Life Sciences 174 (2017) 8–14

Contents lists available at ScienceDirect

Life Sciences

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

Deciphering molecular properties and docking studies of hepatitis C and non-hepatitis C antiviral inhibitors – A computational approach

Arthi Venkatesan, Febin Prabhu Dass J *

Department of Integrative Biology, School of Bio Sciences and Technology, VIT University, Vellore 632014, Tamil Nadu, India

ARTICLE INFO

Article history:
Received 14 November 2016
Received in revised form 24 February 2017
Accepted 28 February 2017
Available online 2 March 2017

Keywords:
Antiviral inhibitors
Binding energy
Hepatitis C virus
Hierarchical clustering
Molecular descriptor
Docking studies

ABSTRACT

Background: Hepatitis C is an infectious liver disease with high mortality rate which is caused by Hepatitis C virus. Several treatment methods have been applied to combat this deadly virus including interferons, vaccine and direct acting antivirals (DAAs). However, the later shows promising effects in HCV treatment with lower adverse effect. Specifically, the DAAs target the non-structural proteins (NS3 and NS5B).
Purpose: The objective of the present study is to hypothesize an alternative antiviral inhibitor for HCV from the available other antivirals.
Methods: Computation of 2D molecular descriptors for the selected antiviral inhibitors followed by clustering the descriptor features. The closely clustered compounds were subjected to the interaction studies against the HCV target protein to validate the cluster result.
Results and discussion: The clustering result showed that indinavir (HIV inhibitor) and AT130 (HBV inhibitor) molecule are close to the HCV inhibitor. The indinavir complexed with NS3 protein shows −5.33 kcal/mol and AT-130 complexed with NS5B protein possess the binding energy of −8.87 kcal/mol. The docking interaction study indicated a better binding affinity than other viral inhibitors.
Conclusion: From the descriptor based feature similarity analysis and the interaction study, it can be concluded that indinavir and AT-130 could be a potential alternative agent for HCV treatment.

1. Introduction

Hepatitis C virus (HCV) is a deadly disease that affects more than 180 million people worldwide. Individuals with HCV infection are at high risk to develop liver cirrhosis and hepatocellular carcinoma that may lead to death [1,2]. HCV is a single, positive-stranded RNA virus that belongs to the genus *hepacivirus* of the Flaviviridae family. It encodes a polyprotein of approximately 3010 amino acids. Polyprotein comprises of four structural and six nonstructural proteins. HCV structural proteins consist of core protein (C), envelope glycoprotein (E1 and E2) and transmembrane protein (P7). The nonstructural proteins are transmembrane protein (NS2), serine protease (NS3), cofactors (NS4A and NS4B), interferon resisting protein (NS5A) and RNA-dependent RNA polymerase (NS5B) [3,4]. Non-structural protein NS3, a serine protease is made up of three domains that possess helicase and protease activity. Similarly, NS4A acts as a cofactor thereby directing the helicase and protease activities [5]. NS5B is an essential polymerase enzyme involved in the lifecycle of the HCV. The role of polymerase enzyme is to amplify the genetic material in endoplasmic reticulum through RNA-dependent RNA polymerase (RdRp). NS5B is composed of three domains such as palm, thumb and finger domain [6]. Antiviral drug development for HCV is mainly focused on serine protease NS3/4A, RNA-dependent RNA polymerase (RdRp) and non-structural proteins [7,8].

The anti-HCV treatments are expensive, consequently, the affected patients do not receive proper treatment [9]. The unavailability of vaccines against this disease makes the treatment strategy incompetent. Vaccine development are still in the early stages of clinical trial [10,11]. The only treatment strategies that prevailed until recent years were the use of ribavirin (RBV) and pegylated interferon (PegIFN). The RBV and PegIFN treatments were mostly unsuccessful because of their severe adverse effects [12]. The first DAA was developed against HIV protease. This is followed by telaprevir and boceprevir, the first approved protease inhibitor drugs targeted against HCV NS3/4A proteins [13,14]. The combination of these two DAAs along with the PegIFN and RBV, a triple therapy depicts a successful range of anti-HCV treatment. Currently, an array of compounds is available against HCV, target protein. For NS3/4A protease, the present drugs are telaprevir, boceprevir, simeprevir, and paritaprevir. Similarly, the interferon

Abbreviations: 2D, Two-Dimensional; 3D, Three-Dimensional; DAAs, Direct Acting Antivirals; FDA, Food and Drug Administration; HCV, Hepatitis C Virus; HBV, Hepatitis B Virus; HCMV, Human Cytomegalo Virus; HIV, Human Immunodeficiency Virus; HSV, Herpes Simplex Virus; Inf A and B, Influenza A and B; NS, non-Structural; PegIFN, Pegylated Interferon; RBV, Ribavirin (RBV); RdRp, RNA dependent RNA Polymerase; RSV, Respiratory Syncytial Virus.

* Corresponding author.
E-mail address: mail2febin@gmail.com (J. Febin Prabhu Dass).

http://dx.doi.org/10.1016/j.lfs.2017.02.014
0024-3205/

resisting NS5A protein inhibitors are namely daclatasvir, ledipasvir, and ombitasvir. NS5B, nucleoside proteins are targeted by sofosbuvir and non-nucleoside target inhibitor is dasabuvir. All these DAAs are approved by US Food and Drug Administration (FDA) [15].

The drawback of triple therapy involves high production cost, time duration and affordability for all the HCV patients. To address this issue, initial computational ligand screening and designing has become integral and inevitable part in the stages of drug designing [16]. The foremost step in small molecule research is the exploration of available drug data knowledge [17]. The major advantage of computational analysis in the drug development process are cost effective as well as time consumption is very less when compared to *in vitro* and *in vivo* experiments. Similarity-based rational drug design approach helps in finding out the similar drugs based on the molecular similarity and molecular diversity. Ultimately these methods help in therapeutic development of HCV. One of the most common features of a drug molecule is its descriptor, which ranges from physical, chemical or structural information. In fact, recent years have witnessed the cross usage of drugs for different diseases because of the feature similarities, a very classic and notable instance is the competitive inhibition of protein targets due to the presence of similar function or structural features in small molecules [18]. Additionally, small molecules display numerous similarities in the form of basic nucleus, ring structure, functional group, and molecular properties, later it has been established in many researches and have yielded promising results [19]. Though there are many descriptors extracting software available as open source and proprietary. In this work, Dragon (6.0.26), is used as a commercial software package that computes around 4885 molecular descriptors that are categorized into 29 logical blocks. This paper explores the similarity in molecular descriptors between HCV inhibitors and other antiviral inhibitors. This large number of features can be classified in machine learning methods. Clustering is one such technique comes under the unsupervised machine learning method. It is a widely used data mining technique that divides the data of a similar and dissimilar group of clusters based on the common features within the dataset. Hierarchical cluster analysis is an agglomerative and unsupervised data mining method which divides or combines existing groups based on the features [20]. These multidimensional data can be easily visualized and the distance between them can be computed. Here, we computed 2168 descriptors to cluster the selected antiviral inhibitors. Further, the cluster based selected non-HCV and HCV inhibitors were analyzed through interaction and affinity analysis with the HCV protein targets.

2. Methodology

2.1. Dataset

A dataset of 82 antiviral inhibitors (44 HCV and 38 other antiviral) are mined from the PubMed literature database. The other antiviral inhibitors includes HBV (14), HIV (14), HSV (5), HCMV (2), Inf A and B (2) and RSV (1). The antiviral inhibitors taken in this study has been further defined with respect to their protein targets across the viral species. The selected inhibitors were classified based on the protease and polymerase of nucleoside and non-nucleoside inhibitors. The corresponding 2 Dimensional (2D) structures of these antiviral inhibitors were retrieved from PubChem database. The final dataset is depicted using their PubChem IDs, molecule name, chemical formula, virus and its corresponding target proteins as shown in Table S1 [21]. A schematic representation of the workflow is presented in Fig. 1.

2.2. Computation of molecular descriptor for antiviral inhibitors

The two-Dimensional (2D) descriptors can be derived from the molecular graphs also called as topological descriptors which are conformationally independent [22]. A total of 2168 descriptors were calculated for all 82 inhibitors using Dragon software. There are 18 logical blocks which include constitutional indices, ring descriptors, topological indices, walk and path counts, connectivity indices, information indices, 2D matrix based descriptors, 2D autocorrelations, Burden eigenvalues, P_VSA-like descriptors, ETA indices, edge adjacency indices, functional group counts, atom-centered fragments, atom-type E-state indices, 2D atom pairs, molecular properties and drug like indices [24]. These descriptor features of every antiviral inhibitors are subjected to unsupervised clustering techniques.

2.3. Hierarchical clustering

Hierarchical clustering is a commonly used technique in data mining also widely used in cheminformatics methods that aim to build a hierarchy of clusters. The principle behind the similarity analysis of molecular

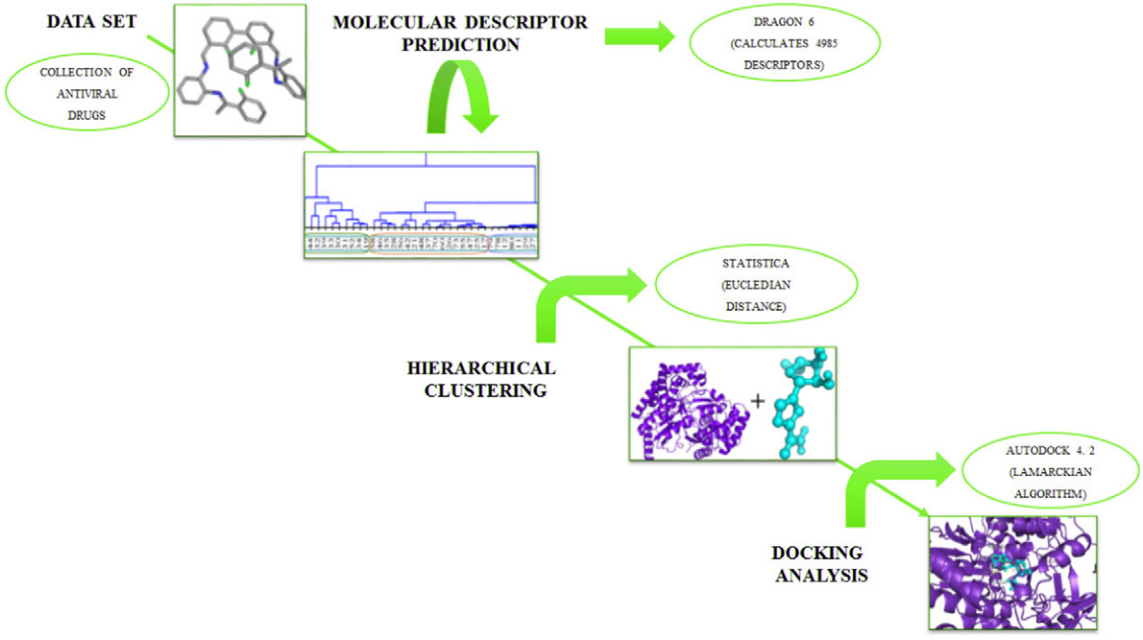

Fig. 1. Outline representation of the protocol.