IOSR Journal of Applied Chemistry (IOSR-JAC)
e-ISSN: 2278-5736.Volume 8, Issue 8 Ver. I (Aug. 2015), PP 57-64
www.iosrjournals.org

Qsar of Ketones Derivatives Using Genetic Function Approximation

AjalaAbduljelil* AdamuUzairu Suleiman Idris
Department Of Chemistry, Ahmadu Bello University Zaria, Kaduna.

Abstract: *Quantitative electronic structural-activity relationship(QSAR) analysis of a series of Ketones and its derivatives against skin disease causing micro-organism(Candida albicans) have been conducted using Molecular weight(MW), Conolly Accessible Area(CAA), cLogp, Molar Refractivity(MR), Molar Volume(MV), Polarizability(POL)as the descriptors. The descriptors were obtained using computational chemistry method (semi-empirical PM3). Ketones and its derivatives activities, were taken as the activities of the molecules against skin disease causing micro-organism(Candida albicans) and are presented as the values of ln(1/IC_{50}), where IC_{50} is an effective concentration inhibiting 50% of the parasite growth. Genetic Function Approximation method was used to generate models that correlate molecular properties of Ketones and its derivatives against skin disease causing micro-organism(Candida albicans) . The best QSAR model generated i.e model-1 has good R^2= 0.986, R^2_{adj} =0.976, R^2(cv)=0.928, $R^2_{predict}$ = 0.968,F-value=96.62. and LOF = 0.0231.The QSAR study provides important structural insights in designing of potent antifungal agents.*

Key words*: QSAR analysis, antifungal agents, Semi-empirical method, Ketones and its derivativesand Genetic Function Approximation(GFA).*

I. Introduction

This chapter addresses medicinal aspects of the treatment of fungal diseases of all types, but because most recent research has been directed toward the treatment of systemic infections, emphasis is placed on this aspect. Candida spp. Invasive candidiasis is the most common nosocomial mycosis, perhaps because the causative organism is a component of the endogenous flora of the human alimentary tract. There has been debate over the significance of positive blood cultures (candidemia) in the progression of fungal disease. Given the high mortality rates (up to 75%) in cases of invasive candidiasis, the current consensus is that all high risk patients with candidemia should receive therapy (Edwards *et al*,1997).Candidiasis is recognized worldwide as an opportunistic infection, this emerged in recent years as troublesome organisms, challenging the supremacy of Candida- albicans in candida infections. The incidence of such disease is relatively constant. Because these are infections of the skin, many of the available agents are applied topically, although a few notable oral alternatives are now available. Despite the relative triviality of the infections, however, their eradication is often problematic and requires many weeks or months of therapy(Donald,2003).

The quantitative structure–activity relationship (QSAR) is an attempt, based on the Structure -activity relationship (SAR) approach, to remove the element of luck from drug discovery. It uses physicochemical properties (parameters) to represent drug properties that are believed to have a major influence on drug action. Parameters must be properties that are capable of being represented by a numerical value. These values are used to produce a general equation relating drug activity with the parameters. This equation enables medicinal chemists to predict the activity of analogues and, as a result, determine which analogue is most likely to produce the desired clinical response. Its use takes some of the guess work out of deciding which analogues of a lead to synthesis. This has the knock-on effect of reducing cost, a major consideration in all commercial companies .Structure–activity relationship studies are usually carried out by making minor changes to the structure of a lead to produce analogues and assessing the effect that these structural changes have on biological activity. The success of the SAR approach to drug design depends not only on the knowledge and experience of the design team but also a great deal of luck. QSAR is an attempt to remove this element of luck from drug design by establishing a mathematical relationship in the form of an equation between biological activity and measurable physicochemical parameters. These parameters are used to represent properties such as lipophilicity, shape and electron distribution, which are believed to have a major influence on the drug's activity. They are normally defined so that they are in the form of numbers derived from practical data that are thought to be related to the property that the parameter represents. This makes it possible either to measure or to calculate these parameters for a group of compounds and relate their values to the biological activity of these compounds by means of mathematical equations using statistical methods such as regression analysis .The main properties of a drug that appear to influence its activity are its, lipophilicity, the electronic effects within the molecule and the size and shape of the molecule (steric effects). Recently, sophisticated computing hardware is readily available, along with a growing number of software packages that are capable of carrying out highly evolved computational

chemistry; the determination of molecular descriptors for use in predictive models of molecular activity is a relatively easy task. In this study, genetic function approximation (GFA) which is a statistical modeling algorithm which builds functional models of experimental data. Since its inception, several applications of this algorithm in the area of quantitative structure–activity relationship modeling have been reported (Roggerss, D. 1996).

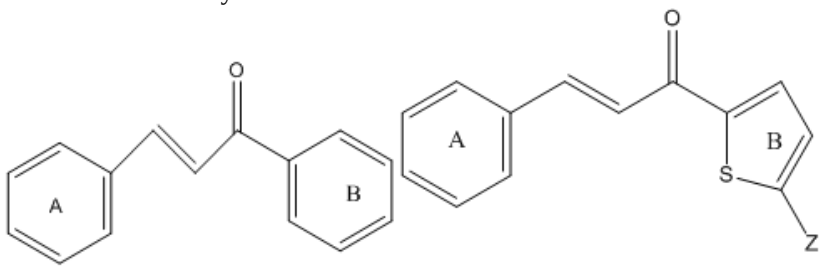
II. Materials And Methods

The purpose of the present work is to perform a quantum chemical QSAR study of the Ketone derivatives on the series (Wanda Pereira Almeida *et al 2011*) to investigate the experimental activities of the compounds as an Antifungal Agents and obtain a linear model by using Genetic function Approximation (GFA) method.

2.1Chemical Data:

Biological data on the activity of Ketone derivatives has been obtained from Wanda Pereira Almeida *et al* (Table 1). The activity data refers pMIC, which indicates the biological activity of compounds experimentally determined, necessary for the inhibition of *candidaalbicans* resistant. The –log MIC(molar) scale refers pMIC .Fluconazole for Table i: was used as controls in the assays.

Table i:
Data set from the literature used in the Quantum Chemical QSAR analysis;
pMIC= -log MIC =Minimum inhibitory concentration in molar



Compound 1-12
Type X

Compound 13- 20
Type Y

Compound	Type	Ring A	Ring B	Substituent Z	pMIC
1	X	4-SCH_3	4-F	-	4.53
2	X	4-SCH_3	4-Cl	-	4.16
3	X	4-SCH_3	4-Br	-	3.52
4	X	4-SCH_3	2,4-Cl	-	3.21
5	X	4-SCH_3	4-NO_2	-	3.48
6	X	4-SCH_3	4-OCH_3	-	4.15
7	X	4-SCH_3	H	-	3.80
8	X	4-SCH_3	4-OH	-	3.83
9	X	4-SCH_3	2-OH	-	4.13
10	X	4-SCH_3	3-OH	-	3.83
11	X	4-SCH_3	4-phenyl	-	3.12
12	X	2,3-OCH_3	4-OCH_3	-	4.47
13	Y	4-SCH_3	-	H	4.72
14	Y	4-SCH_3	-	Br	3.83
15	Y	3,4-OCH_3	-	H	4.14
16	Y	3,4-OCH_3	-	Br	3.55
17	Y	4-phenyl	-	H	3.06
18	Y	4-phenyl	-	Br	3.17
19	Y	4-OCH_3	-	H	4.09
20	Y	4-OCH_3	-	Br	3.51

2.2 Computational And Statistical Details

QSAR studies of Ketone derivatives was carried out on windows7, Intel Pentium operating system by Spartan '14v112 for windows, Macintosh and Linux. PaDEL-Descriptor(A software to calculate molecular descriptors and fingerprints),version: 2.21 and Chem3D pro, version 12.0.2 1076. The molecular structures of the dataset was sketched using Chem Draw Ultra, version 12.0.2.1076 developed by CambridgeSoft.

The first step consisted in obtaining the molecular geometry of all the derivatives from the dataset (Table-i) was energy minimization (Williams *et al*,2002) and geometry optimization using Merck Molecular Force Field (MMFF) in semi-empirical PM3 Method(Halgren,1996).

Table ii: Parameters for Energy Minimization

PARAMETERS	VALUE
Force Field	MMFF(Merck Molecular Force Field)
Maximum no of Cycles	100,000
Convergence Criteria	0.001cal/molÅ
Dielectric constant	1(in a vacuum)
Gradient Type	Analytical

The genetic function approximation (GFA) algorithm offers a new approach to the problem of building quantitative structure-activity relationship (QSAR) and quantitative structure-property relationship (QSPR) models. Replacing regression analysis with the GFA algorithm enables the construction of models competitive with or superior to those produced by standard techniques and makes available additional information not provided by other techniques. Unlike most other analysis algorithms, GFA gives multiple models, where the populations of the models are created by evolving random initial models using a genetic algorithm. GFA can build models using not only linear polynomials but also higher-order polynomials, splines, and other nonlinear functions.

The genetic algorithms are search algorithms that take inspiration from natural genetics and evolution. In this section, the ideas underlying genetic algorithms are briefly described, emphasizing the aspects relevant to the genetic function approximation (GFA) approach to model building. The GFA algorithm itself applies these ideas to the problem of function approximation (Hopfinder, A.J. 1994)given a large number of potential factors influencing a response, including several powers and other functions of the raw inputs, to find the subset of terms that correlates best with the response. The central ideas of genetic algorithms are simple. The region to be searched is coded into one or multiple strings. In the GFA, these strings are sets of terms – powers and splines of the raw inputs. Each string represents a location in the search space.The algorithm works with a set of these strings, called a population. This population is evolved in a manner that leads it toward the objective of the search. This requires that a measure of the fitness of each string, corresponding to a model in the GFA, be available

Following this, three operations are performed iteratively in succession: selection, crossover, and mutation. Newly added members are scored according to a fitness criterion. In the GFA, the scoring criteria for models are all related to the quality of the regression fit to the data. The selection probabilities must be re-evaluated each time a new member is added to the population.

Stability and convergence In common with other iterative minimization algorithms, there are issues with the stability and convergence of the GFA algorithm. An indication of the stability of the GFA algorithm can be obtained by generating a plot showing the evolution of variable usage withtime. Such a plot shows the number of occurrences of each variable in the population for each generation of the evolution. For practical reasons, to reduce the amount of data that would be collected, such a plot is generated only for those variables that occur most commonly in the final population and the data are not normally collected for every generation. The GFA algorithm is assumed to have converged when no improvement is seen in the score of the population over a significant length of time, either that of the best model in each population or the average of all the models in each population. When this criterion has been satisfied, no further generations are calculated.

Advantages of GFA

The GFA algorithm approach has several important advantages over other techniques [28]:

1. It builds multiple models rather than a single model.
2. It automatically selects which features are to be used in the models.
3. It is better at discovering combinations of features that take advantage of correlations between multiple features.
4. It incorporates Friedman's lack-of-fit (LOF) (Friedman J.H. 1991)error measure, which estimates the most appropriate number of features, resists overfitting, and allows control over the smoothness of fit.
5. It can use a large variety of equation term types in construction of its models, e.g., splines, step functions, high order polynomials.
6. It provides, through study of the evolving models, additional information not available from standard regression analysis, such as the preferred model length and useful partitions of the data set. The procedure continues for a user-specified number of generations, unless convergence occurs in the interim. Convergence is triggered by lack of progress in the highest and average scores of the population.

III. Result And Discussion

QSAR addresses two fundamentally important questions in scientific research; what structural and electronic properties of a molecule determine its activity and what can be altered to improve this activity? Computational tools allow researchers to identify chemicals with optimal physico-chemical properties in silico, before expensive experimentation. This saves both time and money, and allows the discarding of inferior candidates well in advance of reaching the laboratory (Khaled, K. F. 2011).

Table iii: List of descriptors used in this study.

Descriptors	Type	Significance
Molecular Weight	Structural	Used as the descriptors in systems such as transport studies where diffusion is the mode of operative.
Conolly Accessible Area	Geometrical	It combines shape and electronic information to characterize molecules.
cLogp	Thermodynamic	Octanol/Water partition coefficient
Molar Refractivity	Thermodynamic	It dependent on the spatial array of the aromatic ring in the synthesized compounds also necessary to study the interaction of the ligand with the receptor.
Molar Volume	Geometrical	It describes the volume inside van der waals area of the molecule
Polarizability	Quantum-chemical	It provide precise quantitative descriptions of molecular structure and their chemical properties.

Table iv: Experimental pMICand GFA PredictedpMIC for test set.

Compound	Observed	Predicted	Residual
4	3.21	3.44	-0.23
5	3.48	3.44	0.04
11	3.12	3.07	0.05
16	3.55	3.63	-0.08
18	3.17	3.12	0.05

3.1 QSAR Study

To investigate the Observed data, the distribution of the data must be first investigated. Most regression algorithm relies on the data that is being normally investigated, in case the data are not normally distributed, we should consider applying a numerical transformation to achieve a normal distribution.Observed data in Table ishow acceptable normal distribution, so no need to perform a numerical transformation. Table vshows a univariate analysis for the inhibition data. Table vcontains several statistical measures that describe the corrosion inhibition data. The most important parameters in Table vare the skewness and kurtosis. Skewness is the third moment of the distribution, which indicates the symmetry of the distribution.

Table v: Univariate analysis of the Observed data.

		IC_{50}
1	Number of sample points	15
2	Range	1.55
3	Maximum	4.72
4	Minimum	3.17
5	Mean	3.99
6	Median	4.09
7	Variance	0.16
8	Standard deviation	0.41
9	Mean absolute deviation	0.33
10	Skewness	-0.12
11	Kurtosis	-0.77

Constructing QSAR model is a process that takes a set of inputs and provides a set of outputs. For example, an energy minimization is a model which takes a structure as input and provides an optimised structure as output. At this point in a typical QSAR study, calculation of descriptors occurs. These are models which take a single structure as an input and provide a single number or group of closely related numbers as outputs. Table vi.Shows the experimental pMIC and the predicted pMIC using the GFA approach of the training set. This shows how the GFA method predicted the pMIC.

Table vi: Experimental PMIC and GFA Predicted PMIC for the training set.

Compounds	Actual values for A : IC50	Equation 1: predicted values	Equation 1: residual values
1	4.53	4.56	-0.031
2	4.16	4.097	0.0626
3	3.52	3.608	-0.088
6	4.15	4.139	-0.061
7	3.80	3.7762	-0.1076
8	3.83	3.85	-0.1142
9	4.13	4.1946	0.0518
10	3.83	3.855	-0.025
12	4.47	4.485	0.098
13	4.72	4.720	0.0462
14	3.83	3.789	0.0737
15	4.14	4.0995	0.0123
17	3.51	3.534	-0.0329
19	4.09	3.998	0.092
20	3.51	3.535	-0.024

Table vii:shows the genetic function approximation analysiswhich gives a summary of the input parameters used for the calculation. Also, it reports whether the GFA algorithm converged in the specified number of generations. Convergence is achieved when there has been no improvement in the scoring function for a number of generations. It can be seen from Table vii:that the accuracy of the model, indicated by the R^2 value, is reasonably high therefore the predictive power of the model, as indicated by the adjusted R^2 and cross validated R^2 values, is also, high, even though the regression is significant according to F-test. In Table vii, the Friedman's lack-of-fit (LOF) score (Friedman, J.H 1991), which evaluates the QSAR model by considering the number of descriptors as well as the quality of fitness, is chosen: the lower the LOF, the less likely it is that GFA model will fit the data. The significant regression is given by F-test, and the higher the value, the better the model.

Fig. ia-bshows a relation between the predicted values using the equation in Table ivand the experimental data in Table i. Also, the distribution of the residual values against the Observed activity values. A residual can be defined as the difference between the predicted value in the generated model and the measured value for corrosion inhibition. To test the constructed QSAR model, potential outliers have been identified in Fig. iia-b. An outlier can be defined as a data point whose residual value is not within two standard deviations of the mean of the residual values. Although the number of outliers can vary depending on the quality of the dataset (e.g., incorrect measurements of physical properties or errors in molecular structures will reduce the data set quality), it still a good test of QSAR model is to identify potential outliers.

Fig. iia-b contains two charts. One contains the residual values plotted against the Observed pMICand the other displays the residual values plotted against Table iiiraw number. Each chart contains a dotted line that indicates the critical threshold of two standard deviations beyond which a value may be considered to an outlier. Inspection of Fig. iishows that there is no points appeared outside the dotted lines which make the QSAR model acceptable.

In fig. iii:, the Y-axis represents the different molecular descriptors used in this study as shown on the right side of the graph. On the other hand, the X-axis represents the number of thegenerations we could generate for each of these molecular descriptors.

According to fig iii, at each step, the GFA uses the current population to create the children that makes up the next generation. The algorithm selects a group of individuals in the current population , called parents , who contribute their genes –the entries of their vectors – to their children . The algorithm usually selects individuals that have better fitness values as parents . User can specify the function that the algorithm uses to select the parents. The GFA creates three types of children for the next generation: Elite Children , Crossover Children and Mutation Children. In our QSAR study ,the algorithm stops when the number of generations reaches the value of 500 Generations.

Table vii:Validation of the genetic function approximation

		Equation 1
1	Friedman LOF	0.0231
2	R-squared	0.9864
3	Adjusted R-squared	0.9762
4	Cross validated R-squared	0.9284
5	Significant Regression	Yes
6	Significance-of-regression F-value	96.6172
7	Critical SOR F-value (95%)	3.5968
8	Lack-of-fit points	8
9	Min expt. error for non-significant LOF (95%)	0.0457

Table viii: best model:

Y= -0.01015a-0.007566b-0.98076c-55.9669d+0.02751e+141.3109f+8.76211

a =molecular weight,b=conollyaccessibl area,c=cLogp,d=molar refractivity,e=molar volume,f=polarizability and Y:pMIC

Fig. ia -b. Plot of predicted activity and residuals versus Observed activity

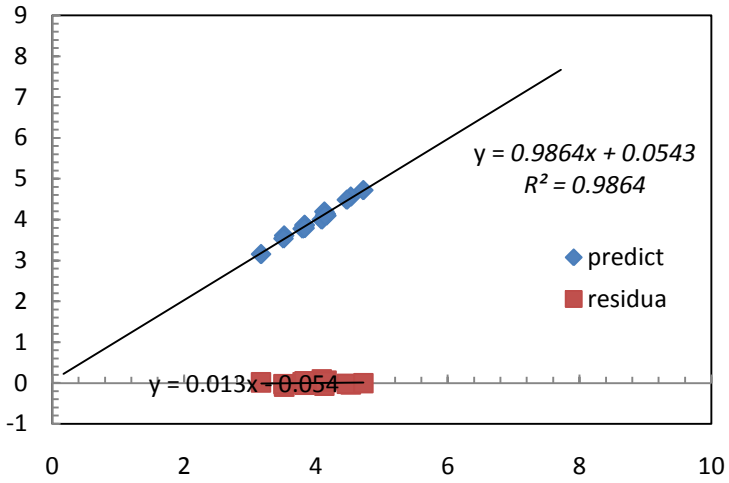


Fig iia-b :OUTLIER ANALYSIS

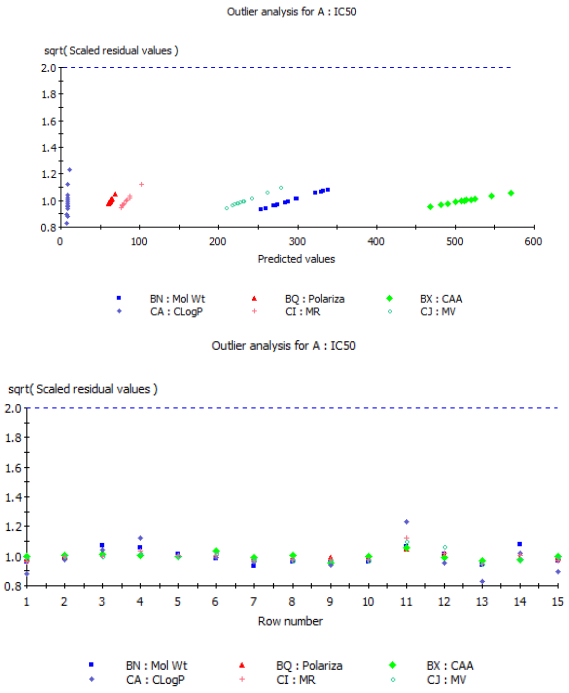
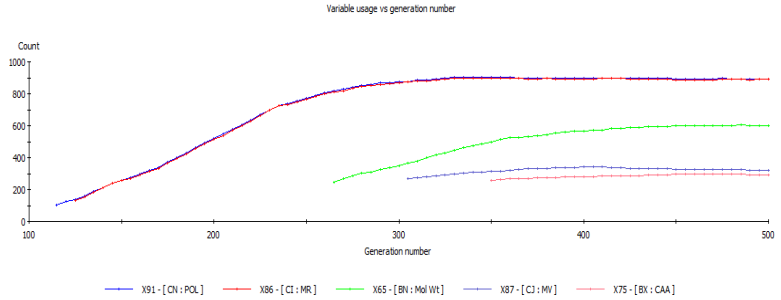


Fig iii: The graph of the variable usage against generation number



IV. Conclusion

A genetic function approximation method was used to run the regression analysis and establish correlation's between different types of descriptors and measured Chemotheraphy activities of Ketones derivatives.

References

[1]. Ajmani, S., Jadhav, K., &Kulkarni, S. A. (2006). Three-dimensional QSAR using the k- nearest neighbor method and its interpretation.Journal of chemical information and modeling, 46(1), 24-31.
[2]. Allen, B.R.and Stanley, S.Y. (2008). An Introduction to QSAR Methodology. Network Science Corporation 1116 Miller Mountain Road Saluda and GlaxoWellcome Research Five Moore Drive Research Triangle Park, New Jersey.
[3]. Delley, B. (1990). An all-electron numerical method for solving the local density functional for polyatomic molecules.The Journal of chemical physics, 92(1), 508-517.
[4]. Delley, B. (2000). From molecules to solids with the DMol3 approach. The Journal of chemical physics, 113(18), 7756-7764.
[5]. Devillers, J. (1996).PRINCIPLES OF QSAR AND DRUG DESIGN:GENETIC ALGORITHMS INMOLECULAR MODELING. Academic Press Limited, 24- 28 Oval Road, London, pp:20- 21
[6]. Dorsett, H and White,A.(2000). Overview of Molecular Modelling and AbinitioMolecular Orbital Methods Suitable for Use with Energetic Materials. DSTO Aeronautical and Maritime Research Laboratory,Salisbury South Australia,Australia.
[7]. Eldred, D. V., Weikel, C. L., Jurs, P. C., & Kaiser, K. L. (1999). Prediction of fathead minnow acute toxicity of organic compounds from molecular structure. chemical Research in Toxicology, 12(7), 670-678.
[8]. Elvis, A. M and Rakesh, R.S.(2012). Drug Designing, Discovery and Development Techniques Available from: http://www.intechopen.com/books/promising- pharmaceuticals/drug-designing- discovery-and-development techniques.
[9]. Eriksson, L., & Johansson, E. (1996). Multivariate design and modeling in QSAR. Chemometrics and Intelligent Laboratory Systems, 34(1), 1-19.
[10]. Franke, R.(1984). Theoritical Drug Design Methods.Elseiver, Amsterdam.pp. (256).
[11]. Friedman, J. H. (1991). Multivariate adaptive regression splines.The annals of statistics, 1-67
[12]. Fujita, T.(1990). The extral thermodynamic approach to drug design Comprehensive Medicinal Chemistry.Pergamon, Oxford. (4), 497-561.
[13]. Gareth,T. (2003). Fundamentals of Medicinal Chemistry. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England.
[14]. Gonzalez. C.E.,VenzonD.,Lee S.,Meller B.U.,PizzoP.A.,Walsh T.J.,(1996).Risk factors for fungemia in children infected with HIV:a case control study.Clinical Infected Dissertation. (23): 515.
[15]. Graybill, J.R.(1992).Future directions of antifungal Chemotheraphy.Clinical Infected Dissertation. 14 (1):170.
[16]. Haigren T.A.,(1996). Merck molecular force field-11.MMFF94 van der Walls and electrostatic parameters for intermolecular interactions, Journal of Computational Chemistry, (17), 520 – 552.
[17]. Haigren T.A.,(1996). Merck molecular force field-111.Molecular Geometries and vibrational frequencies for MMFF94, Journal of Computational Chemistry, (17): 553 – 586.
[18]. Haigren T.A.,(1996). Merck molecular force field-1V. Conformational energies and geometries for MMFF94, Journal of Computational Chemistry, (17): 587 – 615.
[19]. Halgren T.A.,(1996). Merck molecular force field-1. Basis, form, scope, parametrization and performance of MMFF94, Journal of Computational Chemistry, (17): 490 – 519.
[20]. Hammett, L. P. (1935). Some Relations between Reaction Rates and Equilibrium Constants. Chemical Reviews, 17(1), 125-136.
[21]. Hansch, C.,Rockwell, S.D.,Joe, P.Y.,Leo. A.,Steller, E.,(1997). Substituents constants for correlation analysis.Journal of Medicinal Chemistry. (2). 304.
[22]. Holland, J. H. (1975). Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. U Michigan Press.
[23]. Homaifar, A., Qi, C. X., & Lai, S. H. (1994).Constrained optimization via genetic algorithms Simulation, 62(4), 242-253.
[24]. Jitender, K Malik, HimeshSoni, Singhai A K and Harish Pandey. (2013). QSAR Application in Drug Design.International Journal of Pharmaceutical Research and Allied Sciences.2,(1): 1-13.
[25]. Joanna, J and Nina, N,(2004). Review of methods for assessing the applicabilty domains of sars and qsars. The European Commission - Joint Research Centre Institute for Health & Consumer Protection Italy.
[26]. Kamalakaran, A.S., et al (2009). Studies of N-aryl Derivative Activity Towards Alzheimer's Disease. Journal of molecules, 14: 1448-1455
[27]. Katritzky, A. et al.,(1995). QSAR: the correlation and quantitative prediction of chemical and physical properties from structure .Chemical Society Review, 24: 279 – 287.
[28]. Khaled, K. F. (2011). Modeling corrosion inhibition of iron in acid medium by genetic function approximation method: A QSAR model.Corrosion Science, 53(11), 3457-3465.
[29]. Khaled, K.F. and Abdel-Shafi, N.S.(2011).QUANTITATIVE STRUCTURAL AND ACTIVITY RELATIONSHIP MODELING STUDY OF CORROSION INHIBITORS: GENETIC FUNCTION APPROXIMATION AND MOLECULAR DYNAMICS SIMULATION METHODS. International Journal of Electro chemical Science. 6: 4077- 4094
[30]. Khaled, K.F., and El-Sherik, A.M.(2013).Using molecular Dynamics Simulations and Genetic Function Approximation to model corrosion Inhibition of iron in chloride solutions. International journal of Electrochemical Science, 8: 10022- 10043.
[31]. Kubiny, H. (1994). Variable selection in QSAR studies. I. An evolutionary algorithm. Quantitative Structure-Activity Relationships, 13(3), 285-294.
[32]. Kubinyi, H. (1994). Variable selection in QSAR studies. II. A highly efficient combination of systematic search and evolution. Quantitative Structure-Activity Relationships, 13(4), 393-401.
[33]. Kubinyi, H. (1997). QSAR and 3D QSAR in drug design Part 1: methodology. Drug discovery today, 2(11), 457-467.
[34]. Leardi, R. (2001). Genetic algorithms in chemometrics and chemistry: a review. Journal of chemometrics, 15(7), 559-569.
[35]. Leardi, R., & Gonzalez, A. L. (1998). Genetic algorithms applied to feature selection in PLS regression: how and when to use them. Chemometrics and intelligent laboratory systems, 41(2), 195-207.
[36]. Murray P.R.,Rosenthal K.S., and Pfaller M.A.,(2006).MicrobiologiaMedica.Elsevier. Brazil. Pp:693 702.
[37]. Nantasenamat, C., Isarankura-Na-Ayudhya, C., Naenna, T., &Prachayasittikul, V. (2009).A practical overview of quantitative structure-activity relationship.EXCLI J, 8(7).

[38]. Niculescu, S. P. (2003). Artificial neural networks and genetic algorithms in QSAR.Journal of Molecular Structure: THEOCHEM, 622(1), 71-83.

[39]. Odds, F.C.,(1993). Chemotherapeutic agents.Journal of AntimicrobialChemotheraphy., (31): 463- 471.

[40]. Ravichandran V., Mourya V.K., Agrawal R.K.,(2007). QSAR prediction of HIV-1 reverse transcriptase inhibitory activity of benzoxazinone derivatives, internet Electron.Journal of Molecula. Descriptors., (6): 363-374.

[41]. Ray, S., &Pratim Roy, P. (2012). A QSAR study of biphenyl analogues of 2-nitroimidazo-[2, 1- b][1, 3]-oxazines as antitubercular agents using genetic function approximation. Medicinal Chemistry, 8(4), 717-726.

[42]. Rogers, D. (1991). G/Splines: A hybrid of Friedman's multivariate adaptive regression splines (mars) algorithm with Holland's genetic algorithm.

[43]. Rogers, D. (1996). Approximation with Comparison to Evolutionary Techniques.Genetic algorithms in molecular modeling, 87

[44]. Rogers, D., &Hopfinger, A. J. (1994).Application of genetic function approximation to quantitative structure activity relationships and quantitative structure-property relationships. Journal of Chemical Information and Computer Sciences, 34(4), 854- 866

[45]. Rogers, J. L., McCulley, C. M., &Bloebaum, C. L. (1996).Integrating a genetic algorithm into a knowledge-based system for ordering complex design processes (pp. 119- 133). Springer Netherlands.

[46]. Taft, R. W., & Newman, M. S. (1956). Steric effects in organic chemistry.Wiley, New York, NY, 597.

[47]. Thomsen M and Carlsen,L. (2001).Evaluation of Empirical Contra non-empirical Descriptors. SARS and QSAR in Environmental Chemistry , in press.

[48]. Tomasz, P., Jerzy, L. and Mark, T.D.C. (2010). Challenges and advances in Computational Chemistry and physics. Recent Advances in QSAR Studies.Methods and Applications.Springer Dordrecht Heidelberg London New York.8:3- 60.

[49]. Tong W., Hong H., Xie Q., Shi L., Fang H. and Perkins R. (2005). Current Computer – Aided Drug Desig (7) :195 – 205.

[50]. Topliss, J. G., & Costello, R. J. (1972).Chance correlations in structure-activity studies using multiple regression analysis.Journal of Medicinal Chemistry, 15(10), 1066-1068.

[51]. Tropsha A., Gramatica p., Gombar V.K.,(2003). The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSAR models, QSAR Combinatory. Science.(22): 69-77.

[52]. Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C. P., &Agrawal, R. K. (2011). Validation of QSAR models-strategies and importance.International Journal of Drug Design & Discovery, 3, 511-519

[53]. Vijay, M.G. and Vithal, M.K. (2000). Understanding the Antifungal Activity OfTerbinafine Analogues Using Quantitative Structural – Activity Relationship (QSAR) Models. BIOORGANIC AND MEDICINAL CHEMISTRY.8: 2487- 2499.

[54]. Viral, P., et al (2013).QSAR Study of Series of 4-amino-3,5-di(substituted)thiazol-2(3H)- thiore as Antitubercular agent. PHARMAGENE Research Article, 1:2.

[55]. Walsh T. J.,Georgopa N.P.,(1992). Emerging Targets in Antibacterial and Antifungal Chemotherapy. Chapman & Hall, New York London, pp. 349-373.

[56]. Weinstein, J.N. (2001).Quantitative Structure-Antitumor Activity Relationships of Camptotheci Analogues.Journal of medicinal Chemistry, 44: 3254-3263.

[57]. William, J.W. and Thomas, E.R.(2003). Burger's Medicinal Chemistry and Drug Discovery.I :Donald,J.A.(Ed). Antifungal Agents. John Wiley & Sons, Inc,Califonia. Pp.881 – 900.

[58]. Williams D.A.,Lemke T.L.F. (2002). Principles of Medicinal Chemistry(Ed). Lippincott Williams Wilkins,Baltimore. Pp 81.