBCH. BORPE and NEIGO form the second cluster at a score of 0.6719, implying a high structural similarity of the enzyme in the two organisms. The two strains of *H. pylori*:

HELPY and HELPH form the third cluster at 0.04. The *H. hepaticus* strain, HELHP separates out from the other two strains of the same genus, as is the case for DHQD enzyme, suggesting structural differences in the enzyme SDH of this strain with that of the other two strains of *H. pylori* even though belonging to the same genus (Figure 2D).

5.1.5. Shikimate Kinase (SK)

Two clusters are formed within the cut-off NiRMSD score of 1.97. The first one is formed by the two strains of *H. pylori*: HELPY and HELPH at a score of 0.4796. *H. hepaticus* strain, HELHP is only distantly related to these two strains at NiRMSD score of 6.928. The second cluster is formed by ECOLI and SHIDS. These two are predicted show a very high structural similarity as they group together at a NiRMSD score of 0.00 (Figure 2E).

5.1.6. 5-enolpyruvate shikimate-3-phosphate synthase (EPSPS)

Four clusters are formed for EPSPS within the cut-off score of 2.335. ACIBA and STRR6 form the first cluster at NiRMSD score of 2.107. SALTI, SHIDY, ECOLI, SHIDS, VIBCH, HAEIN and BORPE form the second cluster at 1.732. The two strains of *H. pylori*: HELPY, HELPH group together at a score of 1.147. HELHP shows a similar trend of clustering as is for DHQD, SDH and SK and is distantly related to the other two strains at NiRMSD score of 3.136. Fourth cluster is formed by the two strains of *L. monocytogenes*: LISMO and LISMF which group together at NiRMSD score of 0.0245 signifying close structural relatedness amongst the two strains of the species (Figure 2F).

5.1.7. Chorismate Synthase (CS)

With the least of mean, median and standard deviation, CS shows the maximum structural conservedness across the different species considered amongst all the seven enzymes of the Shikimic acid pathway. The median cut-off NiRMSD score for CS is 1.17. The clusters that are grouped within this score are predicted to have a significantly high structural relatedness. Four clusters are formed within the cut-off. ACIBA, ECOLI, SALTI, SHIDS, HAEIN, VIBCH and BORPE form the first cluster at 1.017 followed by CORDI and MYCTU clustering at a score of 0.8453. Two separate clusters are formed by *L. monocytogenes*: LISMO and LISMF and STRR6 at score 0.00 and the two strains of *H. pylori*: HELPY and HELPH again at score 0.00. This pattern might appear trivial except for the observation that the strains of the same species do not always tend to group together. HELHP, which is considered in the clustering analysis of CS, also shows a similar trend to separate out from the two other strains as is observed for DHQD, SDH, SK and EPSPS. With a very small NiRMSD score for clustering, the organisms belonging to each of the clusters are predicted to have high structural similarity (Fig. 2G).

Interestingly, through these clustering analyses, it could be observed that even though yeast (YEAST) is distantly related to bacteria in the phylogeny, the structural relatedness of the enzymes of Shikimate pathway is considerably high among them. On the other hand, the structures are so different in case of *Plasmodium falciparum* (PLAFA/PLAF7), that it is completely eliminated out of the structural hierarchy with the bacteria. The different strains of same species of bacteria tend to cluster together due to their close structural resemblance.

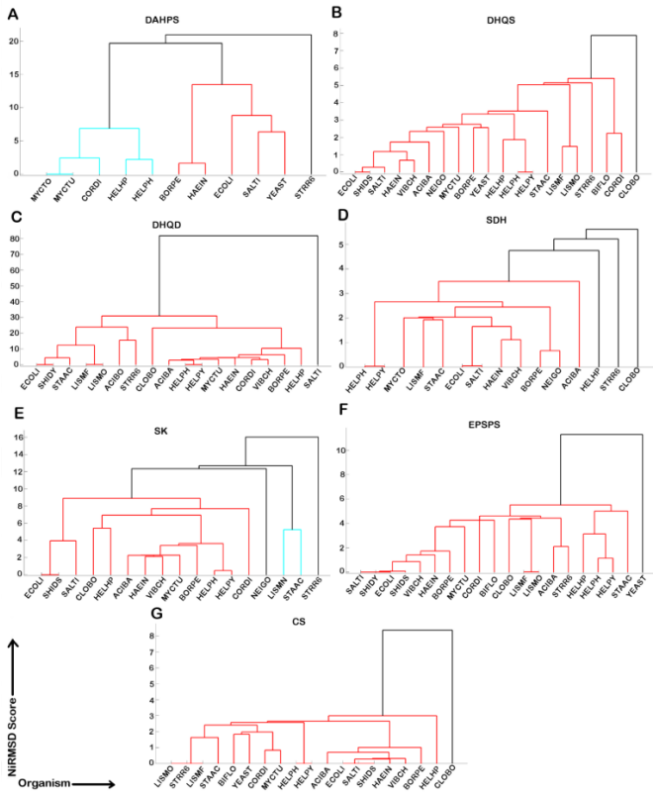

Fig 2 – NiRMSD distance based hierarchical clustering of organism for the seven enzymes of Shimikate pathway. Clusters formed for: (A) 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase (DAH7PS), (B) Dehydroquinate synthase (DHQS), (C) 3-Dedydroquinate Dehydratase-Shikimate Dehydrogenase (DHQD), (D) Shikimate dehydrogenase (SDH), (E) Shikimate Kinase (SK), (F) 5-enolpyruvate shikimate-3-phosphate synthase (EPSPS) and (G) Chorismate Synthase (CS).

5.2. Structural relatedness

To demonstrate the structural similarity amongst the clustered organism based on their NiRMSD scores, we have further illustrated our work by comparing the available PDB structure of enzymes for the clustered organisms. For this we have chosen the organism clustered within the median cut-off NiRMSD score 1.017 for the enzyme CS. This cluster shows the maximum number of diverse organism and is further analyzed to evaluate the structure relatedness with each other. The UniProt sequences of the clustered organism are used for the homologous 3D structure search against PDB using protein BLAST and are summarized in the Supplementary Table S2. The obtained result illustrates that 3D structure is resolved and deposited only for the CS of ACIBA with PDB ID 4LJ2 (DOI: 10.2210/pdb4lj2/pdb) among the entire cluster of organisms. For the remaining organisms of the cluster, i.e. ECOLI, SALTI, SHIDS, HAEIN, VIBCH and BORPE, the structures are not resolved and are not available in the PDB entries. Meanwhile, for all these remaining organisms, ACIBA (PDB ID: 4LJ2) is showing maximum sequence similarity with identity greater than 60%.

Additionally, the study has also included the analysis of the availability of drugs/inhibitors targeting CS enzyme of different organisms of the cluster. Supplementary Table S2 explains the related diseases of the clustered organism along with the available drugs for CS of each organism. It is found that among all seven clustered organism, drugs are only available for ECOLI and SALTI (Pitchandi *et al.*, 2013) that targets CS. Hence, based on the structural relatedness analysis one can consider to study the effect of already available effective drug/inhibitor of CS of one organism to other related organism.

5.3. Analysis of Binding Site

In order to evaluate the response of an inhibitor or a drug of one organism over the same enzyme of other organisms, identification of functionally important inhibitor binding site is very important. The site is evaluated with the help of prior knowledge of conserved binding site of the homologous protein/enzyme. The experimentally solved 3D structure of CS from STRR6 with 2.0 A˚ of resolution (PDB ID "1QX0"), with a known EPSP binding site, has been considered for binding site identification (Barcellos *et al.*, 2009). The sequence of ACIBA, ECOLI and SALTI are aligned with the sequence of STRR6 and are showed in Supplementary Figure S3. The Percent Identity Matrix obtained shows that STRR6 aligned with ACIBA, ECOLI and SALTI with 35.98%, 37.89% and 37.04% identity, respectively**.** Meanwhile, most of the conserved EPSP binding sites residues of STRR6 are common for CS from other aligned organism (see Supplementary Figure S3**).** The highlighted residues are the EPSP binding site residues of STRR6. From the obtained alignment result it has been observed that EPSP binding site is almost conserved in all the four organisms, except for two residue mismatches for ACIBA and one mismatch for ECOLI

and SALTI with respect to the known STRR6 EPSP binding site. The highlighted residues are considered as the binding site for EPSP in ACIBA to check the effectiveness of the already available drugs of ECOLI and SALTI.

5.4. Investigation of Drug Binding

Review of available drugs revealed that, among the clustered organism, drugs are only available for ECOLI and SALTI. Meanwhile, the 3D structure is only available for ACIBA. Hence, to analyze the effect of available drugs of CS from one organism to other organism the PDB structure of ACIBA (PDB ID: 4LJ2) is considered as a receptor to perform the docking study. Three reported drugs are tested for binding with the ACIBA i.e. the only drug of CS from ECOLI ((6R)-6-Fluoro-EPSP) and two different drugs used for the treatment of Typhoid Fever caused by SALTI (2-(3-(-(S)-5-((S)-1-Amino-3-(3-chlorophenyl)-1-oxopropan-2-ylamino)-4-(3-hydroxy-4-methyl-2-nitrobenzamido)-5-oxopentylcarbamoyl)-phenoxy)acetic acid and N-((S)-1-((S)-1-Amino-3-(3-chlorophenyl)-1-oxopropan-2-ylamino)-1-oxopropan-2-yl)-3-hydroxy-4-methyl-2 nitrobenzamide). Detailed descriptions of the techniques are explained in the Supplementary Materials and Methods subsection "Molecular Docking of Available Inhibitors".

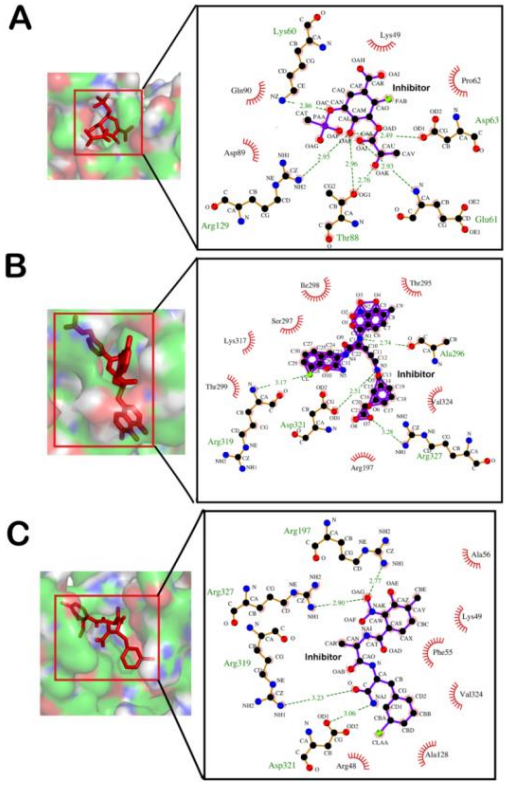

Fig 3 – Docking analysis of Chorismate Synthase of *A. baumannii* with (A) *E. coli* inhibitor Pubchem CID: 46926329, (B) Inhibitor of *S. typhi*, Pubchem CID: 46221610 and, (C) Inhibitor of *S. typhi*, Pubchem CID: 46221613. Dotted lines represent the hydrogen bonding between the residues of the inhibitors and the enzyme along with their bond length.

The result illustrates efficient binding of all the three inhibitors considered with ACIBA, with respect to the binding energy and interacting residues. The detailed results are explained in Figure 3 and Supplementary Table S3**.** Figure 3 shows the interacting residues (hydrogen bonds and hydrophobic interactions) of the enzyme and the inhibitors. Supplementary Table S3 includes the name of the inhibitors, Binding Energy, Hydrogen bonding residues along with the bond length and the hydrophobic interaction observed between inhibitor and the protein.

6. Discussion

In the present work, with the help of modern *in-silico* techniques, we have tried to explore an approach towards extending the utilization of the already available drugs into a wider range of organisms. The use of such techniques is helpful in preliminary drug discovery, as they narrow down the range of search from a huge pool of lead molecules. However, limitation of such approaches is the availability of structural details for the targets. Hence, we have tried to demonstrate the usage of available structural and sequence information to predict structural similarities of targets in different organisms, using the core Shikimic acid pathway enzymes as targets. 23 orthologs of *M. tuberculosis*, all having disease relevance, have been chosen. The selected organisms have been hierarchically clustered on the basis of their NiRMSD scores for all the seven enzymes of Shikimic acid pathway. The organisms, which cluster together within the decided cut-off, have been considered to be structurally similar. As a demonstration of feasibility of the present objective, we have chosen a subset of organisms, which clustered together within a cut-off of 1.17 for chorismate synthase, and have tried to check the drug binding efficiency of the enzyme within this subset. The subset includes *Acinetobacter baumannii* ATCC 19606 (ACIBA),

Escherichia coli (strain K12) (ECOLI), *Salmonella typhi* (SALTI), *Shigella dysenteriae* (strain Sd197) (SHIDS), *Haemophilus influenzae* (strain ATCC 51907) (HAEIN), *Vibrio cholerae* (strain ATCC 39315) (VIBCH) and *Bordetella pertussis* (strain ATCC BAA-589) (BORPE) clustering at NiRMSD score of 1.017. To verify of the structural similarity exhibited by this subset of organisms, an alignment of the Uniprot sequences of these organisms was performed by a BLAST analysis against PDB. Results revealed the availability of only one resolved structure for chorismate synthase of *A. baumannii* (PDB ID 4LJ2). BLAST results also revealed that *A. baumannii* has maximum sequence similarity with all other organisms of the cluster with identity greater than 60%. A thorough review of the available drugs against chorismate synthase in these organisms has been done (see Supplementary Table S2), which showed the availability of drugs only against *E. coli* and *S. typhi*. With the identification of probable binding site in the structure of chorismate synthase for *A. baumannii* (with identification of conserved residues as available for EPSP-binding site in the chorismate synthase of *S. pneumoniae*) efficiency of binding of the drugs available for *E. coli* and *S. typhi* have been checked using molecular docking. The results showed efficient drug binding to the available structure with a minimum binding energy of -9.5 Kcal/mol for the inhibitor (6R)-6-Fluoro-EPSP against chorismate synthase of *E. coli* and -9.12 Kcal/mol and -8.51 Kcal/mol for the inhibitors 2-(3-(-(S)-5-((S)-1-Amino-3-(3-chlorophenyl)-1-oxopropan-2-ylamino)-4-(3-hydroxy-4-methyl-2-nitrobenzamido)-5-oxopentylcarbamoyl)-phenoxy)acetic acid and N-((S)-1-((S)-1-Amino-3-(3-chlorophenyl)-1-oxopropan-2-ylamino)-1-oxopropan-2-yl)-3-hydroxy-4-methyl-2 nitrobenzamide against chorismate synthase of *S. typhi* (see Figure 3 and Supplementary Table S3).

Broadening the range of use of the presently available therapeutics will largely reduce the time and effort of *de novo* drug designing and discovery. The present requirement is to make the maximum use of the available knowledge and information of the presently available therapeutic agents. The use of *in-silico* techniques, such as the one proposed through the present review and analyses will be helpful. The benefit of such an approach is not limited by the availability of resolved 3D structural details of the targets. Addressing the drug resistant species requires designing and development of new drugs and antibiotics. However, the utility of the presently available drugs cannot be overlooked. The possible use of these drugs can be widened, by testing their efficacy over other related organisms. Hence, in the present work, we have tried to demonstrate a possible utilization of the available information to make use of the presently available drugs to a wider range of organisms so as to provide a new direction to this research area.

References

Barcellos, G. B., Caceres, R. A., and de Azevedo, W. F. (2009). Structural studies of shikimate dehydrogenase from Bacillus anthracis complexed with cofactor NADP. *Journal of Molecular Modeling*, 15(2): 147, doi: 10.1007/s00894-008-0403-z

Bentley, R. and Haslam, E. (1990). The shikimate pathway—a metabolic tree with many branche. *Critical Reviews in Biochemistry and Molecular Biology*, 25(5): 307-384, doi: 10.3109/10409239009090615

Bordás, B., Kőmíves, T. and Lopata, A. (2003). Ligand-based computer-aided pesticide design. A review of applications of the CoMFA and CoMSIA methodologies. *Pest Management Science*, 59(4): 393-400, doi: 10.1002/ps.614

Campbell, H. and Nair, H. (2015). Child pneumonia at a time of epidemiological transition. *The Lancet Global Health*, 3(2): e65-e66, doi: 10.1016/S2214-109X(14)70308-0

Chong, C. R. and Sullivan, D. J. (2007). New uses for old drugs. *Nature*, 448(7154): 645-646, doi:10.1038/448645a

Coggins, J., Abell, C., Evans, L., Frederickson, M., Robinson, D., Roszak, A. and Lapthorn, A. (2003). Experiences with the shikimate-pathway enzymes as targets for rational drug design. *Portland Press Limited*, 31 (3) 548-552, doi: 10.1042/bst0310548

Cohen, M. L. (2000). Changing patterns of infectious disease. *Nature*, 406(6797): 762-767, doi:10.1038/35021206

Corbett, E. L., Watt, C. J., Walker, N., Maher, D., Williams, B. G., Raviglione, M. C. and Dye, C. (2003). The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Archives of Internal Medicine*, 163(9): 1009-1021, doi:10.1001/archinte.163.9.1009

Crump, J. A., Luby, S. P. and Mintz, E. D. (2004). The global burden of typhoid fever. *Bulletin of the World Health Organization*, 82(5): 346-353, doi: 10.1590/S0042-96862004000500008

Crump, J. A. and Mintz, E. D. (2010). Global trends in typhoid and paratyphoid fever. *Clinical Infectious Diseases*, 50(2): 241-246, doi: 10.1086/649541

Dahl, E. L. and Rosenthal, P. J. (2008). Apicoplast translation, transcription and genome replication: targets for antimalarial antibiotics. *Trends in Parasitology*, 24(6): 279-284, doi: 10.1016/j.pt.2008.03.007

Debono, M. and Gordee, R. S. (1994). Antibiotics that inhibit fungal cell wall development. *Annual Reviews in Microbiology*, 48(1): 471-497, doi: 10.1146/annurev.mi.48.100194.002351

DiMaio, F., Tyka, M. D., Baker, M. L., Chiu, W. and Baker, D. (2009). Refinement of protein structures into low-resolution density maps using rosetta. *Journal of Molecular Biology*, 392(1): 181-190, doi:10.1016/j.jmb.2009.07.008

Durán, Á., Martínez, G. C. and Pastor, M. (2008). Development and validation of AMANDA, a new algorithm for selecting highly relevant regions in molecular interaction fields. *Journal of Chemical Information and Modeling*, 48(9): 1813-1823, doi: 10.1021/ci800037t

Eady, E. A. and Cove, J. H. (2003). Staphylococcal resistance revisited: community-acquired methicillin resistant Staphylococcus aureus-an emerging problem for the management of skin and soft tissue infections. *Current Opinion in Infectious Diseases*, 16(2): 103-124, doi: 10.1097/01.aco.0000065071.06965.ca

File, T. M. (2006). Clinical implications and treatment of multiresistant Streptococcus pneumoniae pneumonia. *Clinical Microbiology and Infection*, 12: 31-41, doi: 10.1111/j.1469-0691.2006.01395.x

File, T. M. (2004). Streptococcus pneumoniae and community-acquired pneumonia: a cause for concern. *The American Journal of Medicine Supplements*, 117(3): 39-50, doi: 10.1016/j.amjmed.2004.07.007

Fridkin, S. K., Hageman, J. C., Morrison, M., Sanza, L. T., Como-Sabetti, K., Jernigan, J. A., Harriman, K., Harrison, L. H., Lynfield, R. and Farley, M. M. (2005). Methicillin-resistant Staphylococcus aureus disease in three communities. *New England Journal of Medicine*, 352(14): 1436-1444, doi: 10.1056/NEJMoa043252

Hendlich, M., Rippmann, F. and Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *Journal of Molecular Graphics and Modelling*, 15(6): 359-363, doi: 10.1016/S1093-3263(98)00002-3

Herrmann, K. M. (1995). The shikimate pathway: early steps in the biosynthesis of aromatic compounds. *The Plant Cell*, 7(7): 907, doi: 10.1105/tpc.7.7.907

Hess, B., Kutzner , C., Van Der Spoel, D. and Lindahl, E. (2008). GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3): 435-447, doi: 10.1021/ct700301q

Jakinovich, A. and Sood, S. K. (2014). Pertussis: still a cause of death, seven decades into vaccination. *Current Opinion in Pediatrics*, 26(5): 597-604, doi: 10.1021/ct700301q

Jha, P., Chaloupka, F. J., Moore, J., Gajalakshmi, V., Gupta, P. C., Peck, R., Jamison, D., Breman, J., Measham, A. and Alleyne, G. (2006). Disease control priorities in developing countries. *Disease Control Priorities in Developing Countries*, doi: 10.2105/AJPH.81.1.15

Kays, M. B., Smith, D. W., Wack, M. F. and. Denys, G. A (2002). Levofloxacin treatment failure in a patient with fluoroquinolone-resistant Streptococcus pneumoniae pneumonia. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 22(3): 395-399, doi: 10.1592/phco.22.5.395.33185

Kohanski, M. A., Dwyer, D. J., Hayete, B., Lawrence, C. A. and Collins, J. J. (2007). A common mechanism of cellular death induced by bactericidal antibiotics. *Cell*, 130(5): 797-810, doi: 10.1016/j.cell.2007.06.049

Laurie, A. T. and Jackson, R. M. (2005). Q-SiteFinder: an energy-based method for the prediction of protein–ligand binding sites. *Bioinformatics*, 21(9): 1908-1916, doi: 10.1093/bioinformatics/bti315

Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., Maciejewski, A., Arndt, D., Wilson, M. and Neveu, V. (2014). DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Research*, 42(D1): D1091-D1097, doi: 10.1093/nar/gkt1068

Lönnroth, K., Jaramillo, E., Williams, B. G., Dye, C. and Raviglione, M. (2009). Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Social Science & Medicine*, 68(12): 2240-2246, doi: 10.1016/j.socscimed.2009.03.041

Miller, L. G., Perdreau-Remington, F., Rieg, G., Mehdi, S., Perlroth, J., Bayer, A. S., Tang, A. W., Phung, T. O. and Spellberg, B. (2005). Necrotizing fasciitis caused by community-associated methicillin-resistant Staphylococcus aureus in Los Angeles. *New England Journal of Medicine*, 352(14): 1445-1453, doi: 10.1056/NEJMoa042683

Misura, K. and Baker, D. (2005). Progress and challenges in high-resolution refinement of protein structure models. *Proteins: Structure, Function, and Bioinformatics*, 59(1): 15-29, doi: 10.1002/prot.20376

Moran, G. J., Amii, R. N., Abrahamian, F. M. and Talan, D. A. (2005). Methicillin-resistant Staphylococcus aureus in community-acquired skin infections. *Emerging Infectious Diseases*, 11(6): 928, doi: 10.3201/eid1106.040641

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 19(14): 1639-1662, doi: 10.1.1.113.6936

Nachega, J. B. and Chaisson, R. E. (2003). Tuberculosis drug resistance: a global threat. *Clinical Infectious Diseases*, 36(Supplement 1): S24-S30, doi: 10.1086/344657

Nettleman, M. D. (2005). Multidrug-resistant tuberculosis: news from the front. *Jama*, 293(22): 2788-2790, doi: 10.1001/jama.293.22.2788

Pang, K., Wan, Y. W., Choi, W. T., Donehower, L. A., Sun, J., Pant, D. and Liu, Z. (2014). Combinatorial therapy discovery using mixed integer linear programming. *Bioinformatics*, 30(10): 1456-1463, doi: 10.1093/bioinformatics/btu046

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L.and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16): 1781-1802, doi: 10.1002/jcc.20289

Pitchandi, P., Hopper, W. and Rao, R. (2013). Comprehensive database of Chorismate synthase enzyme from shikimate pathway in pathogenic bacteria. *BMC Pharmacology and Toxicology*, 14(1): 29, doi: 10.1186/2050-6511-14-29

Raval, A., Piana, S., Eastwood, M. P., Dror, R. O. and Shaw, D. E. (2012). Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins: Structure, Function, and Bioinformatics*, 80(8): 2071-2079, doi: 10.1002/prot.24098

Ryan, E. T. (2011). The cholera pandemic, still with us after half a century: time to rethink. *PLOS Neglected Tropical Diseases*, 5(1): e1003, doi: 10.1371/journal.pntd.0001003

Šali, A. and Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, 234(3): 779-815, doi: 10.1006/jmbi.1993.1626

Sattler, C. A., MASON JR, E. O. and Kaplan, S. L. (2002). Prospective comparison of risk factors and demographic and clinical characteristics of community-acquired, methicillin-resistant versus methicillin-susceptible Staphylococcus aureus infection in children. *The Pediatric Ifectious Disease Journal*, 21(10): 910-916, doi: 0.1097/01.inf.0000034268.59915.6f

Sinha, N., Chowdhury, S. and Sarkar, R. R. (2016). Structural Insight of NICD-MAML Interactions: Virtual Screening, Docking and Molecular Dynamics Study for Identification of Potential Inhibitor. *Letters in Drug Design & Discovery*,13(4): 301-313, doi: 10.2174/1570180812666150819003634

Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational methods in drug discovery. *Pharmacological Reviews*, 66(1), 334-395, doi: 10.1124/pr.112.007336

Strambu, I. (1999). Therapeutic pneumothorax-an effective adjuvant method in treating multidrug-resistant tuberculosis. *Pneumologia (Bucharest, Romania)*, 49(2): 129-136

Su, J., Zhang, L., Zhang, W., Choi, D. S., Wen, J., Jiang, B., Chang, C. C. and Zhou, X. (2014). Targeting the Biophysical Properties of the Myeloma Initiating Cell Niches: A Pharmaceutical Synergism Analysis Using Multi-Scale Agent-Based Modeling. *PLoS ONE*, 9(1): e85059, doi:10.1371/journal.pone.0085059

Trott, O. and Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2): 455-461, doi: 10.1002/jcc.21334

Wysowski, D. K. and Swartz, L. (2005). Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. *Archives of Internal Medicine*, 165(12): 1363-1369, doi: 10.1001/archinte.165.12.1363

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, 9(1): 40, doi: 10.1186/1471-2105-9-40

Zhao, C., Zhang, H., Zhang, X., Liu, M., Hu, Z. and Fan, B. (2006). Application of support vector machine (SVM) for prediction toxic activity of different data sets. *Toxicology*, 217(2): 105-119, doi: 10.1016/j.tox.2005.08.019

Zignol, M., Hosseini, M. S., Wright, A., Weezenbeek, C. L. v., Nunn, P., Watt, C. J., Williams, B. G. and Dye, C. (2006). Global incidence of multidrug-resistant tuberculosis. *The Journal of Infectious Diseases*, 194(4): 479-485, doi: 10.1086/5058

ПЕРСПЕКТИВА ЭЛЕКТРОННОЙ ПЦР С ТОЧКИ ЗРЕНИЯ ЦЕЛЕВОГО ВОЗДЕЙСТВИЯ СУЩЕСТВУЮЩИХ ПРЕПАРАТОВ ПРИ ЛЕЧЕНИИ ИНФЕКЦИОННЫХ ЗАБОЛЕВАНИЙ: ВОЗМОЖНАЯ СТРАТЕГИЯ

Благодарности

Эта работа была осуществлена при поддержке Совета по научным и промышленным исследованиям (CSIR), XII пятилетний план проекта «ГЕНЕЗИС» ("GENESIS") [грант № BSC0121].

Конфликт интересов

Не указан.

Рупа Бхомвик (Rupa Bhowmick) [1,2], Ноопур Синха (Noopur Sinha) [1,2],

Рам Руп Саркар (Ram Rup Sarkar) [1,2,*]

[1]CSIR-Национальная химическая лаборатория (NCL), Пуна, Махараштра, Индия-411008

[2]Академия научных и инновационных исследований (AcSIR), Кампус CSIR-NCL, Пуна, Индия-411008

*Корреспондирующий автор.

Редактор: Джанкарло Кастельяно

Получена 01 Июня 2017, доработана 26 Июня 2017, принята 13 Июля 2017.

Аннотация

Мотивация: *На раннем этапе развития клинической медицины несколько подходов электронной ПЦР на основе структуры и лигандов способствовали открытию и разработке современных лекарств. Тем не менее, такие методы ограничены наличием выделяемых трехмерных структур целей и лигандов. В то же время растущая проблема резистентности к лекарственным препаратам требует не только создания новых, но и рационального использования уже существующих препаратов. В такой ситуации использование уже существующих препаратов, нацеленных на определенные молекулы, для воздействия на другие гомологичные молекулы более широкого спектра организмов представляется более перспекивным для лечения. Для этого требуется подтверждение структурного сходства мишеней (ферментов и белков-мишеней) в этих организмах.*

Результаты: *В настоящем исследовании на основе структурного сходства ферментов-мишеней у различных патогенных микроорганизмов, мы рассмотрели получение электронной модели эффективности нескольких имеющихся лекарственных препаратов в отношении более широкой группы организмов. Результаты предполагают эффективную аффинность связывания таких препаратов с ферментами организмов, принадлежащих к кластеру, сформированному на основе структурного сходства.*

Реализация: *Такой подход может быть принят для использования имеющихся в настоящее время препаратов для более широкого круга патогенных микроорганизмов.*

Дополнительная информация: *http://journal-biogen.org/article/view/58/25*

Ключевые слова: *Инфекционных болезней, путь шикимовой кислоты, мишень лекарственного препарата, структурное сходство, NiRMSD.*