Available online at www.sciencedirect.com
ScienceDirect
Journal of Molecular Graphics and Modelling 26 (2008) 884–892
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Pharmacophore mapping of arylbenzothiophene derivatives for MCF cell inhibition using classical and 3D space modeling approaches☆

Subhendu Mukherjee [1], Shuchi Nagar [1], Sanchita Mullick, Arup Mukherjee, Achintya Saha *

Department of Chemical Technology, University of Calcutta, 92, A.P.C. Road, Kolkata 700009, India

Received 27 February 2007; received in revised form 11 June 2007; accepted 11 June 2007
Available online 16 June 2007

Abstract

Considering the worth of developing non-steroidal estrogen analogs, the present study explores the pharmacophore features of arylbenzothiophene derivatives for inhibitory activity to MCF-7 cells using classical QSAR and 3D space modeling approaches. The analysis shows that presence of phenolic hydroxyl group and ketonic linkage in the basic side chain of 2-arylbenzothiophene core of raloxifene derivatives are crucial. Additionally piperidine ring connected through ether linkage is favorable for inhibition of breast cancer cell line. These features for inhibitory activity are also highlighted through 3D space modeling approach that explored importance of critical inter features distance among HB-acceptor lipid, hydrophobic and HB-donor features in the arylbenzothiophene scaffold for activity.

Keywords: Pharmacophore mapping; Arylbenzothiophene; MCF cell-line inhibition; QSAR study; Space modeling

1. Introduction

Estrogen is known to play an important role in reproductive endocrinology, involved in growth and function of other tissues, such as the skeleton, cardiovascular and central nervous systems in both male and female [1,2]. It is also important for supporting homeostasis in a women's body, as evidenced by the progressive changes that occur at menopause, when ovarian estrogen synthesis stops around the age of 45. The decreased production of estrogen leads to certain postmenopausal pathologies, such as osteoporosis and coronary artery diseases (CAD) [3,4]. Hormone replacement therapy (HRT, specifically estrogen replacement therapy) is primarily used for the treatment of postmenopausal diseases and is found to be very beneficial for the treatment of osteoporosis. Despite the beneficial effects of HRT, it is observed that this therapy also encourages the development and growth of cancer in the breast and uterus. Thus search began for the treatment of the women reproductive cancer.

Selective estrogen receptor modulators (SERMs) [5,6] that show tissue-dependent agonistic or antagonistic behavior [7], are used as first line treatment for estrogen responsive breast cancer and for therapy against osteoporosis. The pure estrogen antagonists (anti-estrogens) are currently in clinical development for breast cancer treatment. SERMs are used as alternative and research is still on to obtain a variety of non-steroidal compounds, which interact with the estrogen receptor. These molecules are found to fully antagonize the effects of estrogen on uterine and mammary tissues, while mimicking the effects on bone and cardiovascular system [8]. Several group of compounds have been developed as SERM's, one such compound being raloxifene [9]. It is known that raloxifene is the first SERM to be available for the treatment of osteoporosis and is also being tested as a preventive measure for breast cancer and coronary heart disease (CHD) [9–11]. The present work has been designed to explore the pharmacophore features for MCF cells inhibition of 2-arylbenzothiophene core (Fig. 1) of raloxifene by implementing both classical as well as space modeling studies.

The pharmacophore concept is based on the kinds of interaction observed in molecular recognition, i.e., hydrogen bonding, charge and hydrophobic interaction and alternatively

☆ Presented in the 11th Annual International Conference on Research in Computational Molecular Biology (RECOMB 2007), San Francisco Bay Area, California, USA, April 21–25, 2007.

* Corresponding author. Tel.: +91 33 23508386; fax: +91 33 23519755.
E-mail address: achintya_saha@yahoo.com (A. Saha).

[1] Both the authors equally contributed in the work.

1093-3263/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2007.06.003

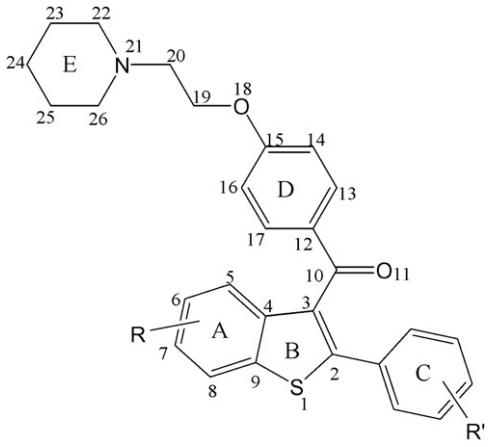

Fig. 1. General structure of arylbenzothiophene derivative.

can be used as a query in a 3D database search to identify new structural classes of potential lead compounds; and it can serve as a template for generating alignment for 3D QSAR analysis [12]. Two types of pharmacophore hypothesis are well established: receptor-based and receptor-independent pharmacophore. Receptor based pharmacophore mapping of SERMs [13–17] are commonly employed, but now a days receptor-independent pharmacophore mapping is growing interestingly for deriving bioactivities of diverse group of compounds, such as aromatase inhibitors [18], serotonin inhibitors [19] and antithrombin [20] agents and are now increasingly being handled by automated computational methods, such as *CATALYST* [21], *GASP*, *DISCO* which are commercially available programs [22–24]. Consequently, the present work is taken up to study the arylbenzothiophene scaffold as a small ligand [25] with a view to deduce the active pharmacophore signals based on receptor-independent hypothesis, using the CATALYST program [21], that can eventually aid in apprehending the tissue-specific effects of different compounds containing this unit.

2. Materials and methods

In the present study, 69 compounds of a series of raloxifene analogs which contain modifications of the 2-arylbenzothiophene core [25] are considered (Table 1) and are randomly segregated into training (Tr) and test (Ts) sets. For classical QSAR studies, the biological activity is expressed as logarithm to the base 10 of IC_{50} (pIC_{50}) to MCF-7 cells. Molecular (partition coefficient [26], hydrophobicity [26], steric [27,28] and thermodynamic factors [29], bulk, moments and orbital energies), electronic (Wang-Ford atomic charge [30] and extended Huckel partial charge [31–33] functions) and electrotopological (*E*-state indices) [34] properties have been explored for deriving classical QSAR models of MCF-7 cell-line inhibition of arylbenzothiophene derivatives using JAVA based program *ETSA–CS* [35], *TSAR* 3.3 [36], *CAChe* 6.1 [37] and *Chem 3D Pro* [38]. The common atoms of this group of compounds have been numbered for computation of charge and electro-topological functions. The indicator variables used for modeling the bioactivities are $I_{7\text{-OH}}$ and $I_{\text{Subs\_6}}$ that signify the presence of hydroxy substitution at atom C_7 and substitutional requirement at atom C_6, respectively.

Statistical analysis is performed by *Statistica* 5.0 [39] using standard and forward stepwise multiple regression methods. The different statistical parameters of the regression equation considered are: r or R (correlation coefficient), EV (explained variance), F (variance ratio) with d.f. (degree of freedom), s (standard error of estimate) and AVRES (average of absolute values of residuals). Leave-One-Out (LOO) cross-validation [40] is performed that generated PRESS (predictive residual sum of squares), SDEP (standard deviation of error of predictions), Pres_{av} (average of absolute value of predicted residuals) and Q^2 (cross-validated variance).

Pharmacophore space modeling study is also performed with this dataset using *CATALYST* 4.11 [21]. The biological activity is expressed as inhibitory concentration (IC_{50}) to MCF-7 cell-line. The pharmacophore models (hypotheses), generated by CATALYST [21], consist of an array of features necessary for bioactivity of the ligands arranged in 3D space that can explain the variance in activity of the molecules w.r.t. geometric localization of the chemical features present in them. To be retrieved as a hit, a candidate ligand must possess appropriate functional groups which can simultaneously reside within the respective tolerance spheres of the pharmacophoric features. Each feature is associated with a weight (a measure of its proposed importance to the pharmacophore as a whole), and the better the overall superimposition of functional groups of the molecule to the appropriate features of the pharmacophore, the higher the score of the fit [22].

Different control parameters employed for hypothesis generation (called a HypoGen process) are spacing, uncertainty and weight variation. Spacing is a parameter representing the minimum inter-features distance that may be allowed in the resulting hypothesis. In the present work, the value of spacing used is 600 pm to accommodate maximum number of chemical features in the hypothesis. In the generated hypothesis, each feature signifies some degree of magnitude of the compound's activity. The level to which this magnitude is explored by the hypothesis generator is controlled by the weight variation parameter. This is varied in some cases from 1 to 2. In other cases, the default value of 2 is generally considered. The uncertainty parameter reflects the error of prediction and denotes the standard deviation of a prediction error factor, called the error cost. A default value of 3 is considered as the uncertainty parameter in the present work. While generating hypothesis, a total cost function is minimized comprising of three terms, viz. weight cost, error cost and configuration cost. Weight cost is a value that increases as the weight variation in the model deviates from an ideal value of 2. The deviation between the estimated activities of the training set and their experimentally determined values is the error cost. A fixed cost that depends on the complexity of the hypothesis space being optimized is denoted as the configuration cost. The configuration cost is equal to the entropy of hypothesis space. The CATALYST program also calculates the cost of null hypothesis (null cost) that assumes no relation in the data, and the