Bioorganic Chemistry 59 (2015) 91–96

Contents lists available at ScienceDirect

Bioorganic Chemistry

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

3D-QSAR and molecular modeling studies on 2,3-dideoxy hexenopyranosid-4-uloses as anti-tubercular agents targeting alpha-mannosidase

Priyanka Shah [a,1], Mohammad Saquib [b,1,2], Smriti Sharma [b], Irfan Husain [b], Sandeep K. Sharma [c], Vinayak Singh [c], Ranjana Srivastava [c], Arun K. Shaw [b,*], Mohammad Imran Siddiqi [a,*]

[a] Molecular and Structural Biology Division, CSIR-Central Drug Research Institute (CDRI), Lucknow 226031, India
[b] Medicinal and Process Chemistry Division, CSIR-Central Drug Research Institute (CDRI), Lucknow 226031, India
[c] Microbiology Division, CSIR-Central Drug Research Institute (CDRI), Lucknow 226031, India

ARTICLE INFO

Article history:
Received 1 October 2014
Available online 11 February 2015

Keywords:
Homology modeling
Docking
Molecular dynamics
CoMFA
CoMSIA
Hierarchical clustering

ABSTRACT

Ligand-based and structure-based methods were applied in combination to exploit the physicochemical properties of 2,3-dideoxy hex-2-enopyranosid-4-uloses against *Mycobacterium tuberculosis* H37Rv. Statistically valid 3D-QSAR models with good correlation and predictive power were obtained with CoMFA steric and electrostatic fields ($r^2 = 0.797$, $q^2 = 0.589$) and CoMSIA with combined steric, electrostatic, hydrophobic and hydrogen bond acceptor fields ($r^2 = 0.867$, $q^2 = 0.570$) based on training set of 33 molecules with predictive r^2 of 0.808 and 0.890 for CoMFA and CoMSIA respectively. The results illustrate the requirement of optimal alkyl chain length at C-1 position and acceptor groups along hydroxy methyl substituent of C-6 to enhance the anti-tubercular activity of the 2,3-dideoxy hex-2-enopyranosid-4-uloses while any substitution at C-3 position exert diminishing effect on anti-tubercular activity of these enulosides. Further, homology modeling of *M. tuberculosis* alpha-mannosidase followed by molecular docking and molecular dynamics simulations on co-complexed models were performed to gain insight into the rationale for binding affinity of selected inhibitors with the target of interest. The comprehensive information obtained from this study will help to better understand the structural basis of biological activity of this class of molecules and guide further design of more potent analogues as anti-tubercular agents.

1. Introduction

Today tuberculosis (TB) continues to be one of the leading causes of death globally, from a single infectious agent [1–3]. The recent emergence of multi-drug resistant TB (MDR-TB), extensively drug resistant TB (XDR-TB) and co-infection with HIV have rendered the existing chemotherapeutic approaches progressively ineffective [4] raising fears that TB may become uncontrollable [5]. This has prompted a new sense of urgency toward the search of novel chemical entities as anti-TB molecules having potent activity, reduced toxicity, rapid mycobactericidal action, shortened duration of therapy and new site of action to minimize the chances of drug resistance [6,7]. In recent decades there has been a renewed interest in the use of carbohydrates scaffolds in drug discovery due to their unique properties [8]. However, a major impediment in the development of carbohydrate molecules as therapeutics has been their rapid degradation in the body by glycosidases. One way to overcome this pharmacokinetic difficulty is to use modified sugars which may not be recognized by the regular complement of glycosidases present in the body [9,10]. In this milieu, we recently reported the design and synthesis of 2,3-dideoxy hex-2-enopyranosid-4-uloses which are small, modified sugar molecules as promising anti-tubercular agents [11].

1.1. Ligand-based molecular modeling

Quantitative structure activity relationship (QSAR) can offer some valuable suggestions to improve the drug activity on the basis of QSAR analyses fitted to the activity data of training set of molecules aligned in three dimensional spaces [12–15]. Therefore, ligand based 3D-QSAR models were built using Comparative Molecular Field Analysis (CoMFA) [16] and Comparative Molecular Similarity Indices Analysis (CoMSIA) [17] to understand the key

* Corresponding authors.
E-mail address: mi_siddiqi@cdri.res.in (M.I. Siddiqi).
[1] Joint first authors.
[2] Present affiliation: Department of Chemistry, University of Allahabad, Allahabad - 211002, U.P., India.

http://dx.doi.org/10.1016/j.bioorg.2015.02.001
0045-2068/

structural elements capable to produce more effective inhibitors. Moreover, training set used to develop the predictive QSAR model is supposed to encompass diverse chemical entities as much as possible to train highly predictive models [18]. According to Maggiora and Johnson similarity–property principle [19], chemical similarity among molecules can be related to the biological activity. This correlation is mostly dependent on the method of similarity measures used in the study and varies between different methods which are classified according to the information used to calculate the similarity between two structures [20,21]. Structural or topological fingerprints, like [22] are simple and rapid methods to encode the information in the form of fingerprints based on features that are included in the structures. These fingerprints can be used to measure the degree of dissimilarity between all pairs of molecules to build a diversified training set. Therefore, to rationalize the selection method in the current study, training set of 33 molecules and test set of 5 molecules were selected from clusters built from hierarchical clustering using pvclust [23] R package to ensure the molecular diversity in both the training and test sets. 2D descriptors [22] were utilized in calculating dissimilarity matrix to cluster the whole dataset.

1.2. Structure-based molecular modeling

On the other hand, structure-based drug design depends entirely on information obtained from three dimensional structures of the biological targets of interest. The proper analyses of spatial arrangements of ligands present in the protein's active site can offer valuable rationale for activity profile of structural analogs in terms of favorable and unfavorable intermolecular interactions. We obtained the inhibition activity of three inhibitors from the present dataset against the alpha-mannosidase enzyme where one of the compounds corresponds to the most active anti-tubercular among other compounds in this dataset.[24] In the absence of X-ray crystallographic structure of the protein, homology modeling was attempted to generate three-dimensional coordinates of the target protein *Mycobacterium tuberculosis* alpha-mannosidase. Cloning and expression studies of Rv0648 gene have demonstrated in previous studies that activity of alpha-mannosidase is important for biosynthesis of mannosylated glycoconjugates in *M. tuberculosis* [25]. Sequence annotation shows that alpha mannosidase enzymes belong to the family of 38 glycosyl hydrolase and catalyze the cleavage of alpha-mannose. Due to the limited sequence similarity with full protein, homology model has been generated only for alpha-mannosidase domain. This model is further used for docking of selected inhibitors followed by nanosecond duration molecular dynamics simulations. Biological activities of three selected inhibitors against *M. tuberculosis* alpha mannosidase were considered in the present study to understand the molecular basis of rationalization behind binding affinity.[24] On the basis of ligand based QSAR analyses in conjunction with molecular docking and molecular dynamics (MD) simulations on homology model of alpha mannosidase co-complexed with inhibitors, key points were highlighted which can act as guidelines for design of new inhibitors against *M. tuberculosis* alpha-mannosidase.

2. Materials and methods

2.1. Ligand-based studies

Thirty-eight molecules from our previously reported work [11] (Table S1) with MIC values <25 μg/mL were converted to the corresponding pMIC (-log MIC) to be used as dependent variable for developing the 3D-QSAR model. Correct alignment of selected conformations of molecules is the most crucial step of development of reliable 3D-QSAR models. Therefore, global energy minimum conformation of most active compound **38** was obtained through the simulated annealing where molecule was heated up to 700 K for 1000 fs followed by annealing the molecule to 200 K for 1000 fs and, further followed by BFGS energy minimization. This energy minimized molecule was used as the alignment template, and the rest of the molecules were built and aligned on it subsequently by using the database alignment module (Fig. 1) using Tripos force field [26] and Huckel partial atomic charges implemented in [27]. Further, the dissimilarity matrix was built between all pairs of compounds measured by tanimoto similarity coefficient between the MACCS fingerprints calculated from rcdk package [28]. Hierarchical clustering was done using ward's linkage method [29] for all the thirty-eight molecules using R 2.12.2 package available at [30] to remove subjective bias in the variable selection procedure. Approximately unbiased (AU) *p*-values were obtained via multiscale bootstrap resampling for clusters using pvclust package in R 2.12.2 for assessing the uncertainty in hierarchical cluster analysis. Three-dimensional quantitative structure–activity relationships (3D-QSAR) methodologies: comparative molecular field analysis (CoMFA) [31] and comparative molecular similarity indices analysis (CoMSIA) [31,32] were applied with default values of different parameters on the training set molecules divided on the basis of clustering. Further, the predictive correlation coefficient (r^2_{pred}) based on the test molecules, is computed by using formula

$$r^2_{\text{pred}} = (\text{SD} - \text{PRESS})/\text{SD}$$

where SD is the sum of the squared deviations between the biological activities of the test set and mean activities of training set molecules and PRESS is the sum of squared deviation between predicted and actual activity for every molecule in test set.

2.2. Receptor-based studies

Amino acid sequence for *M. tuberculosis* alpha mannosidase (Uniprot id: P96937) was evaluated further to determine the presence of conserved domains or signature sequences using Interpro (http://www.ebi.ac.uk/interpro/) and pfam (http://pfam.sanger.ac.uk/) databases. The server for homology detection, HHpred [33] which utilizes hidden markov model to search for suitable templates by profile–profile alignment methods together with predicted secondary structure to produce a high quality alignment even for distant homologs, was used for template searching. Next, a multi-template alignment with selected templates (PDB id: 3BVX and 3LVT) together with zinc and swainsonine co-complexed PDB 2WYI was used by modeler 9.10 [34] to generate ten 3D coordinates of the *M. tuberculosis* alpha-mannosidase. Top model was selected on the basis of Dope score and checked by procheck for presence of structural discrepancies. Low complexity regions with no similarity with templates were deleted and rest of the portion was considered for further studies. Zinc binding site was confirmed

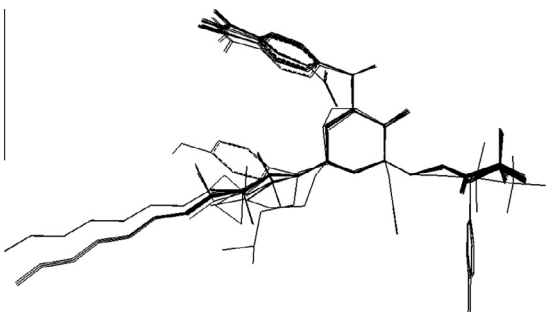

Fig. 1. Database alignment of 38 molecules used in QSAR studies.