Available online at www.sciencedirect.com

ScienceDirect

Journal of Taibah University for Science 10 (2016) 563–570

Journal of Taibah University for Science

www.elsevier.com/locate/jtusci

2D QSAR studies of the inhibitory activity of a series of substituted purine derivatives against c-Src tyrosine kinase

Mukesh C. Sharma *

School of Pharmacy, Devi Ahilya University, Takshila Campus, Khandwa Road, Indore, MP 452 001, India

Received 18 May 2015; received in revised form 10 June 2015; accepted 2 November 2015
Available online 21 December 2015

Abstract

A series of 34 substituted purine analogues derivatives were subjected to quantitative structure-activity relationship analyses as inhibitors of c-Src tyrosine kinase. Partial least squares regression was applied to derive QSAR models, which were further validated for statistical significance by internal and external validation. The best QSAR model developed had a good predictive correlation coefficient (r^2) of 0.8319, a significant cross-validated correlation coefficient (q^2) of 0.7550, and an r^2 for the external test set (pred_r^2) of 0.7983. It was developed from the PLS method with descriptors including the $SsCH_3E$-index, H-Donor Count, T_2_Cl_3, and negative correlation with SsOHcount. The current study provides better insight into the future design of more potent c-Src tyrosine kinase inhibitors prior to synthesis.

Keywords: Purine; Quantitative structure activity relationship (QSAR); 2D descriptors; c-Src tyrosine kinase; Partial least squares (PLS)

1. Introduction

Protein tyrosine kinases are enzymes that catalyze the transfer of the terminal ATP phosphate to specific tyrosine residues present on a target substrate [1]. Protein tyrosine kinases regulate signalling pathways for a broad spectrum of cellular processes, including responses to growth factors, neurotransmitters and hormones; activation of the immune response; and the regulation of cell-cell and cell-extracellular matrix interactions, as well as development, oncogenesis, and angiogenesis [2,3]. The protein c-Src kinase is a nonreceptor tyrosine kinase that acts as a signal transduction inhibitor and is a critical component of multiple signalling pathways controlling cell growth, proliferation, invasion, and apoptosis. While c-Src kinase is highly regulated and active only at low levels in most normal cells, studies have shown its upregulation in many human tumour types [4,5]. Protein tyrosine kinases (PTKs) are enzymes responsible for the phosphorylation of other proteins and can catalyze the transfer of the γ-phosphate group of ATP to protein phenolic groups (on Tyr). PTKs play a central role in signal transduction pathways and are involved in the immune, endocrine, and nervous system physiology and pathology [6]. The enzyme c-Src

* Tel.: +91 9826372944.
E-mail address: mukeshcsharma@yahoo.com
Peer review under responsibility of Taibah University

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jtusci.2015.11.002

tyrosine kinase plays versatile roles in cell responses induced by platelet-derived growth factor (PDGF), including cell growth, cell cycle progression, cell survival, cell migration, actin cytoskeleton rearrangement, DNA synthesis and receptor endocytosis [7,2]. Src kinases play crucial roles in signal transduction pathways and in regulating various cellular functions such as cell proliferation and cell differentiation [8]. The activity and structural conformation of the Src-family protein kinases are mainly regulated by phosphorylation events [9]. Src kinase is a protooncogenic tyrosine kinase [10] and has been implicated in the genesis and progression of multiple types of human cancer, including colon, breast, lung, and other cancers [11]. Quantitative structure–activity relationships (QSARs) are an attempt to correlate the structural or property descriptors of compounds quantitatively with biological activities. The traditional 2D-QSAR model is only a rough approximation of the real relationships, as it mainly uses molecular descriptors. QSAR models, mathematical equations relating chemical structure to biological activity, provide useful information for drug design and medicinal chemistry [12]. Quantitative structure–activity relationships (QSARs) help to predict the biological activity of new structures and may reveal useful information on structural modification at several substitutional positions of c-Src-binding molecules [13–16]. This work was undertaken to find a correlation between physicochemical parameters and the biological activity of a series of novel purine derivatives as c-Src tyrosine kinase analogues. This paper is an attempt at a predictive technique based on partial least squares regression, which identifies key structural features responsible for governing c-Src tyrosine kinase.

2. Materials and method

2.1. Computational method

Computational studies were performed on an HP with Windows 7 Home Basic running on an Intel® core processor. The molecular structures of the compounds in the data set were sketched using the V-life MDS (Molecular Design Suite)^{TM} 3.5 software supplied by V-life Sciences Technologies [17]. Analogues of purine derivatives reported to have potent and selective inhibitory activity against c-Src tyrosine kinase were taken from the literature [18]. The biological assay used to test the activity of all of the molecules was the same, and hence, the inhibition values indicated by IC_{50} are comparable. The biological activities represented by IC_{50} were converted into the corresponding pIC_{50} values ($-\log IC_{50}$), which were used as dependent variables in the QSAR analysis. For this study, a total of 34 purine derivatives were divided into training and test sets consisting of twenty-six and eight compounds (Table 1), respectively.

The sphere exclusion method [19] was adopted to divide the training and test data set comprising of twenty-six and eight compounds, respectively, with a dissimilarity value of 8.6, where the dissimilarity value gives the sphere exclusion radius. This algorithm allows the construction of training sets covering all descriptor space areas occupied by representative points. Eight compounds, namely 3, 6, 8, 10, 19, 23, 27 and 31, were used as the test set, while the remaining molecules were used as the training set. Initially, the data set was split into training (70%) and test sets (30%) using the MDS software. Care was taken to achieve an even distribution of activities in both sets (training and test).

To perform the QSAR analysis, the structures of the compounds in the data set were sketched, and the physicochemical descriptors of the molecules were calculated using the V-life MDS (molecular design suite) software. All of the compounds were batch optimized to minimize energies and optimize the geometry using Merck molecular force fields, followed by considering the distance-dependent dielectric constant of 1.0, the convergence criterion or root mean square (RMS) gradient of 0.01 kcal/mol Å and the iteration limit of 10,000 [20].

2.2. Two-dimensional QSAR

A large number of theoretical descriptors, such as SA Most Hydrophilic (most hydrophilic value on the vdW surface), SA Most Hydrophobic–Hydrophilic Distance (distance between most hydrophobic and hydrophilic point on the vdW surface), SA Hydrophilic Area (vdW surface descriptor showing hydrophilic surface area) and SK Most Hydrophilic, the radius of gyration, Wiener's index, moment of inertia, semi-empirical descriptors, HUMOEnergy (highest occupied molecular orbital), heat of formation and ionization potential, as well as constitutional, physicochemical, electrostatic, topological and semi-empirical descriptors have been computed from chemical structures with a view to developing the structure-activity relationships of purine compounds, which would, in turn, predict their biological activity.

The independent variables (*i.e.*, descriptors) were preprocessed by removing the invariable values (constant column), which resulted in a total of 280 descriptors for use in QSAR analysis. Descriptors with the same value