Journal of Molecular Graphics and Modelling 45 (2013) 1–12

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Molecular modeling and simulation of FabG, an enzyme involved in the fatty acid pathway of *Streptococcus pyogenes*

Rajamohmed Beema Shafreen, Shunmugiah Karutha Pandian*

Department of Biotechnology, Alagappa University, Karaikudi, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 13 March 2013
Received in revised form 28 July 2013
Accepted 30 July 2013
Available online 13 August 2013

Keywords:
FabG (β-Ketoacyl Carrier Protein Reductase)
Streptococcus pyogenes
Homology modeling
ADMET
Cerulenin
Epigallocatechin

ABSTRACT

Streptococcus pyogenes (SP) is the major cause of pharyngitis accompanied by strep throat infections in humans. 3-keto acyl reductase (FabG), an important enzyme involved in the elongation cycle of the fatty acid pathway of *S. pyogenes*, is essential for synthesis of the cell-membrane, virulence factors and quorum sensing-related mechanisms. Targeting SPFabG may provide an important aid for the development of drugs against *S. pyogenes*. However, the absence of a crystal structure for FabG of *S. pyogenes* limits the development of structure-based drug designs. Hence, in the present study, a homology model of FabG was generated using the X-ray crystallographic structure of *Aquifex aeolicus* (PDB ID: 2PNF). The modeled structure was refined using energy minimization. Furthermore, active sites were predicted, and a large dataset of compounds was screened against SPFabG. The ligands were docked using the LigandFit module that is available from Discovery Studio version 2.5. From this list, 13 best hit ligands were chosen based on the docking score and binding energy. All of the 13 ligands were screened for Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET) properties. From this, the two best descriptors, along with one descriptor that lay outside the ADMET plot, were selected for molecular dynamic (MD) simulation. *In vitro* testing of the ligands using biological assays further substantiated the efficacy of the ligands that were screened based on the *in silico* methods.

1. Introduction

Streptococcus pyogenes is a gram-positive, non-motile, spherical bacterium. It is responsible for various human diseases, ranging from trivial to lethal [1]. Although penicillin is the drug of choice for the treatment of *S. pyogenes* infections, macrolides are used for patients who are allergic to β-lactams. The emergence of macrolide resistance in *S. pyogenes* is an increasing problem worldwide [2,3]. Furthermore, the indiscriminate usage of synthetic antimicrobials to treat infections caused by macrolide-resistant *S. pyogenes* has resulted in a strain that has multiple resistances to a major class of antibiotics [4]. In addition, the high cost and adverse side effects (*e.g.*, hypersensitivity, allergic reactions and immunosuppression) of the synthetic antimicrobials used for treating infectious diseases are major issues [5]. Today, there is an urgent and continuous need to identify compounds that are biologically active and incur minimal side effects.

Antibiotics of plant origin have vast therapeutic potential. They have proven effective when used for treatment against infectious diseases, simultaneously extenuating many of the side effects associated with synthetic antimicrobials [6]. The positive effect of potential antimicrobials from natural products with few or no side effects might depend on the structure of the compound that interacts with the toxin or pathogen and not with other molecules in the physiology of the host. This approach has become the rationale for drug design studies as a new field of research.

An extensive literature survey suggests that consuming tea (*Camellia sinensis* L.) enhances the immune system capacity to fight infectious diseases [7–9]. Important micronutrients, vitamins and minerals that are present in green tea produce its free radical capturing (antioxidant), invigorating (caffeine) and detoxifying antibacterial properties. Although the detailed mechanism of the antimicrobial activity of catechins present in green tea remains unexplored, the common target of catechins is the cell membrane of the pathogen (with broad-spectrum activity), in addition to specific targets for each pathogen [10].

Type II FAS pathway was previously reported as a significant antimicrobial target [11,12]. Findings of Brinster and Balemans [13] have opened up a demanded debate, whether FAS II pathway is suitable drug target. Furthermore, targeting FAS II pathway varies, even though the bacteria are closely related to each other. Researchers have well demonstrated that FAS II enzymes may not be suitable target during intraperitoneal infection caused by *Streptococcus agalactiae*, but these enzymes might be suitable target for eradication of cutaneous antibiotic-resistant bacteria involved in

* Corresponding author at: Department of Biotechnology, Alagappa University, Karaikudi 630 003, Tamil Nadu, India. Tel.: +91 4565 225215; fax: +91 4565 229334.
E-mail address: sk_pandian@rediffmail.com (S.K. Pandian).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.07.009

cutaneous infection [14]. Additionally, bacterial FAS II pathway is a desirable for drug discovery because the pathway is not targeted by existing drugs and is therefore likely to produce compounds that inhibit the bacteria which show resistance to the known antibiotics [15]. Recently, FabK of *Streptococcus pneumoniae* is described as an attractive and potential target for developing selective antibacterial agents [16,17]. Based on these highlights, identification of antimicrobial agents against FAS II pathway of *S. pyogenes* was chosen as desirable target for the present study. The *fabG* gene product, β-ketoacyl-acyl carrier protein (ACP) reductase (FabG), plays a key role in the synthesis of fatty acids [18]. FabG is highly conserved and ubiquitously expressed in all bacteria and is the only known isozyme that catalyzes the essential ketoreduction steps in the elongation cycle of FAS II [19–21]. Therefore, FabG represents a valid target that is yet to be explored for the development of broad-spectrum antimicrobial agents. The elongation-condensing enzymes play an important role in the regulation of fatty acid biosynthesis; because natural products are known to target this step in the pathway, FabG is clearly a desirable target for therapeutics development [22].

Natural compounds such as epigallocatechin gallate (EGCG; *e.g.*, green tea catechin) and C-3 gallic acid esters of the catechins have been reported to be potent inhibitors of FabG in *Escherichia coli* and *Plasmodium falciparum* [21,23]. To date, inhibitors against FabG of *S. pyogenes* have not been explored. As a prelude to such exploration, a three-dimensional (3D) homology model of FabG from *S. pyogenes* was developed, and a large dataset of compounds was screened against the model. The compounds screened (based on the structure-based virtual screening) were experimentally tested using biological assays.

2. Materials and methods

2.1. Homology modeling

The FabG protein sequence of *S. pyogenes* (ACZ64265) was retrieved from the NCBI database, and a basic local alignment search tool (BLAST) [24] analysis for the enzyme (FabG) was performed against the protein data bank (PDB) [25] using the default parameters to find a suitable template for homology modeling. Based on the maximum identity with a high score and low *e*-value, PDB ID: 2PNF (from *Aquifex aeolicus*) at a resolution of 1.8 Å was selected as the most appropriate template for homology modeling [26]. The sequence alignment between the target and the template was calculated using the DS Modeler, module from Discovery Studio 2.5 program, distributed by Accelrys Software Inc., San Diego, CA, USA [27]. Further energy minimization was performed using DS CHARMm to remove the geometric restraints of the SPFabG model. The secondary structure of the model was analyzed using PDBsum (available from http://www.ebi.ac.uk/pdbsum). On the SAVES server, Ramachandran's map was drawn using PROCHECK v.3.0, and non-bonded interactions between different atom types were calculated using ERRAT graph. The refined structure was chosen for further study.

2.2. Binding pocket identification

The "Define and edit binding site" protocol in Discovery Studio was used to identify the potential binding site in the protein [28]. A ligand-based search strategy was also applied, in which the known ligand (NADPH) binding active site is defined as the binding pocket for the other test ligands. Based on the earlier report of Miller et al. [29], a detailed comparison of structures from *Pseudomonas aeruginosa* (RhlG) and *E. coli* (FabG) indicated three active site residues (Ser, Tyr, and Lys) with identical roles in catalysis. The multiple sequence alignment of the FabG sequences (*P. aeruginosa*, *E. coli*, *Staphylococcus aureus*, *A. aeolicus* and *S. pyogenes*) showed highly conserved active site residues in FabG of *S. pyogenes*. This identification has assisted in locating the exact binding site in the homology model of *S. pyogenes.*

2.3. Ligand database generation and molecular docking

ACD/ChemSketch an integrated software package from Advanced Chemistry Development Inc. was used for drawing the chemical structure of CID 65064 (Epigallocatechin gallate (EGCG)). The 3D optimization algorithm in ChemSketch was used to rapidly translate the 2D planar structure into a sensible 3D structure. The 3D optimization algorithm based on CHARMM force field was applied to optimize the ligand. The 3D structure of EGCG was used as a search key against PubChem, a database of chemical and small molecule structures. The search key retrieved a series of chemical substructure with the similarity percentage of 80% and the Tanimoto score of >0.9 representing the increased possibility of these selected substructure with same bioactivity. The search resulted in 60,485 similar structures were used as library for virtual screening (VS).

Prior to VS, the homology model (SPFabG) and the library of compounds was minimized to their low energy state using the CHARMm force field implemented in DS 2.5. The convergence gradient was set to 0.01 kcal mol^{-1}, and 10,000 steps of a steepest descent algorithm were performed following 50,000 steps of conjugate gradient algorithm. A spherical cut-off of 14 Å was used for non-bonding interactions, and the other parameters were set to their default values. To validate the docking protocol setup prior to screening, the known inhibitor CID: 65064 was docked in the ligand-binding site, and the results were compared with earlier output. Then, the database was subjected to virtual screening against the homology model. The LigandFit module in DS 2.5 was used for docking the compound database. The LigandFit docking procedure follows two major elements: (i) defining the binding sites of the receptor specific for docking and (ii) docking the ligands to the specified site [30]. The top 10 docked poses were allowed to be saved. The successful poses were evaluated using a set of scoring functions, such as LigScore1, LigScore2, PLP1 and PLP2 were implemented in the DS 2.5 program. The ligands in the binding site were prioritized according to the Dock-Score function. The binding free energies ($\Delta G_{binding}$) of protein–ligand complexes were calculated [31] using the following equation:

$$\Delta G_{binding} = \Delta G_{complex} - \Delta G_{drug} - \Delta G_{target} \quad (1)$$

The top 10 poses obtained after docking were used for calculating the binding energies.

2.4. ADMET prediction

The ADMET descriptor describes the kinetics of drug exposure to tissue and pharmacological activity of the compounds. The ADMET program available in DS 2.5 requires 2D or 3D molecular structure information in either SDF or MOL file formats. The program parses the structure and calculates the values of molecular descriptors. The four ADMET properties were tested for the ligands; the human intestinal absorption descriptor illustrates the absorption of the orally administered drugs in the intestine. The compounds classified based on the predicted values such as "0", "1" and "2" are defined with good, moderate and low absorption respectively. The blood–brain barrier defines the penetration of molecules after oral administration. The cutoff value of "0", "1", "2" and "3" defines very high, high, moderate and low penetration of the compounds. The cytochrome P450 2D6 inhibition predicts the inhibition of the enzyme using 2D chemical structure as an input. The probability