Research Article

Open Access

Computational Methods to Identify Drug Target in Multidrug Resistant Human Pathogens Klebsiella Pneumonia, Vibrio Cholerae and Yersinia Pestis

Iqra Ahmad*, Noor Ul ain Sajid and Sundus Iqbal

Masters in Linguistics form Kashmir ,University, Dargah. Sanat Nagar, Kashmir, India

Corresponding author: Iqra Ahmad, Masters in Linguistics form Kashmir ,University, Dargah. Sanat Nagar, Kashmir, India.

E-mail: iqra_ahmad@ymail.com

Citation: Iqra Ahmad et al. (2016), Computational Methods to Identify Drug Target in Multidrug Resistant Human Pathogens Klebsiella Pneumonia, Vibrio Cholerae and Yersinia Pestis. Int J Pharm Sci & Scient Res. 2:4, 107-119.

DOI: 10.25141/2471-6782-2016-4.0184

Copyright: ©2016 Iqra Ahmad et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Received: August 20, 2016; **Accepted:** September 19 , 2016; **Published:** September 29, 2016

Abstract

Drugable genome of gram negative, multidrug resistant Klebsiella pneumoniae JM45, Vibrio cholerae O395 and Yersinia pestis D182038, to find the therapeutic potential by using various in silico approaches. After successful screening of druggable candidates, on the basis of set parameters, a potential drug target candidate entE present in all bacterial strains was selected. EntE is crucial for the synthesis of enterobactin which belongs to biosynthesis of siderophore pathway. Siderophores have applications in medicine for antibiotics for improved drug targeting. Therefore, it is one of the most viable candidates for drug development. To this aim molecular modeling was carried out to gain insights into active site and modeling was performed via MODELLER and web servers. Best modeled structure was selected and used for molecular docking studies. Total 106 compounds library were prepared for docking and compound 103 was the best docked compound with a GOLDScore of 75.7. Moreover, time dependent dynamic behaviour of docked complexes were analysed using MD simulation studies. MD trajectories analysis revealed the flexibility of loop region to stabilize the binding of ligand and target protein and hydrogen bonding pattern was also rearranged. These conformational changes suggested the potential of compound 103 to act as lead compound.

Keywords: K. Pneumonia, V. Cholerae, Y. Pestis, Homology Modeling, Molecular Docking, Md Simulations

Introduction:

Klebsiella pneumoniae (K. pneumoniae) is the most prevalent and clinically significant human pathogen amongst rod shaped and cylindrical members of genus Klebsiella ^[1]. Classic manifestation of K. pneumoniae infections includes pneumonia and respiratory diseases, and blood stream infections in neonates, elderly and immune-compromised patients^[2]. Yersinia pestis is a gram-negative, rod-shaped coccobacillus, a facultative anaerobic bacterium that can infect humans and animals and is the causative agent of bubonic plague^[3]. V. cholerae is the causative agent of the diarrheal disease cholera. V. cholerae is a water-borne pathogen that is attained through drinking of polluted water or eating contaminated food^[4]. Assisted by a diverse assortment of pathotypes, multidrug resistant (MDR) K. pneumonia, V. cholerae

and Y. pestis associated infections exhibit a globally varying pattern of prevalence and disease associated mortalities ^[5].

The issue of increasing incidence of K. pneumonia, V. cholerae and Y. pestis infections has been complemented by the rapid dissemination of resistance against all major classes of antibiotics among different strains^[6]. More worrisome is the acquisition of virulent genes producing strains have been associated with large scale epidemics in health care facilities world-wide. Recent researches in Pakistan have revealed persistence of MDR K. pneumonia, V. cholerae and Y. infections in health care facilities ^[7]. Researches have been instigated to evaluate the limitations of existing treatment options and to formulate novel therapeutic interventions against these resilient pathogen^[8]. Large scale genome

and proteome comparative analysis can aid the identification of pathogenicity associated factors among bacterial pathogens. However, with the shift from genetic to genomic approach, the conventional route to drug target mining associated with strenuous procedure and technical expense, is no longer desirable.

The advent of resilient bacterial species and inefficacy of frontline antibiotics has necessitated the application of highly efficient computational paradigm of drug design in order to equip modern medicine with novel antibacterial drugs. Trending application of bioinformatics expertise has immensely accelerated the standard procedure of drug design^[9]. In silico drug target mining lays the groundwork in this regard and has been favoured in numerous researches concerning pathogenic bacteria^[10]. CADD is a whole genome comparative approach in which microbial genome is extensively traversed through a series of bioinformatics tools in order to mine therapeutically important, unique, druggable targets^[11]. Such unique targets can be focused upon by therapeutic drugs in an attempt to achieve a more positive outcome. Moreover, use of CADD methods such as 3D structural elucidation of drug targets, molecular docking and molecular dynamics simulations analysis contribute substantially towards the identification of potent antibacterial compounds and determination of time dependant nature of drug-target interactions^[12].

Current research is driven by the necessity to address the incessantly increasing health-risk posed by life threatening nosocomial infections. *K. pneumonia* JM45, *V. cholera* O395, *Y. pestis* D182038, are MDR strains, harbouring a diverse repertoire of drug resistance genes. The study focuses on a common pathogenic drug target in these human pathogens. Moreover, the research entails 3D structure determination of target protein and dissection of molecular interactions between inhibitory compound and identified target that can lead to potential bactericidal activity against the bacteria. The feasibility of drug target interactions and time dependant mechanics in real environment has been simulated in the final section. The design of this study incorporates an integrated in silico methodology that yields essential, pathogen specific drug targets. The selected druggable target, entE which is an enzyme and gene name is entE. 2, 3-dihydroxybenzoate-AMP ligase is an enzyme that catalyzes the chemical reaction where the two substrates of this enzyme are ATP and 2, 3-dihydroxybenzoate. This enzyme participates in siderophore biosynthesis. Siderophores have applications in medicine for antibiotics for improved drug targeting^[13]. Understanding the mechanistic pathways of these enzymes has led to opportunities for designing small-molecule inhibitors that block siderophore biosynthesis and therefore bacterial growth and virulence in iron-limiting environments. Targeting these enzymes has growth inhibitory effect on bacteria as inhibitors halt the biosynthesis of amino acids. Thus, entE is used as a potential drug target. Following target selection, a 3D structure of the target protein is determined via homology modeling; a step elemental to the structure based analytical routines of previously unexplored biomolecules. Next, molecular docking employs the 3D structure

of target protein to identify a potent drug against pathogenic target and assess the native configuration of drug target complex. The scope of current work is broadened by the application MD simulations, for monitoring time dependant dynamics of docked drug target interactions. Insights from this study will be conducive to pathogen specific drug designing protocols which in turn can be extrapolated to improve the existing healthcare systems.

Materials and Methods

Figure 1 outlines the methodology applied in current study which will be detailed in the following section.

Homology Modeling

Model Building

3D structure of selected drug target, entE, was modeled. PDB Id: 1MDB, was used as the template to model *V. cholerae* O395. The experimentally derived structure was in complex form with a bound 2, 3-dihydroxybenzohydroxamoyl adenylate inhibitor. All non-standard molecules were removed from the template structure. MODELLER v.9.14 [14] was employed for command based generation of the model structure. For a comparative analysis, web based tools i.e. I-TASSER^[15], ModWeb^[16], 3D-JIGSAW^[17], SWISSMODEL^[18] and EsyPred3D^[19] were also employed to model 3D structures of target protein. Template 1MDB was specified for model building in each case.

Model Quality Assessment

An assortment of structure validation tools: PROCHECK^[20], ERRAT^[21], Verify-3D^[22] and ProSA^[23], were used to evaluate stereochemical properties and model quality of generated models. Root mean square deviations (RMSD) of generated models were calculated individually by superimposing upon reference structure at UCSF Chimera^[24]. For further structural refinement, the model with best stereochemistry was subjected to energy optimization at Chimera. The structure was assigned Gasteiger-Huckel charges and a total of 1500 energy minimization steps were performed by applying Tripos force-field (TFF) in order to remove residual steric clashes.

Docking Protocol

Active Site Mapping

Literature references to significantly similar known homologous structures of entE, belonging to *K. pneumonia*, *V. cholerae* and *Y. pestis* were scrutinized to determine the primary ligand binding residues. These multiple sequences were aligned with *V. cholerae* O395 using clustal omega (clustalo) [25] in order to manually map the corresponding active site residues in target protein.

Ligand Selection and Preparation

Selection of the ligands was largely guided by the information gathered from BRENDA [26]. The database was explored by querying EC no. 2.7.7.58 for the enzymatic target protein in order to acquire a list of compounds with established inhibitory activity

Figure 1. A flowchart of the integrative methodology applied in the study.

against the protein. Some recently explored classes of inhibitory agents were also taken into consideration ^[27]. ChemDraw Ultra 8.0 application of the integrated suite, ChemOffice 2004^[28] was utilized for drawing scientifically accurate structures of selected ligands. 2D images were then employed to obtain 3D atomic coordinate files (.pdb) of the ligand on Chem3D Ultra 8.0. Energy minimization of ligands was carried out at Chimera using ff03r.1 force field.

Molecular Docking

GOLD^[29] program was employed to predict the structure of protein-ligand docked complexes. Default parameterization settings were opted for exploring the binding orientation of the ligands within the 10 Å radius of user defined active site residue, which was specified to be GLY 195. Genetic algorithm was used to explore optimal docking solutions, of which 10 best scoring solutions were retained. GoldScore function was utilized to calculate and score the 'fitness' of ligands into binding pocket. Additionally, molecular docking was performed using AutoDock Vina^[30] in order to obtain binding affinities of ligands for the receptor protein. Lamarckian Genetic Algorithm was employed to perform rigid receptor dockings, with grid box centred at the ligand, having dimensions -1.408, 25.029, 35.761. Docking results were visualized for bonded and non-bonded interactions using Visual Molecular Dynamics (VMD)^[31], Ligplot^[32], Discovery Studio (DS) Visualizer 3.5 software^[33].

MD Simulations

MD Simulations were performed for docked entE complex by using different modules of Amber suite^[34].

System Preparation

Basic simulation environment for docked system was modeled by Leap module of Amber 12. Ff03r.1 force field was used to define entE in the system. For docked system, ligand force field parameter file defining the topology and connectivity of ligand was created by employing GAFF (general amber force field) in Antechamber program of Amber 12. The parameter file was loaded in Xleap, the graphical interface to the Amber 12 module, Solvated environment was simulated using TIP3P implicit water model by adding 6 net neutralization sodium ions (Na⁺) to the docked system. A total of 8549 water molecules, with a distance of 10 Å between solute and solvent box margins (Figure 2).

Simulations

MD simulations for both systems were performed using Sander module of Amber 10. To remove possible steric clashes, the protein was subjected to a total of 5000 minimization steps, comprising of 2500 conjugate gradient and 2500 steepest descent gradient runs. Heating of the systems from 0 to 300 K at 1 atm was performed for 10 ps. Langevin Dynamics were applied for controlling the temperature of the system. Systems were equilibrated under similar conditions for 100 ps at constant temperature (300 K) and

pressure. Periodic boundary conditions were applied and Shake Algorithm was used to constrain the calculation of hydrogen bonds. 70 ns simulations were performed for docked systems and trajectories for the respective systems were saved for a time-step of 2 picoseconds. Command based Ptraj module of Amber10 was used to calculate physical properties of the systems and Xmgrace was employed for graphical analysis of the output files.

Results

Templates were available for all identified targets. Analytical observations revealed entE as the most suitable drug target to be pursued in current study.

Homology Modeling

2.01 Å resolution, X-ray crystallographic structure of PDB Id: 1MDB was identified as the most suitable template for modeling entE. Based on the template, 3D structural models of entE were acquired using MODELLER v.9.14 and other web-based software. Parameters for the quality assessment of the generated models are listed in Table 1. Ensuing comparison of the model quality and stereochemistry guided the selection of the best model. With best PROCHECK statistics (94.8% residues in the core favourable region), highest ERRAT quality (81.70) and RMSD of 0.237 Å (Figure 3), MODELLER model was deemed the most reliable and high quality representative 3D structure of entE. Physicochemical properties of selected model are also given in Table 2.

Docking Protocol

Primary binding site of entE was identified. Goldscores and binding affinities of the top ten docked compounds are listed in Table 3. GOLD fitness scores of the ligand poses spanned the range of 35.31 to 75.7. Binding affinities of the inhibitor compounds obtained

from AutoDock Vina ranged from -4.7 and -7.8 kcal/mol. Ranking of the docked poses brought forth plausibly potent inhibitors against entE. With highest GOLD fitness score of 75.7 and binding affinity of -7.8 kJmol⁻¹, binding orientation of compound 103 exhibited best complementary binding in the active site of entE. Molecular surface view of the best docked complex is presented in Figure 4. Molecular interaction analysis of compound 103-entE complex by VMD showed that potentially engaged the bound ligand in an elaborate network of hydrogen bonding at distances ranging from 1.5 - 3.5 Å. The identified bonds are listed in Table 4. 2D image of DS Visualizer and LIGPLOT is represented in Figure 5a and 5b respectively. It reveals an extensive coordination between ligand and entE characterized by multiple hydrogen and polar bonds. Ligand oxygen moiety formed two hydrogen bonds with residue Glu 338 having 2.94 Å and 2.75 Å distance and His 238 atom developed hydrogen bonding with ligand at the distance of 2.81 Å. Moreover, Asp526 making a hydrogen bond of length 3.02 Å. The residues Pro 201, Val 522, Gly 416, Ile 525 along with other residues were involved in hydrophobic interactions. Pi-Pi interactions were observed between the aromatic rings of Gly 195 and ligand acetyl chain.

MD Simulations

Root mean square deviation (RMSD) graph for 70 ns simulation of docked entE is represented in Figure 6a. Backbone RMSD of Cα atoms showed a steady increase in initial 10 ns and was stabilized till 50th ns. Overall, simulation run of 70 ns revealed an average root mean square deviation of 2.84 Å for docked entE structure with maximum value of 3.81 Å observed in the 16th ns. Average root mean square fluctuation (RMSF) of the docked system was 1.49 Å with maximum value of 5.82 Å observed for residue 58. Peaks indicated in RMSF graph (Figure 7a), represent residues

Figure 2. Solvation box surrounding docked protein.

with highest atomic oscillations whereby peaks correspond to loop regions. β -Factor graph in Figure 7b is in accordance with RMSF trends, whereby average value of β -Factor for docked entE was 77.1 Å² with a maximum value of 2055 Å² corresponding to residue 58.

Compactness and stability of a protein structure is determined by its radius of gyration (Rg). Reduction of radius of gyration values specified the stability of the system. The average value of 22.74 Å observed for docked protein complex, with a maximum value of 23.23 Å, denotes stability of the protein structure (Figure 8). This

implies that docked complex exhibits stable and compact system according to the value of radius of gyration.

Discussion

Pathogens responsible for different diseases are found everywhere in our environment, from hot springs to snow glaciers. The use of antibiotics is increasing day by day to encounter the attack of the pathogens. With the passage of time the pathogens are evolving and they are developing resistivity leading to multi drug resistant strains. Therefore, the treatment of infectious diseases is becoming

Structure Resource	Number of Residues				G Factor	Bad Contacts	Z Score
	[A,B,L] Allowed region	[a,b,l,p] Addi- tionally allowed region	[~a,~b,~l,~p] Generously al- lowed region	Disallowed region			
MODELLER (1)	93.5%	5.6%	0.4%	0.0%	0.4	0.2	-9.78
MODELLER (2)	93.5%	6.9%	0.8%	0.0%	0.8	0	-9.84
MODELLER (3)	92.3.4%	5.8%	0.4%	0.0%	0.6	0.1	-9.94
MODELLER (4)	93.5%	6.1%	0.4%	0.0%	0.2	0.2	-9.97
MODELLER (5)	93.5%	6.1%	0.2%	0.2%	0.5	0.3	-9.86
I-TASSER	79.0%	17.2%	1.0%	1.7%	0.4	0.8	-10.6
ModWeb	90.3%	6.3%	0.4%	0.0%	0.7	0.5	-9.01
Swiss-Model	85.4%	8.5%	1.5%	0.6%	0.2	0.6	-10.72

Table 1. Stereochemical evaluation parameters for entE structure models generated by MODELLER and web based servers

Figure 3. Superimposed structures of template IMDB (blue) and target entE (pink).

Physicochemical Properties	Values
Number of amino acids	543
Molecular weight	60123.4
Theoretical pI	6.02
Instability index	36.48
Aliphatic index	92.43
Grand average of hydropathicity (GRA-VY)	-0.144
Total number of negatively charged residues (Asp + Glu)	55
Total number of positively charged residues (Arg + Lys)	47

Table 2 Physicochemical properties of entE using ExPASy ProtParam tool.

Sr. No.	Compound No.	GoldScore	Binding Affinity (kcal/mol)
1.	103	75.7	-7.5
2.	76	70.4	-7.4
3.	106	70.2	-6.8
4.	81	69.5	-7.8
5.	96	69.4	-7.6
6.	79	58.0	-6.9
7.	105	57.7	-7.2
8.	63	56.8	-7.1
9.	20	55.9	-6.8
10.	15	55.3	-7.4

Table 3. Docking results of inhibitors arranged in descending order of GOLDScore with corresponding binding affinities

difficult. Various biotypes and strains of bacteria's exist. The horizontal gene transfer of the virulence genes and pathogenicity islands is the prominent cause of the emergence of new strains with increased drug resistivity (Kovach, Shaffer and Peterson, 1996). Thus, more promising drug designing techniques and approaches are been developed.

In current work in silico screening of *K. pneumonia* JM45, *V. cholerae* 0395, *Y. pestis* D182038 bacterial genome for novel targets possessing principle qualifying features of an effective therapeutic agent. Exclusive to gram negative bacteria, target selected in current study was 2, 3-dihydroxybenzoate-AMP ligase (entE), involved in "Biosynthesis of siderophore" pathway. This gene was common to all the three bacteria. Selected entE gene was used for computational homology modeling to identify the possible binding modes of the ligands.

Comparative homology modeling was performed to model entE using MODELLER9.14 and web based servers. The generated model was evaluated with various tools including ERRAT, PDBSum and Verify3D, which tested the quality of structures generated. Template selected for modeling having the PDB ID: 1MDB had 100% coverage and 59% sequence identity. Model 2 was selected as it had minimal bad contacts 0 and highest G factor 0.8. MODELLER model was deemed the most reliable and high quality representative 3D structure of entE, for conducting the docking and simulation studies. Among the web servers ModWeb generated the best model. Selected model 2 generated via MODELLER9.14 was monomer and its subcellular location was cytoplasm. RMSD is calculated to check the similarity between two proteins, which analyze the ensembles of backbone atoms for proper conformation [35]. RMSD value calculated for target and

Protein Interacting Atom	Ligand Interacting Atom	Distance (Å)
194 SER O	UNK N	2.98
194 SER HG	UNK N	3.89
194 SER O	UNK H	3.45
195 GLY O	UNK N	2.21
195 GLY HA2	UNK O	3.09
195 GLY HA3	UNK N	3.88
196 GLY HA2	UNK N	3.49
197 THR N	UNK H	2.57
238 HIS ND1	UNK N	2.81
413 TYR HH	UNK O	3.20
432 ARG HD3	UNK N	2.70
432 ARG HE	UNK N	2.88
432 ARG NE	UNK O	2.77
432 ARG HD3	UNK O	2.53
432 ARG NH1	UNK H	3.82
434 LYS HE2	UNK O	2.47
524 LYS NZ	UNK H	3.42
526 ASP OD2	UNK H	3.62

Table 4. H- bond interactions between entE and compound 103 identified by using VMD

Figure 6.

(a) RMSD plot of docked entE protein complex for 70 ns simulation run.

(b) Snapshots taken of docked protein entE at 0 ns, 25 ns, 50 ns and 70 ns timescale. The red circle highlights the changes that were observed over the simulation.

Figure 7.

(a) RMSF plot of docked entE protein over 70 ns simulation run.

(b) β -Factor graph of docked entE protein over 70 ns simulation run

Figure 8. Radius of gyration of docked protein entE over 70 ns simulation time period.

template was 0.237 Å, low value means generated model was of good quality and had similar main chain fold as that of template chain. Detailed physicochemical properties of the modeled structure are mentioned Table 2.

Molecular docking is an *in silico* method for exploration of binding modes of two molecules along with their binding affinities [36]. To dock molecule some site is crucial thus, in current study Gly195 used as an active residue. Total 106 ligands were docked in active site present in entE. Docking of ligand was performed via AutoDock Vina and GOLD. Compound 103 i.e. 2,3-dihydroxybenzohydroxamoyl adenylate, was the best docking ligand and showed the highest GOLDScore 75.7 with binding affinity of -7.5 kcal/mol. Interaction studies of compound 103 with protein was visualized by LIGPLOT, DS visualizer and UCSF Chimera. Ligand oxygen moiety formed two hydrogen bonds with residue Glu 338 having 2.94 Å and 2.75 Å distance and His 238 atom developed hydrogen bonding with ligand at the distance of 2.81 Å, while DS Visualizer highlighted the hydrogen bonds with Asp526, Glu338 and His238. The residues Pro 201, Val 522, Gly 416, and Ile 525 along with other residues were involved in hydrophobic interactions. Moreover, π -interactions and sigma interactions were also present between ligand and target protein. The docking study of entE, further needed an understanding of the structural adjustments made upon ligand introduction into the

system. However, it provided this information within the context of a static environment. In order to insinuate the dynamic conduct, simulation protocol was carried out that provided eloquent insights into the structural basis of druggability potential of entE. This was followed by trajectory analysis to assess various characteristics of the docked system. For normal physiological functioning of biomolecules aqueous medium is required. In the similar way, inhibition also required solvated medium. Molecular dynamics simulation allows to unveil the time dependent dynamic behaviour of biomolecules. It provides the insights about the region of molecule which is involved in dynamic behaviour of system [37]. In current study solvated medium was provided to the docked complex of entE to understand the behaviour of inhibition process, with respect to time.

Properties namely RMSD, RMSF, B-factor and radius of gyration were plotted as a function of time to implicate the biomolecular arrangements within a solvated environment. Analysis of protein, whilst in ligand bound form helped unravel the ligand induced variability and led to evaluation of organizational alterations of the protein-ligand system [38]. The deviation of the backbone C α atoms was noticed for a time period of 70 ns. The inhibitor bound entE RMSD behavior over the studied time scale, shows slight shifts in the structure and geometry of the protein-ligand complex. It also denotes that the ligand placement is well supplemented

within the active site and does not disrupt the protein chemistry.

The average RMSD value of 2.84 Å was observed. Overall, the pattern of RMSD graph supports slight shifts within the structural framework of the protein-ligand complex. As a measure of atomic fluctuations, RMSF helped distinguish between the structurally flexible and rigid loci of the drug target. The average Ca fluctuation is 1.49 Å. Detailed analysis of the trajectories led to identification of those protein substructures that are responsible for the obtained RMSF trend. In accordance with the RMSD graph, conspicuous structural alterations occurred at a time of 10ns, 25ns and 50ns, after which the protein acquired a stable conformation. Figure 6b represents the structural and conformational alterations that occurred in the docked protein complex, at different time lapses. The graph indicates an alternate increasing and decreasing trend at the 20thns. This is due to the structural shift of an alpha helix into a turn at one location. Obvious structural changes occurred from 25thns to the 50thns. The protein structure however, maintains the alterations in its conformity, from the 50thns onwards, till the 70thns. Thus, simulating the docked protein, aided the protein-ligand complex achieve its stable conformation.

The variations underscored in the RMSD plot, were additionally analyzed for the fluctuations at the atomic levels, revealing the individual residues, being the cause of the alterations in the simulated protein-ligand complex. The RMSF graph designates the residue locations, instigating the changes. The highest peak nominates 11 residues in total that form a loop, notably Val62, Thr63, Glu64, Asn65, Thr66, Thr67, Glu68, Lys69, Ala70, Ala71 and Thr72. Changes in this loop are the major source of the peaks in the RMSD graph. The other loci being altered are Leu176, Val167, Asn168, Ala169 and Lys239, Asp240. Amino acid residues that incur the fluctuations are away from the active site pocket of the protein. Thus no major changes occur at the active site. Changes in the RMSD and RMSF plots, thus illustrate the fact that the conformational deviations and alterations are in a way worthy for the protein-ligand complex. The unwavering active site residues ascertain the active site locus to remain stable throughout the simulation.

β-Factor explains the flexibility and thermal stability of a protein in terms of time. The pattern of beta factor for the docked protein complex is consistent with the RMSF trend. Regions having greater fluctuations as identified in the RMSF analysis exhibit average beta factor of 77.1 Å² for ligand bound protein, presenting conformational homogeneity. As an assessment constraint of the structural compactness, radius of gyration was also calculated as a time function for the 70 ns simulation of ligand-protein complex. The average value of 22.74 Å observed for docked protein complex, denotes stability of the protein structure. Stability of docked complex was explained by physical properties calculation over the time period of 70 ns for docked protein. These findings were explaining dynamics of protein.

In silico approach adopted in current study highlighted the significant results at various stages of analysis. The structure dynamics of docked protein and its simulation analysis provided

significant insights, which can be practiced to increase the efficacy of drug against the three bacteria along with drug target specificity and selectivity and to cure infections caused by pathogens *K. pneumonia*, *V. cholerae* and *Y. pestis*.

Conclusion

Current work culminates into a successful realization of outlined objectives by systematic in silico application of drug design modules. The strategic route applied for the scrutinizing therapeutic candidates in the emerging and evolving gram negative, multi drug resistant bacterial pathogens, *K. pneumonia*, *V. cholerae* and *Y. pestis* has provided findings that are of great pharmacological prominence. Selection of entE, that is common and unique as it has a decisive role in formation of enterobactin, a siderophore necessary for the survival of this bacteria in iron limiting conditions. Homologues of enterobactin are present in *K. pneumonia*, *V. cholerae* and *Y. pestis*. Thus, targeting this protein will halt the bacterial growth and thus will cause the bacteria to die. As a result diseases caused by these bacteria namely cholera by *V. cholerae*, pneumonia by *K. pneumonia* and plague by *Y. pestis* will be reduced greatly. Efficacious application of the comparative homology modeling generated a high quality model for previously structurally uncharacterized entE. Rigorous stereochemical verification endorsed the homology model as fit for usage in high level analytical work such as molecular docking and molecular dynamics. Subsequently, it functioned as underpinning for docking studies which revealed 2, 3-dihydroxybenzohydroxamoyl adenylate as strong inhibitor against entE, exhibiting wide hydrogen bonding. The influence of these interactions was imitated by a Gold-Score of 75.7. The strength of polar inhibitor moieties on the level of protein ligand interactions was concluded from these observations, which coincided with the physicochemical character of the binding pocket denoting a preference for charged residues. Molecular dynamic studies including the MD simulations well explained the dynamic behavior of the docked protein. Beside the side chain fluctuations and minor helix to loop movement, stability of inhibitor and target protein complex was observed. Consequently, coupled with information about key functional residues of drug target, the atomic level structural and dynamic insights can fuel a structure based drug design of novel inhibitors with increased drug selectivity, efficacy and specificity against *K. pneumonia*, *V. cholerae* and *Y. pestis*.

Author Contributions

SSA conceived the idea. He has revised the draft critically for important intellectual content and has given final approval of the version to be published. ASA carried out computational work and has been involved in drafting the manuscript.

Funding Sources

Financial assistance is provided by Higher Education Commission (HEC).

ACKNOWLEDGMENT

Authors are grateful to the supportive group at computational biology lab and employees at National Center for Bioinformatics, Quaid-i-Azam University.

ABBREVIATIONS

MDR, Multidrug resistant; CADD, Computer Aided Drug Design; K. pneumonia, *Klebsiella pneumonia*; V. cholerae, *Vibrio cholerae*; Y. pestis, *Yersinia pestis*; 3D, 3-Dimensional; AMBER, Assisted Model Building with Energy Refinement; UniprotKB, Uniprot Knowledgebase; I-TASSER, Iterative TASSER; DS Visualizer, Discovery Studio Visualizer; GOLD, Genetic Optimization for Ligand Docking; EC, Enzyme Commission; ProSA-web, Protein Structure Analysis-web; RMSD, Root Mean Square Deviation; RMSF, Root Mean Square Fluctuation; β -Factor, Beta Factor; MD, Molecular Dynamics; ns, nanosecond; ps, picosecond; GAFF, General AMBER force field; VMD, Visual Molecular Dynamics.

References

1. Richard, F. In *Taxonomic Guide to Infectious Diseases: Understanding the Biologic Classes of Pathogenic Organisms*; Berman, J. J., Ed.; Academic Press, 2012; Chapter 7, pp 35-47.
2. Moreira, B. M.; Gales, A. C.; Alves, M. S.; da Silva Dias, R. C. In *Molecular Detection of Human Bacterial Pathogens*; Liu, D., Ed.; CRC Press, 2011; Chapter 78, pp 905-919.
3. Bhunia, A.K., 2008. *Yersinia enterocolitica* and *Yersinia pestis*. In *Foodborne microbial pathogens mechanisms and pathogenesis*. New York: Springer, pp. 227–276.
4. Faruque, S.M., Albert, M.J. and Mekalanos, J.J., 1998. *Epidemiology, Genetics, and Ecology of Toxigenic Vibrio cholerae*. *Microbiology and molecular biology reviews*, 62(4), pp.1301-1314.
5. Ko, W. C.; Paterson, D. L.; Sagnimeni, A. J.; Hansen, D. S.; Von Gottberg, A.; et al. *Community-acquired Klebsiella pneumoniae bacteremia: global differences in clinical patterns*. *Emerg. Infect. Dis*. 2002, 8, 160-166.
6. Mantzarlis, K.; Makris, D.; Manoulakas, E.; Karvouniaris, M.; Zakyntinos, E. *Risk factors for the first episode of Klebsiella pneumoniae resistant to carbapenems infection in critically ill patients: A prospective study*. *BioMed Res. Int*. 2013. 2013.
7. Saleem, A. F.; Qamar, F. N.; Shahzad, H.; Qadir, M.; Zaidi, A. K. M. *Trends in antibiotic susceptibility and incidence of late-onset Klebsiella pneumoniae neonatal sepsis over a six-year period in a neonatal intensive care unit in Karachi, Pakistan*. *Int. J. Infect. Dis*. 2013, 17, e961-e965.
8. Hirsch, E. B.; Tam, V. H. *Detection and treatment options for Klebsiella pneumonia carbapenemases (KPCs): an emerging cause of multidrug-resistant infection*. *J. Antimicrob. Chemother*. 2010, 65, 1119-1125.

9. Phoebe Chen, Y. P.; Chen, F. *Identifying targets for drug discovery using bioinformatics*. *Expert Opin. Ther. Target*. 2008, 12, 383-389.
10. Amineni, U.; Pradhan, D.; Marisetty, H. *In silico identification of common putative drug targets in Leptospira interrogans*. *J. Chem. Biol*. 2010, 3, 165-173.
11. Barh, D.; Tiwari, S.; Jain, N.; Ali, A.; Santos, A. R.; Misra, A. N.; Azevedo, V.; Kumar, A. *In silico subtractive genomics for target identification in human bacterial pathogens*. *Drug Dev. Res*. 2011, 72, 162-177.
12. Azam, S. S.; Abbasi, S. W.; Akhtar, A. S.; Mirza, M. L. *Comparative modeling and molecular docking studies of d-Alanine: d-alanine ligase: a target of antibacterial drugs*. *Med. Chem. Res*. 2014, DOI: 10.1007/s00044-014-0970-8.
13. Rusnak, F., Faraci, W.S. and Walsh, C.T., 1989. *Subcloning, expression, and purification of the entErobactin biosynthetic enzyme 2, 3-dihydroxybenzoate-AMP ligase: demonstration of en-zyme-bound (2, 3-dihydroxybenzoyl) adenylate product*. *Biochemistry*, 28(17), pp.6827-6835
14. Eswar, N.; Webb, B.; Marti-Renom, M. A.; Madhusudhan, M. S.; Eramian, D.; Shen, M.; Pieper, U.; Sali, A. *Comparative protein structure modeling using Modeller*. In *Current Protocols in Bioinformatics*; Baxevanis, A. D.; Petsko, G. A.; Stein, L. D.; Stormo, G. D., Eds.; John Wiley & Sons, 2006; Chapter 5, Unit 5.6, pp. 5.6.1-5.6.30.
15. Zhang, Y. *I-TASSER server for protein 3D structure prediction*. *BMC Bioinformatics*. 2008, 9, 40.
16. Pieper, U.; Webb, B.M.; Dong, G.Q.; Schneidman-Duhvny, D.; Fan, H.; Kim, S. J.; Khuri, N.; Spill, Y. G.; Weinkam, P.; Hammel, M.; Tainer, J. A.; Nilges, M.; Sali, A.. *ModBase, a database of annotated comparative protein structure models and associated resources*. *Nucleic Acids Res*. 2014, 42, D336-D346.
17. Bates, P. A.; Kelley, L.A.; MacCallum, R.M.; Sternberg, M. J. *Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM*. *Proteins*. 2001, 45, 39-46.
18. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling*. *Bioinformatics*. 2006, 22, 195-201.
19. Lambert, C.; Léonard, N.; De Bolle, X.; Depiereux, E. *ESyPred3D: Prediction of proteins 3D structures*. *Bioinformatics*. 2002, 18, 1250-1256.
20. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J.M. *PROCHECK: a program to check the stereochemical quality of protein structures*. *J. Appl. Cryst*. 1993, 26, 283-291.
21. Colovos, C.; Yeates, T. O. *Verification of protein structures: patterns of non-bonded atomic interactions*. *Protein Sci*. 1993, 2, 1511-1519.
22. Eisenberg, D.; Lüthy, R.; Bowie, J. U. *VERIFY3D: assessment*

of protein models with three-dimensional profiles. *Methods Enzymol.* 1997, 277, 396-404.

23. Wiederstein, M.; Sippl, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 2007, 35, W407-W410.

24. Pettersen, E.F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera- a visualization system for exploratory research and analysis. *J. Comput. Chem.* 2004, 25, 1605-1612.

25. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011, 7, 539.

26. Schomburg, I.; Chang, A.; Placzek, S.; Söhngen, C.; Rother, M.; Lang, M.; Munaretto, C.; Ulas, S.; Stelzer, M.; Grote, A.; Scheer, M.; Schomburg, D. BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. *Nucleic Acids Res.* 2013, 41, D764-D772.

27. Brown, M. F.; Reilly, U.; Abramite, J. A.; Arcari, J. T.; Oliver, R.; Barham, R. A.; Che, Y.; Chen, J. M.; Collantes, E. M.; Chung, S. W.; Desbonnet, C.; Doty, J.; Doroski, M.; Engtrakul, J. J.; Harris, T. M.; Huband, M.; Knafels, J. D.; Leach, K. L.; Liu, S.; Marfat, A.; Marra, A.; McElroy, E.; Melnick, M.; Menard, C. A.; Montgomery, J. I.; Mullins, L.; Noe, M. C.; O'Donnell, J.; Penzien, J.; Plummer, M. S.; Price, L. M.; Shanmugasundaram, V.; Thoma, C.; Uccello, D. P.; Warmus, J. S.; Wishka, D. G. Potent inhibitors of LpxC for the treatment of gram-negative infections. *J. Med. Chem.* 2012, 55, 914-923.

28. Li, Z.; Wan, H.; Shi, Y.; Ouyang, P. Personal experience with four kinds of chemical structure drawing software: review on ChemDraw, ChemWindow, ISIS/Draw, and ChemSketch. *J. Chem. Inf. Comput. Sci.* 2004, 44, 1886-1890.

29. Jones, G.; Willett, P.; Glen, R. C. Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* 1995, 245, 43-53.

30. Trott, O.; Olson, A. J. Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 2009, 31, 455-461.

31. Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics.* 1996, 14, 33-38.

32. Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* 1995, 8, 127-134.

33. Accelrys Software Inc. Discovery Studio Visualizer, Release 3.5. Accelrys Software Inc. 2012.

34. Weiner, P. K.; Kollman, P. A. AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comput. Chem.* 1981, 2, 287-303.

35. Maiorov, V.N.; Crippen, G.M. Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins. *Journal of Molecular Biology.* 1994, 235(2), 625-634.

36. Huang, S.Y.; Zou, X. Advances and challenges in protein-ligand docking. *Intl. J. of Mol. Sci.* 2010, 11(8), 3016-3034.

37. Azam, S.S.; Uddin, R.; Wadood, A. Structure and dynamics of alpha-glucosidase through molecular dynamics simulation studies. *J. of Mol. Liq.* 2012, 174, 58-62.

38. Azam, S.S.; Saroosh, A.; Zaman, N.; Raza, S. Role of N-acetylserotonin O-methyltransferase in bipolar disorders and its dynamics. *J. of Mol. Liq.* 2013, 182, 25-31.