Journal of Molecular Structure xxx (2016) 1–8

Contents lists available at ScienceDirect

Journal of Molecular Structure

journal homepage: http://www.elsevier.com/locate/molstruc

QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines

Vijay H. Masand [a, *], Nahed N.E. El-Sayed [b, f], Devidas T. Mahajan [a], Andrew G. Mercader [c], Ahmed M. Alafeefy [d], I.G. Shibi [e]

[a] Department of Chemistry, Vidya Bharati College, Camp, Amravati, Maharashtra, 444 602, India
[b] Department of Chemistry, College of Science, "Girls Section", King Saud University, P.O. Box 2457, Riyadh 11451, Saudi Arabia
[c] Instituto de Investigaciones Fisicoquímicas Teóricas y Aplicadas (INIFTA), UNLP, CCT La Plata-CONICET, Diag. 113 y 64, Sucursal 4, C.C. 16, 1900 La Plata, Argentina
[d] Department of Chemistry, Kulliyyah of Science, International Islamic University, P.O. Box 141, 25710 Kuantan, Malaysia
[e] Department of Chemistry, Sree Narayana College, Chempazanthy, Kerala, India
[f] National Organization for Drug Control and Research, Giza 35521, Egypt

ARTICLE INFO

Article history:
Received 20 August 2016
Received in revised form
2 November 2016
Accepted 3 November 2016
Available online xxx

Keywords:
Anti-HAT activity
QSAR
Substituted 2-Phenylimidazopyridines

ABSTRACT

In the present work, sixty substituted 2-Phenylimidazopyridines previously reported with potent anti-human African trypanosomiasis (HAT) activity were selected to build genetic algorithm (GA) based QSAR models to determine the structural features that have significant correlation with the activity. Multiple QSAR models were built using easily interpretable descriptors that are directly associated with the presence or the absence of a structural scaffold, or a specific atom. All the QSAR models have been thoroughly validated according to the OECD principles. All the QSAR models are statistically very robust (R^2 = 0.80–0.87) with high external predictive ability (CCC_{ex} = 0.81–0.92). The QSAR analysis reveals that the HAT activity has good correlation with the presence of five membered rings in the molecule.

1. Introduction

Sleeping sickness or human African trypanosomiasis (HAT), transmitted by tsetse flies (genus *Glossina*), has a major occurrence in rural populations in sub-Saharan Africa. HAT, considered as a neglected tropical disease, was nearly eradicated in the mid-1960s. The resurgence in the late 1990s, due to poor sanitation and suitable habitats for its vector in the Democratic Republic of the Congo (DRC), Angola, Central African Republic, southern Sudan, and Uganda, received considerable attention of the researchers to develop better diagnosis and treatment for the disease [1–3]. Recent reports indicate that in humans the disease is thought to be mainly caused by *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodesiense*, however the non-human-pathogenic trypanosome species *Trypanosoma brucei brucei*, *Trypanosoma congolense*, and *Trypanosoma evansi* are also responsible in some instances [1]. The disease has two stages; in stage 1 (hemolymphatic) the peripheral infection with non-specific clinical symptoms occur; and in stage 2 the parasite crosses the blood brain barrier (BBB) and intrudes the central nervous system (CNS) [1–3].

Suramin and pentamidine are the recommended drugs for stage 1 infection, whereas for stage 2, the therapeutic options are melarsoprol, eflornithine and the currently used combination therapy NECT (nifurtimox and eflornithine combination therapy). Unfortunately, vaccine cannot be developed due to a high degree of antigenic variation. In addition, the treatment is parasite- and stage-specific, depending on the ability of the compound to cross the BBB. For BBB clearance the drug must be sufficiently lipophilic, which results in poor water solubility, hence, such drugs are mostly toxic and problematical to administer. Consequently, the available drugs for stage 2 of the disease exhibit high toxicity, involve the complexity of administration procedures and progressive loss of efficacy in some geographical regions. Recent efforts identified Fexinidazole, furamidine, DB289 (parafuramidine), CPD-0802 (an aza analogue of parafuramidine) and SCYX-7158 (a boron based compound) as attractive lead/targets in the drug pipeline for

Abbreviations: HAT, Human African Trypanosomiasis; GA, Genetic algorithm; MLR, Multiple linear Regression; QSAR, Quantitative structure-activity analysis; WHO, World health organization; ADMET, Absorption, Distribution, Metabolism, Excretion and Toxicity; OLS, Ordinary Least Square; QSARINS-Chem, QSAR Insubria-Chemistry; OECD, Organization for Economic Co-operation and Development.

* Corresponding author.
E-mail address: vijaymasand@gmail.com (V.H. Masand).

http://dx.doi.org/10.1016/j.molstruc.2016.11.012
0022-2860/

developing better therapeutics for HAT (see Fig. 1). Despite the previous efforts executed high toxicity, poor oral bioavailability and blood–brain barrier penetration are the major obstacles ahead for these clinical candidates. Thence, the search for a drug candidate with adequate activity, ADME and toxicity profile still persists [1–4].

Recently, Tatipaka et al. [5] identified substituted oxazolopyridine **1** (see Fig. 2) as an attractive lead due to good whole-cell activity on *T. brucei*, no cytotoxicity on mammalian cell lines, acceptable exposure in the central nervous system, and satisfactory aqueous solubility. But, its poor metabolic stability in liver microsomes appeared as a severe liability. Later, to design a better analogue of **1** with the desired profile, they synthesized and screened a series of substituted 2-Phenylimidazopyridines. Since, the mechanism of action and the specific target with which these analogues interact is unknown [5]; in such a situation, a good strategy for lead optimization is to employ computer aided drug design (CADD) using the available information. Hence, in the present work, ligand based drug design technique, viz. QSAR (2D- and 3D-) has been executed to determine the structural features having a significant correlation with the HAT activity profile of substituted 2-Phenylimidazopyridines.

In the past decades, CADD has appeared as a thriving option to conventional 'trial and error' methodology of drug design/discovery to unknot the mysteries of structural patterns that govern the activity, pharmacokinetics, pharmacodynamics and toxicity profiles of a drug candidate. CADD is relatively fast, economical and significantly result oriented successful *in-silico* technique [6–10]. It encompasses a combination of different ideas, algorithms, tools and techniques of various scientific fields like computer, mathematics, statistics, etc. Its major emphasis is on simulation of interactions of different molecules, to determine the reasons behind the specific interactions of different molecules and identification of effective structural features associated with activity/toxicity. QSAR, molecular docking, pharmacophore modeling, etc. are established CADD methods, which when used in harmony provide significant and unrivalled information essential for lead/drug optimization [4,11–15]. These methods have been widely used for identification of the structural patterns that govern the specific activity/toxicity of drug candidates and provide better insight into the mechanism of drug action.

The main objective of the present work is to develop statistically robust and easily interpretable, in terms of structural fragments or specific atom, QSAR models with high external predictive ability.

2. Experimental methodology

2.1. Experimental datasets

In the present work, HAT inhibition activities of sixty substituted 2-Phenylimidazopyridines comprising different heterocyclic scaffolds and diverse substituents at various positions covering a meaningful portion of the chemical space were subjected to QSAR modeling [5]. The reported EC_{50} (μM) values for HAT activity were converted to pEC_{50} ($-\log_{10}EC_{50}$) before QSAR analysis. The EC_{50}, pEC_{50} and the substituents on 2-Phenylimidazopyridine moiety have been listed in Table 1.

2.2. Modeling and molecular descriptors calculation

In the present work, a QSAR analysis following the standard procedure recommended by OECD and different researchers was exercised [16–26]. The chemical structures were drawn using ChemSketch 12 freeware followed by energy minimization using MMFF94 force field in TINKER [4,12,15]. The optimized structures were used as input for the calculation of a good number of 1-3D, electro-topological, fingerprints and other descriptors. Two descriptor calculating softwares were used: PaDEL 2.21 and e-Dragon. Since, all the calculated descriptors (>18,000) do not contain significant information; objective feature selection was employed to reduce the descriptor pool. Nearly constant (>95%), constant, and highly correlated (|R| > 95%) descriptors were eliminated before subjective feature selection (SFS) using QSARINS-Chem 2.2.1 [16,17,20]. This resulted in a reduced cluster of 345 descriptors only. The very next step involved the elimination of highly esoteric descriptors, the descriptors for which an exact explanation is not available or it is difficult to interpret it in terms of structural features [26]. This led to a set of only 253 easily interpretable descriptors. The reduced set still consists a wide range of theoretical molecular descriptors that takes into account different structural features, viz. constitutional (0D-), mono-dimensional (1D-), bi-dimensional (2D-) and three-dimensional (3D-), capturing and magnifying the diverse aspects of the chemical structures.

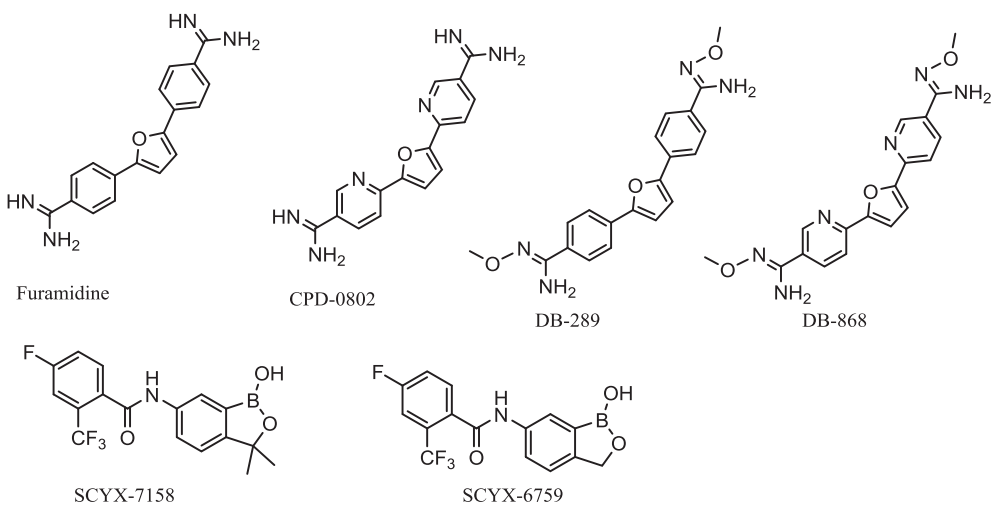

Fig. 1. Chemical structures of clinical drug candidates against HAT.

Please cite this article in press as: V.H. Masand, et al., QSAR modeling for anti-human African trypanosomiasis activity of substituted 2-Phenylimidazopyridines, Journal of Molecular Structure (2016), http://dx.doi.org/10.1016/j.molstruc.2016.11.012