European Journal of Medicinal Chemistry 69 (2013) 99–114

Contents lists available at ScienceDirect

European Journal of Medicinal Chemistry

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

Original article

QSAR modeling of aromatase inhibitory activity of 1-substituted 1,2,3-triazole analogs of letrozole

Chanin Nantasenamat [a,b,*], Apilak Worachartcheewan [a], Supaluk Prachayasittikul [a], Chartchalerm Isarankura-Na-Ayudhya [b], Virapong Prachayasittikul [b]

[a] Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand
[b] Department of Clinical Microbiology and Applied Technology, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand

ARTICLE INFO

Article history:
Received 22 February 2013
Received in revised form
28 July 2013
Accepted 7 August 2013
Available online 19 August 2013

Keywords:
Aromatase
Breast cancer
Triazole
Letrozole
QSAR
Structure–activity relationship
Chemical space

ABSTRACT

Aromatase is an estrogen biosynthesis enzyme belonging to the cytochrome P450 family that catalyzes the rate-limiting step of converting androgens to estrogens. As it is pertinent toward tumor cell growth promotion, aromatase is a lucrative therapeutic target for breast cancer. In the pursuit of robust aromatase inhibitors, a set of fifty-four 1-substituted mono- and bis-benzonitrile or phenyl analogs of 1,2,3-triazole letrozole were employed in quantitative structure–activity relationship (QSAR) study using multiple linear regression (MLR), artificial neural network (ANN) and support vector machine (SVM). Such QSAR models were developed using a set of descriptors providing coverage of the general characteristics of a molecule encompassing molecular size, flexibility, polarity, solubility, charge and electronic properties. Important physicochemical properties giving rise to good aromatase inhibition were obtained by means of exploring its chemical space as a function of the calculated molecular descriptors. The optimal subset of 3 descriptors (i.e. number of rings, ALogP and HOMO–LUMO) was further used for QSAR model construction. The predicted pIC_{50} values were in strong correlation with their experimental values displaying correlation coefficient values in the range of 0.72–0.83 for the cross-validated set (Q_{CV}) while the external test set (Q_{Ext}) afforded values in the range of 0.65–0.66. Insights gained from the present study are anticipated to provide pertinent information contributing to the origins of aromatase inhibitory activity and therefore aid in our on-going quest for aromatase inhibitors with robust properties.

1. Introduction

Breast cancer is considered to be the most common type of cancer and is the leading cause of cancer-related death in women accounting for an estimated 23% of new cases and 14% of all cancer deaths in 2008 [1]. It is widely accepted that the majority of breast cancers are hormone-dependent and that estrogen is a key mediator in the progression and metastasis of breast tumors. Particularly, for postmenopausal women it has been reported that the concentration of 17β-estradiol (E2) in breast tumor can be ten-fold higher than those in plasma [2]. The high concentration of E2 in breast tumors could be attributed to increased uptake from plasma or in situ aromatization of androgens to estrogens [3]. The latter is afforded by aromatase, an enzyme involved in the rate-limiting step of estrogen biosynthesis by catalyzing three consecutive hydroxylation reactions that aromatizes C19 androgens to C18 estrogens. Thus, blockade of the synthesis of estrogens by inhibiting aromatase represents a lucrative therapeutic approach for the treatment of hormone-sensitive breast cancer [4,5].

Aromatase inhibitors (AIs) are comprised of 2 major classes of compounds according to their molecular structure and mechanism of action. Type I or steroidal inhibitors typically contain an androgen core structure and snugly binds to the substrate-binding site while type II or non-steroidal inhibitors are typically comprised of azole moieties that bind to the heme co-factor of the enzyme. The former class usually binds irreversibly to the substrate-binding site and is thus termed suicidal inhibitors as they inactivate the enzyme while the latter class binds reversibly by positioning the nitrogen atom from its azole moiety to coordinate with the iron atom of the heme co-factor.

* Corresponding author. Center of Data Mining and Biomedical Informatics, Faculty of Medical Technology, Mahidol University, Bangkok 10700, Thailand. Tel.: +66 2 441 4371; fax: +66 2 441 4380.
E-mail address: chanin.nan@mahidol.ac.th (C. Nantasenamat).

0223-5234/$ – see front matter © 2013 Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.ejmech.2013.08.015

Three generations of AIs had been introduced over the past three decades. The third generation AIs demonstrated good tolerability profiles, higher selectivity and potency when compared with the first and second generation AIs [6,7]. Particularly, the first and second generation AIs had been shown to inhibit *in vivo* estrogen synthesis by up to 90% whereas the third generation AIs provided greater than 98% inhibition [8]. Letrozole (also known as CGS 20267) is a third generation AI that was first introduced by Novartis (then Ciba-Geigy) as a non-steroidal AI possessing potent pharmacological profile [9]. In 1996 it was approved in Europe and this is followed by its subsequent approval in the United States in 1997 [10]. The compound is marketed as Femara and has an IUPAC name of 4-[(4-cyanophenyl)(1H-1,2,4-triazol-1-yl)methyl]benzonitrile. Letrozole has been particularly demonstrated to be a highly potent AI both *in vitro* and *in vivo*. Particularly, letrozole exhibited higher potency than several AIs such as aminoglutethimide, anastrozole, exemestane and formestane [9,11,12].

The seminal work by Hansch et al. [13] in correlating the biological activity of plant growth regulators with Hammett constants and partition coefficients popularized the concept of quantitative structure–activity relationship (QSAR). Over the years, the QSAR paradigm has been employed in modeling a wide range of biological activities (i.e. spectral properties of green fluorescent protein [14,15], antioxidant activity [16,17], quorum quenching of *N*-acyl homoserine lactone lactonase [18], lipopolysaccharide neutralization activity of anti-endotoxins [19], furin inhibition [20], anti-prion activity [21] and anti-cancer activity [22]), chemical properties (i.e. imprinting factor of molecularly imprinted polymers [23,24]) and medical conditions (i.e. prediction of ischemic heart disease [25] and identification of metabolic syndrome [26,27]). The concepts of QSAR modeling had been reviewed previously [28,29]. Briefly, QSAR constructs statistically validated models that are capable of quantitatively predicting the bioactivity of the explored compounds. This is performed by employing multivariate analysis methods to discern linear or non-linear relationships between the molecular structures (i.e. physicochemical descriptors) and their respective biological activities (i.e. aromatase inhibitory activity).

Herein, a series of 1-substituted mono- and bis-benzonitrile or phenyl analogs of 1,2,3-triazole letrozole are employed for QSAR modeling of the aromatase inhibitory activity. A diverse set of quantum chemical and molecular descriptors were employed to provide numerical description of the investigated compounds. Descriptors accounting for the unique structural features of the compounds were correlated to their respective aromatase inhibitory activity using multiple linear regression (MLR), artificial neural network (ANN) and support vector machine (SVM).

2. Results and discussion

2.1. Structure–activity relationship of 1,2,3-triazole letrozole analogs

Letrozole is a non-steroidal AI known by its IUPAC name of 4-[(4-cyanophenyl)(1H-1,2,4-triazol-1-yl)methyl]benzonitrile. As can be seen from Fig. 1 the chemical structure of letrozole **1** is comprised of a 1,2,4-triazole and 1-substituted bis-benzonitrile. Mechanistically, the azole functional moiety coordinate its N4-atom to interact with Fe^{2+} of the heme prosthetic group while its two phenyl rings provide tighter fit inside the binding cavity by mimicking the steroidal backbone of the natural substrate, androstenedione. Furthermore, the two nitrile groups at the para positions of the phenyl ring mimics the carbonyl group of androstenedione and functions as hydrogen bond acceptors. Previous investigation on structure–activity relationship signifies the importance of electron withdrawing groups at the para position of the phenyl ring while demonstrating that nitrile groups afforded the best activity [30]. This is in concomitant with the findings by Schuster et al. [31] that two aromatic rings along with two hydrogen bond donors are important pharmacophores for strong aromatase inhibition. Moreover, the importance of hydrogen bonding in aromatase inhibition was suggested by Neves et al. [32] from their computational analysis of potent AIs. In a comparative analysis of the structures of letrozole and androstenedione, Doiron et al. [33] observed that distances between nitrogens of the nitrile group and those at positions 3 or 4 of the triazole heterocycle are similar to distances between oxygens in the structure of androstenedione.

Several findings have suggested that letrozole is a highly potent AI both *in vitro* and *in vivo* as well as exhibiting higher potency than several AIs such as aminoglutethimide, anastrozole, exemestane and formestane [9,11,12]. Owing to the success of letrozole, intense efforts have been invested in deriving novel AIs from this structural scaffold. In a series of investigations, Le Borgne et al. [34–36] synthesized several letrozole analogs possessing arylindole moiety with imidazole or 1,2,4-triazole. Furthermore, Farag et al. [37,38] synthesized several pyrazole-based letrozole and celecoxib analogs affording aromatase inhibitory activities. Moreover, Potter et al. [39–43] synthesized an array of sulfamate-containing letrozole, anastrozole and vorozole analogs, which provided interesting dual inhibitory activities toward aromatase and steroid sulfatase. Recently, Doiron et al. [33] sought out to investigate the aromatase inhibitory activity of 1,2,3-triazole instead of the 1,2,4-triazole present in letrozole. In their study, a library of substituents at the N1 position of the triazole heterocycle was examined by observing the influence of one or two phenyl/benzonitrile on the aromatase inhibitory activity. The authors concluded that nitrogen atom at positions 3 or 4 were both important for aromatase inhibition. A subset of compounds from their investigation comprising of fifty-four 1-substituted mono- and bis-benzonitrile or phenyl analogs of 1,2,3-triazole letrozole were compiled as a data set for QSAR investigation in this study (Table 1). A schematic representation of the computational methodology used herein is depicted in Fig. 2.

2.2. Exploring the chemical space of 1,2,3-triazole letrozole analogs

The linkage between the molecular structures of compounds with its respective biological activities is central to the QSAR paradigm. Molecular descriptors play a crucial role in providing numerical description of the physicochemical properties of molecules. In order to properly account for these structural features, it is essential that suitable descriptors be chosen for QSAR investigation. A handbook providing comprehensive coverage of molecular descriptors has been summarized by Todeschini and Consonni [44] while an in-depth review of quantum chemical descriptors was provided by Karelson et al. [45] In spite of the wide availability and relatively large number of molecular descriptors to choose from, this study selected a small subset of descriptors representing the general characteristics of a molecule (i.e. molecular size, flexibility, polarity, solubility, charge and electronic properties as well as chemical reactivity) and whose physicochemical properties can be easily understood. Such descriptors were recently used in exploring the chemical space of all known aromatase inhibitors [46].

Geometry optimization of the molecular structures was performed in a two-step fashion starting from an initial optimization with the semi-empirical AM1 method as to afford reasonably good starting structures followed by a subsequent and more refined optimization at the DFT level. The resulting low-energy conformers served as the basis for the extraction of 6 quantum chemical descriptors followed by a subsequent calculation using the Dragon