Journal of Advanced Research (2016) xxx, xxx–xxx

Cairo University

Journal of Advanced Research

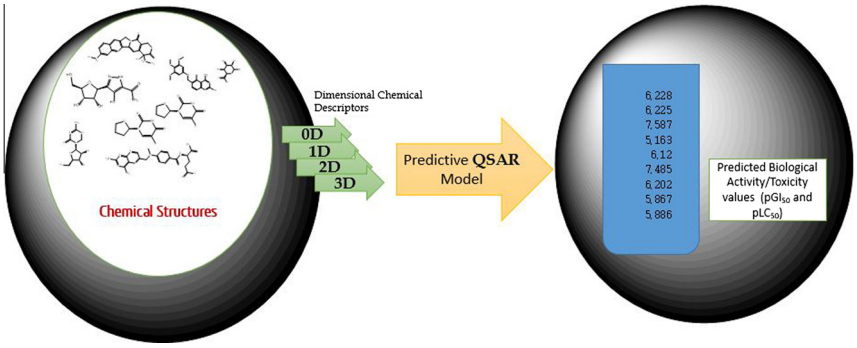
ORIGINAL ARTICLE

# Quantitative structure–activity relationship study on potent anticancer compounds against MOLT-4 and P388 leukemia cell lines

**David Ebuka Arthur** *, **Adamu Uzairu, Paul Mamza, Steven Abechi**

*Department of Chemistry, Ahmadu Bello University (ABU) Zaria, Kaduna State, Nigeria*

## GRAPHICAL ABSTRACT




## ARTICLE INFO

*Article history:*
Received 17 December 2015
Received in revised form 29 March 2016
Accepted 31 March 2016
Available online xxxx

## ABSTRACT

A quantitative structure–activity relationship (QSAR) study was carried out on 112 anticancer compounds to develop a robust model for the prediction of anti-leukemia activity ($pGI_{50}$) against MOLT-4 and P388 leukemia cell lines. The Genetic algorithm (GA) and multiple linear regression analysis (MLRA) were used to select the descriptors and to generate the correlation models that relate the structural features to the biological activities. The final equations consist of 15 and 10 molecular descriptors calculated using the paDEL molecular descriptor software. The GA-MLRA analysis showed that the Conventional bond order ID number of order 1

* Corresponding author. Tel.: +234 8138325431.
E-mail address: hanslibs@myway.com (D.E. Arthur).

Peer review under responsibility of Cairo University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.jare.2016.03.010

*Keywords:*
QSAR method
Anticancer
paDEL descriptors
Applicability domain
Cell lines
NCI database

(piPC1), number of atomic composition (nAtomic), and Largest absolute eigenvalue of Burden modified matrix – n 7/weighted by relative mass (SpMax7_Bhm) play a significant role in predicting the anticancer activities of these compounds. The best QSAR model for MOLT-4 was obtained with $R^2$ value of 0.902, $Q^2_{LOO}$ = 0.881 and $R^2_{pred}$ = 0.635, while for P388 cell line $R^2$ = 0.904, $Q^2_{LOO}$ = 0.856 and $R^2_{pred}$ = 0.670. The Y-scrambling/randomization validation also confirms the statistical significance of the models. These models are expected to be useful for predicting the inhibitory activity (pGI50) against MOLT-4 and P388 leukemia cell lines.


## Introduction

Leukemia is a word that is attributed to cancer of the blood cells, which creates uncontrollable quantities of irregular white blood cells in the blood and bone marrow, swarming out ordinary blood cells. The low level of ordinary blood cells makes it difficult for the body to get oxygen to its tissues, battle diseases and control bleeding [1]. QSAR models in modern time are important tool for predicting the inhibition of such aliments via chemotherapeutic means [2]. There are four regular sorts of leukemia the vast majority of which are gathered in view of how rapidly the sickness deteriorates (chronic) and on the platelet the disease begins in (lymphoblastic or myeloid) [3]. Human leukemia is one of the generally analyzed neoplasms. Most leukemia cell lines gain resistance to the different systems that prompt the human cell demise [4]. Melphalan is a chemotherapy medication fitting in with the class of nitrogen mustard alkylating operators, which moderates the development of growth cells in anyone [5], yet it has been found to be inadequate, and hence these outcomes set the path in the consistent interest and quest for a more successful nontoxic leukemia cell inhibitor [4].

Quantitative structure–activity relationship (QSAR) study performs an urgent part in novel drug design and configuration via a ligand-based approach [6]. Such methodologies are unequivocally judgmental to give not just the solid forecast of particular properties of new analogs, but also illustrate the conceivable molecular mechanism of the receptor–ligand interaction [7]. Quantitative structure–activity relationship (QSAR) [8], has been generally utilized for a long time to give quantitative investigation of structure and biological activity relationships of compounds [9,10]. The importance of QSAR application in pharmaceutical industry and risk assessment cannot be over emphasized as review of its growing applications in these areas was reported by Roy et al. [11].

As of late 2015, computer assisted drug design based on QSAR has been effectively utilized to develop new medications for the treatment of cancer [8,12–14], AIDS, SARS, and other ailments. Selassie et al. [15] analyzed the cytotoxicity of complex mono-substituted phenols toward a fast-developing murine leukemia cell line (L1210). Despite the fact that the interest for "in silico" revelation is clear in every aspect of human therapeutics, the field of anti-infective medications has a specific requirement for computational treatment empowering quick distinguishing proof of novel therapeutic leads [16]. The multidrug-resistance (MDR) of tumor cells to chemotherapeutic operators is a noteworthy issue in the clinical treatment of malignancy [17]. It is the capacity of cancer cells exposed to chemotherapeutics to resist a wide scope of medications [18].

The inability of anticancer drugs to mitigate cancer in some cancer cell lines and their accompanying side effects build the quest for novel treatment choices of this illness [8,19]. Goyal and his colleagues showed that curcumin which destroys a few tumor cell lines can as well suppress the immune system [20].

In this work, the activity of anti-leukemia compounds collected from NCI database (Fig. 1) against P388 and MOLT-4 leukemia cell lines was modeled using several statistical tools, including genetic functional algorithm for variable selection, multiple linear regression (MLR) for modeling and Euclidean based applicability domain for outlier detection.

## Material and methods

### *Experimental dataset*

In this study, a dataset of 112 compounds was used to model the relationship between the chemical fingerprints of the compounds and their anticancer activities on human acute lymphoblastic leukemia (MOLT-4) and multidrug-resistant P388 leukemia cell line. The chemical structures of the dataset, NSC and CAS number were taken from the drug discovery and development arm of the National Cancer Institute (NCI). Eligible compounds were determined by reviewing and curating the raw data collected from the literature (NCI

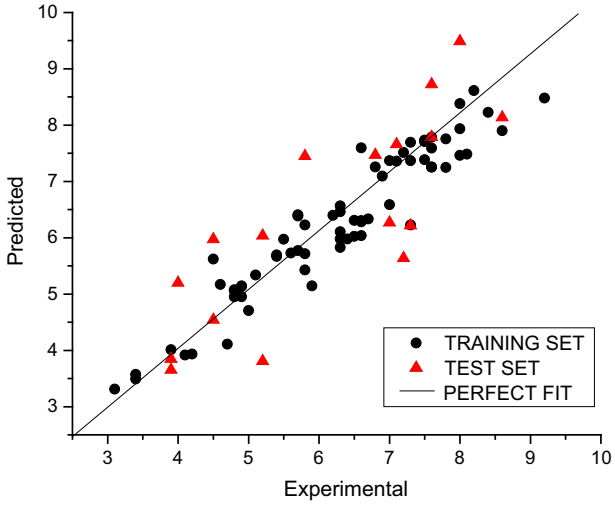



**Fig. 1** The predicted pGI50 against the experimental values for the training and test sets of P388 leukemia cell line.