Gazi University Journal of Science
GU J Sci
24(2):275-282 (2011)

GUJS
www.gujs.org

Defining the Possible Molecular Structure of the Drug to Be Penetrated through Skin Layers Using Genetic Algorithm

Aysun COSKUN[1,♠], Nursal ARICI[2]

[1] *Department of Computer Education, Gazi University Faculty of Industrial Arts Education, Golbası Campus, Ankara, Turkey*

[2] *Department of Electronics and Computer Education, Gazi University Faculty of Technical Education, 06500 Besevler, Ankara, Turkey*

Received: 19.03.2010 Revised: 17.10.2010 Accepted: 16.01.2011

ABSTRACT
The scientists in the past when they were trying to solve the problem of relationship between parameters by using trial and errors method, due to increase of the number of parameters problem of deadlock or non-evaluation of the solution has occurred. They tried to find new techniques in order to solve the problems of parameters and positive results were taken from genetic algorithms of artificial intelligence. Genetic algorithm which has an optimization technique has been identified as a non-traditional type of research techniques. The implementation of genetic algorithm have been realized in the identification of quotients of penetration of chemicals through skin and Delphi 7.0 and MOLGA (MOLecule and Genetic Algorithm) program was set up in this work. Genetic algorithm method was used in solving the problems of multi parameters optimization problems. 11 parameters were taken as the basis of the molecular structure of the chemicals, and random method has chosen in the optimization of the genetic algorithm. As the coherence function for the implementation of genetic algorithm regression equation of penetration through skin quotients based on the parameters have been used. It was seen that, when the quotients identified according to the MOLGA program results and molecular structure were compared there was very little margin of error. At the same time, molecular structure of the chemicals within MOLGA program has to be changed and developed.

Key words: *Quotients of penetration, Genetic algorithms, Molga*

1. INTRODUCTION

There have been considerable developments in our knowledge about the mechanisms and factors affecting parameters of skin permeation. This has been possible because of the latest developments of experimental techniques and increased computational power and technology. The advanced technology and available software have provided an opportunity to determine a relationship between permeability and molecular properties of the penetrant [1]. Available computer programs for molecular modeling have been used to calculate some molecular properties of the drug molecules such as surface area, partial charges etc. Some new approaches have also been incorporated to the skin research including Principal Component Analysis (PCA) [2] and a biologically inspired computer algorithm designed to learn from data in a manner emulating the learning pattern in the brain which is called Artificial Neural Network (ANN) modeling. Using an ANN, Agatonovic-Kustrin et al. [3] developed a quantitative structure–permeability relationship of penetration across polydimethylsiloxane membranes, which were expected to be the model of skin permeation [4] [7]. A set of 254

♠Corresponding author, e-mail: aysunc@gazi.edu.tr

compounds and their maximum steady state flux was collected from the literature. Twelve of 42 molecular descriptors were selected for ANN modeling of maximum steady-state flux by the use of genetic algorithm, that include molecular shape and size, inter-molecular interactions, hydrogen bonding capacity of drugs, and conformational stability [8]. Degim et al. [9] analyzed skin permeability of 40 compounds by an ANN and compared its predictability with the multiple linear regression model obtained by Pugh et al. [2]. According to the linear model of Pugh et al., the partial charges of the penetrants, their molecular weight, and their calculated octanol–water partition coefficient (log P) were used as molecular descriptors. While the linear equation gave a regression coefficient of 0.672, the ANN produced log K values that correlated well with the experimental ones (r^2= 0.997). There is no study has been performed so far to define possible molecular structure of the drug to be penetrated through skin layers using genetic algorithm.

2. GENETIC ALGORITHM

Genetic algorithm was first determined by J. Holland and suggested as an investigation technique built on the basis of genetic sciences. It is a sampling of the method which is used by the biological system for adaptation to environmental conditions. It is present in the basis of evolution theory that all individuals adapt to environmental conditions by changing some characteristic properties in the period of time and new populations occur. Genetic algorithms work on the population constituted by individuals. Population constituted by individuals who are representing possible solutions of a problem. The possible solution is searched from the initial population and better solutions are seeked. Populations are changing by the time but have the same number of constituents [10]. Each solution gets a fitting parameter in the general structure of genetic algorithm. In nature, this fitting parameter is the degree of success of the biologic matter in the environment. If the fitting parameter is high, it means that the biological matter in question will be live and transfer its genetic material to the next generations with high possibility.

Genetic algorithm is an iterative process. The basis of genetic algorithm is to create an initial population randomly and to continue to constitute new generations until reach to the certain population which gives best solution. The steps for the solution of the problem by genetic algorithm are by Karaboga as follows [11]:

1st step: Create an initial population randomly
2nd step: Give fitting parameters to each solutions in initial population
3rd step: If the fitting parameter conforms the required criteria stop the program else apply the operators of genetic algorithm to create new population
3.1st step: Apply natural selection
3.2nd step: Cross wising
3.3rd step: Mutation
4th step: Give fitting parameters to each solutions in new population
5th step: Go to step 3.

The first step for achieving a successful application of genetic algorithm is coding of the independent parameters in the chromosome. This chromosome carries all information about the certain solution of the problem. There are number of different ways for coding. Using a binary system is the general method but, using integers is also possible for each gene on the chromosome.

Generally initial population is constituted randomly, but especially in limited optimization problems, random constituting may cause finding a solution which is not suitable therefore, the intuitional methods may be also useful for solving the problem. Using a sequential constitution is possible but it needs quite long time.

The performance of the chromosome on the function is evaluated by determining the acceptance criteria of the chromosome in the genetic algorithm.

Selection is a process where choosing the individuals for reproduction. It is a process where the possible candidates for solving the problem chosen. The number of youngsters that each individual will have and how the individuals will be selected for reproduction is determined in this process [12]. The roulette circle or tournament methods are used in the literature. The selection can be made by considering the value of the acceptance criteria of the chromosome. Higher values always accepted. Then the genetic operators applied.

The procedures of the selection do not always solve the problem. Selected individuals are kept in a reproduction pool and selection process stops when the population in the reproduction pool reaches to the original population. Selected individuals are ancestors of new generation, developed solutions are obtained by genetic operators at this step. There are two important operators works for creating new solutions namely mutations and cross wising. Two chromosomes coupled randomly to each other and genes on these chromosomes replaced. The aim of this is to get a new chromosome which has got good genes from other two.

It is the most important operation in genetic algorithms; mostly used types are cross wising from single point, double point, multipoint or uniform cross wising. The codes of each individual are differentiating by certain probability to get back lost good properties of the population after several cross wising.

Mutation makes changes of the genes. This operation is secondarily important operation in genetic algorithm as it is seen in nature.

3. METHOD APPLICATION OF GENETIC ALGORITHM AND A DEVELOPED COMPUTER PROGRAM

The genetic algorithm was modeled by Delphi 7.0. Genetic algorithm needs a function to examine whether the chromosome conforms to criteria or not. The function was taken from the literature. The relationship between 11 structural parameters of the 91 compound and skin permeability coefficients (log k_p) has been given as follows [13]:

$$Logk_p=-2,709+0,233C-0,467Ar+0,446Hal\ 1,177N(Amine)1,445N(toO)+0,503O(toN)-0,657NonAr-1,136Ste-0,452OH-0,316O-0,352Amide \quad (1)$$

C : Number of carbon atoms on the molecule
Ar : Number of aromatic ring on the molecule
Hal : Number of halide groups on the molecule
N(Amine): Number of amine groups on the molecule
N(toO) : Number of nitrogen atoms attached to the oxygen atoms on the molecule
O(toN) : Number of oxygen atoms attached to the nitrogen atoms on the molecule
NonAr : Number of non-aromatic ring on the molecule
Ste : Number of steroid rings on the molecule
OH : Number of hydroxyl groups on the molecule
O : Number of oxygen atoms on the molecule
Amide : Number of amide groups on the molecule

A computer program was developed and called MOLGA to define the possible molecular structure of the drug to be penetrated through skin layers using genetic algorithm according to given criteria. Initial population was created by 91 compounds and their properties. These structural parameters were taken from the literature [13]. Structure of the virtual compound is defined by genetic algorithm rules randomly. When MOLGA runs, program asks user to enter permeability range for a virtual compound which though to be penetrated through human skin. MOLGA then produces virtual chemical structures by giving random values to genes according to genetic algorithm. MOLGA then calculates the log k_p using Equation-1 and examines whether its is in the given permeability range or not. If it is in given permeability range program stops. If it is not, program continues to run and searches what is the possible structure which conforms to the given criteria. MOLGA works under following assumptions and rules:

- MOLGA considers Log k_p values of the penetrant,
- 11 structural parameters used and each parameter is represented on one gene at the chromosome (Table-1).

Table 1. The structure of the chromosome used by genetic algorithm

1st gene	2nd gene	3rd gene	4th gene	5th gene	6th gene	7th gene	8th gene	9th gene	10th gene	11th gene
19	0	0	0	0	0	4	1	2	2	0

MOLGA gives numbers randomly to the genes for coding of the chromosome (Table -2).

Table 2. Coding of the chromosome

1st gene	2nd gene	3rd gene	4th gene	5th gene	6th gene	7th gene	8th gene	9th gene	10th gene	11th gene
C(Not=O)	Aromatic	Halide	N(Amine)	NtoO	OtoN	NonAr	Steroid	OH	O	Amide

- MOLGA uses the roulette circle for selection,
- Cross wising operator performs its duty randomly from single and two points at the chromosome.

Some part of the flow chart of the Molga program was given in Fig-1.

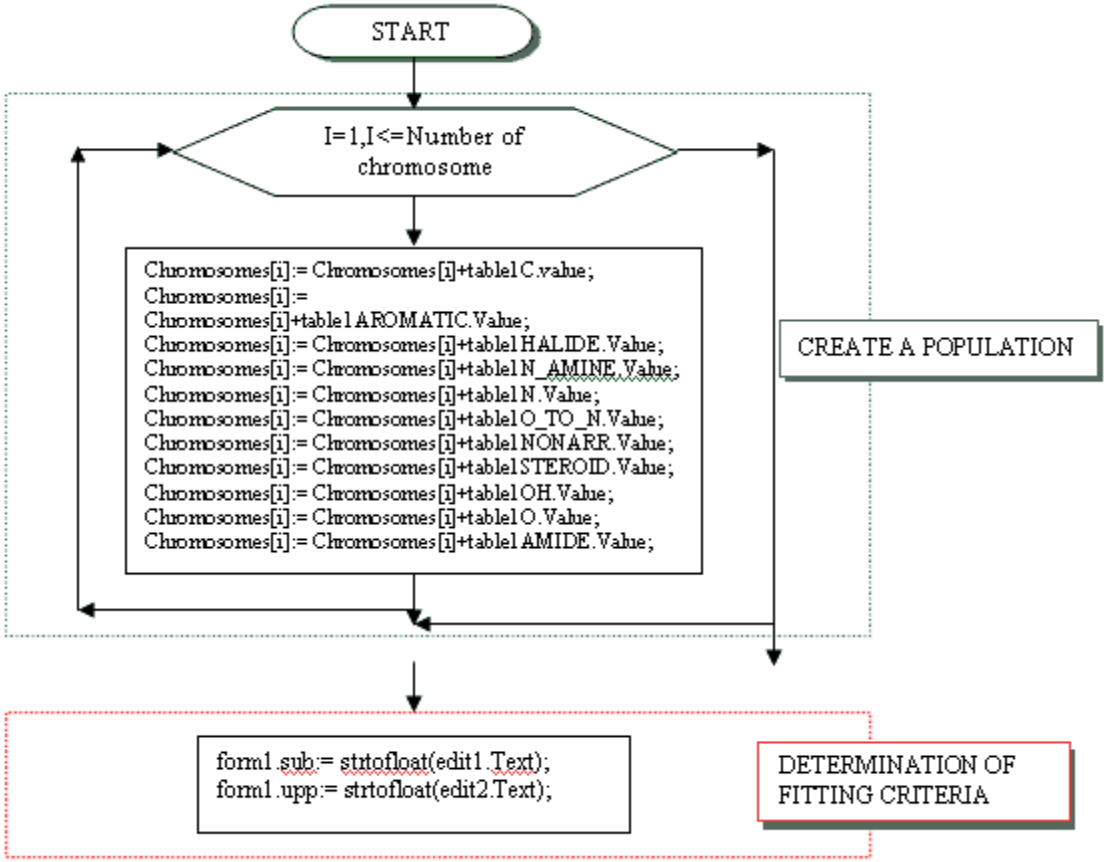

Figure 1a. Create a population-determination of fitting criteria

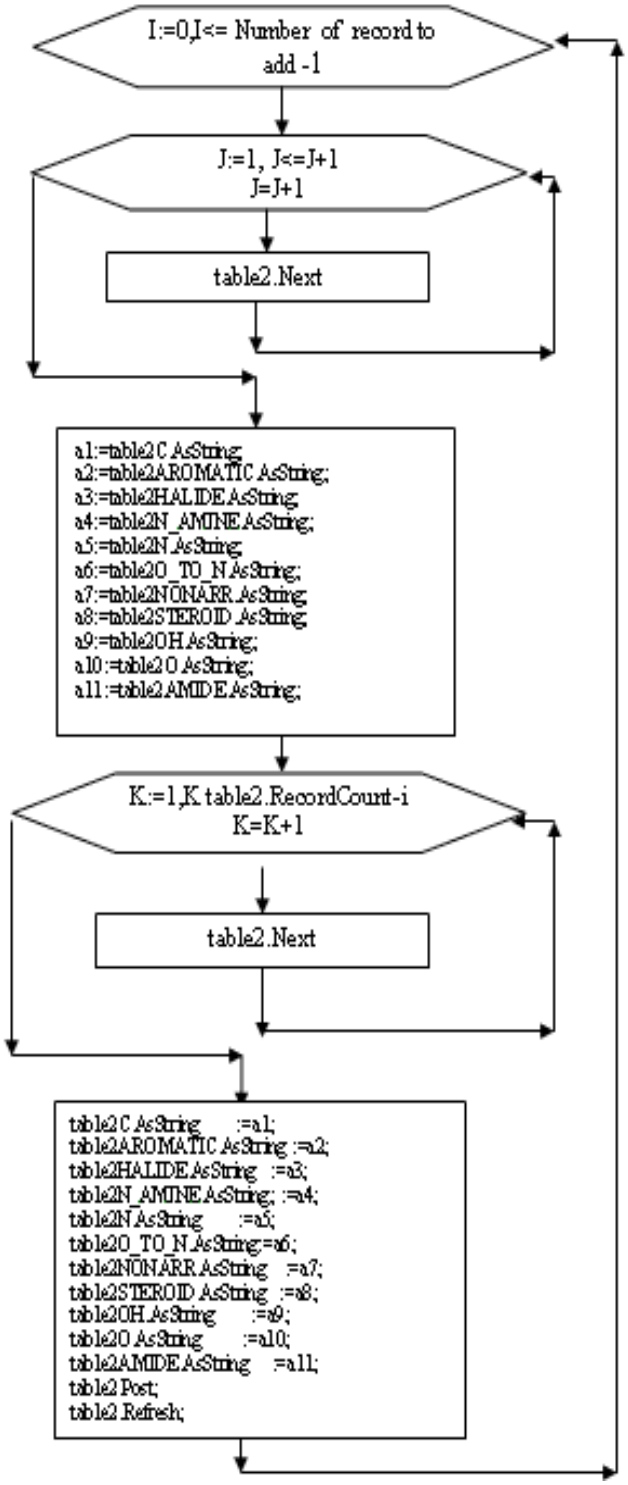

Figure-1b. The best individuals instead of bad ones

4. MOLGA MENUS

Molga menus work as follows.

Defining the criteria
MOLGA asks user to enter a permeability range for a virtual chemical which is going to be penetrated through human skin. MOLGA expect an expert to enter these values for instance whoever searches a cosmeceutics type of chemical which is wanted to be penetrate through human skin well and entirely. User can run the program and understand what type of groups on the molecule can increase or decrease the permeability of the compound.

Data
Some of the data from the literature (for 91 chemicals) have already been entered to the MOLGA. MOLGA allows user to alter, change or remove some or all of the data but, if any of the data changed or altered the function or equation will alter, therefore the function or equation should be changed. It may not be changed by the user but accepted equation must be bear in the mind. These entered data used by MOLGA to create initial population.

Setting up the algorithm
This menu is to set up genetic algorithm operators. Number of chromosomes at the initial population, number of genes on the chromosome and operators for the selection of ancestors can be defined using this menu. Operators for the selection are roulette circle and choosing good individuals instead of bad ones. In the Choosing good individuals instead of bad ones option user can define how many bad individuals will be discarded. For example, if the user enters the value of 10, 10 bad individuals will be discarded and 10 best individuals will be used for next generation. Cross wising operator in the program performs cross wising from single point or two points on the chromosome. User can define which two genes will be sued for cross wising by the program.

Function or equation
MOLGA allows user to enter the function or equation. Default function has already been entered to the MOLGA (Equation-1), if user likes it can be changed. This function will be used to calculate log k_p values of the compounds.

Run window
This window is a interactive window between user and the MOLGA. Possible solutions, suitable chromosomes and results will appear in this window. Number of iterations, given criteria can also be seen. MOLGA finally gives possible structures and user can understand what functional group affects the permeability.

History
MOLGA also able to store last performance results if user wanted.

5. RESULTS AND DISCUSSION

Molga was first run and permeability range was entered as -2.3 to -2.2. The best individuals instead of bad ones and random single point hybridization options were chosen. Molga finally produced some possible structures for these setting and one possible structure was given as follows after 10th iterations:

Calculated Log permeability values: -2.23
Number of carbon atoms: 6,
Number of OH groups: 1,
Number of aromatic ring: 1.

This structure can be phenol (Figure-3). The Log kp values of the phenol was reported in the literature with the value of -2.09 in the literature[13]. This value from Molga calculation was found to be quite close to original value and proposed structure by Molga was found to be realistic.

Molga was then run again and permeability range was entered as -2.5 to -2.4. Rulet circle and random single point hybridization options were chosen. Molga finally produced some possible structures for these setting and one of them was given as follows after 1th iteration:

Calculated Log permeability values: -2.462
Number of carbon atoms: 3,
Number of OH groups: 1,

This structure can be methanol (Figure-3). The Log kp values of the methanol was reported in the literature with the value of -3.30 in the literature[13]. This value from Molga calculation was found to be quite close to original value and proposed structure by Molga was found to be realistic again.

Molga was run again and permeability range was entered as -1.6 to -1.5. The best individuals instead of bad ones and random single point hybridization options were chosen. Molga finally produced some possible structures for these setting and one of them was given as follows after 1th iterations:

Calculated Log permeability values: -1.551
Number of carbon atoms: 7,
Number of OH groups: 1,
Number of aromatic ring: 1.

This structure can be *p*-cresol (Figure-3). The Log kp values of the *p*-cresol was reported in the literature with the value of -1.75 in the literature 13. This structure can be *benzoic acid* (Figure-1). The Log kp values of the benzoic acid was reported in the literature with the value of -1.600 in the literature 14, 15. This value from Molga calculation was found to be quite close to original values and proposed structures by Molga were found to be realistic again.

Molga was run again and permeability range was entered as -2.3 to -2.2. Rulet circle and random single point hybridization options were chosen. Molga finally produced some possible structures for these setting and one of them was given as follows after 1th iteration:

Calculated Log permeability values: -2.229
Number of carbon atoms: 4,
Number of OH groups: 1,

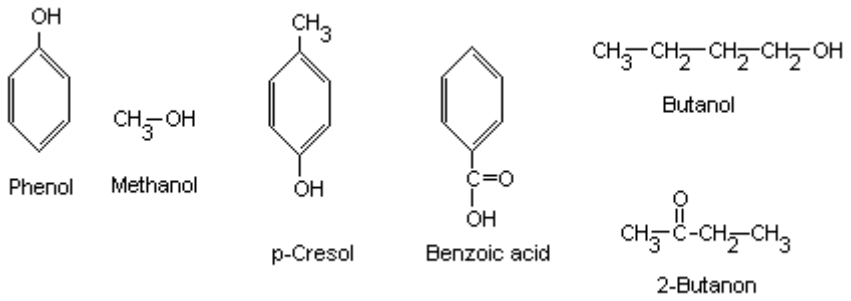

Figure 3: Chemical structures of phenol, methanol, p-cresol, benzoic acid, butanol and 2-butanon

This structure can be butanol (Figure-3). The Log kp values of the butanol was reported in the literature with the value of -2.60 in the literature 13. This value from Molga calculation was found to be quite close to original value and proposed structure by Molga was found to be realistic again.

Molga was run again and permeability range was entered as -2.1 to -2.0. Rulet circle and random single point hybridization options were chosen. Molga finally produced some possible structures for these setting and one of them was given as follows after 1th iterations:

Calculated Log permeability values: -2.093
Number of carbon atoms: 4,
Number of O atoms: 1,

This structure can be 2-butanon (Figure-3). The Log kp values of the 2-butanon was reported in the literature with the value of -2.60 in the literature 13. This value from Molga calculation was found to be quite close to original value and proposed structure by Molga was found to be realistic again.

6. CONCLUSION

As a conclusion, Molga can produce realistic structures when desired permeability ranges are entered. Molga can be very useful and Molga can be used for various aims from searching well penetrated structures to looking for non penetrating structures such as insecticides or pesticides which are not penetrate through human skin well considering farmers health especially at the product development stage. Molga can be used by experienced users who are familiar for the chemical structures. This publication proposed quite different idea which has not been published and it may be developed more and program can produce possible chemical structures in detail.

REFERENCES

[1] Degim, I. T. "Understanding skin penetration: Computer aided modeling and data interpretation", ***Current Computer-Aided Drug Design***, 1, 11-20, (2005).

[2] Pugh, W. J., Degim, I. T. and Hadgraft, J., "Epidermal permeability–penetrant structure relationships: 4, QSAR of permeant diffusion across human stratum corneum in terms of molecular weight, H-bonding and electronic charge", ***Int. J. Pharm.***, 197, 203–211, (2000).

[3] Kustrin A. R. Beresford and Yusof, A. P. "ANN modeling of the penetration across a polydimethylsiloxane membrane from theoretically derived molecular descriptors", ***J. Pharm. Biomed. Anal.,*** 26, 241–254, (2001).

[4] Addicks, W. J., Flynn, G.L., Weiner, N. and Chiang, C. M. "Drug transport from thin applications of topical dosage forms: development of methodology", ***Pharm. Res***. 5, 377– 382, (1988).

[5] Jetzer, W. E., Huq, A. S., Ho, N. F, Flynn, G. L., Duraiswamy, N. and Condie L. Jr. "Permeation of mouse skin and silicone rubber membranes by phenols: relationship to in vitro partitioning", ***J. Pharm. Sci.*** 75, 1098–1103, (1986).

[6] Chen, Y., Vayuhasuwan, P. and Matheson, L. E. "Prediction of flux through polydimethylsiloxane membranes using atomic charge calculations:

application to an extended data set", ***Int. J. Pharm***. 137, 149–158, (1996).

[7] Chen, Y., Yang, W. L. and Matheson L.E. "Prediction of flux through polydimethylsiloxane membranes using atomic charge calculations", ***Int. J. Pharm.*** 94 81–88, (1993).

[8] Cronin, M. T., Dearden, J. C. Gupta, R. and Moss, G. P. "An investigation of the mechanism of uxacrosspolydi-methylsiloxane membranes by use of quantitative structure-permeability relationships", ***J. Pharm. Pharmacol.*** 50, 143–152, (1998).

[9] Degim T., Hadgraft J., Ilbasmis S. and Ozkan Y. "Prediction of skin penetration using artificial neural network (ANN) modeling", ***J. Pharm. Sci.*** 92, 656–664, (2003).

[10] Murty, K. G. Operations Research Deterministic Optimal Models, ***Prentice Hall***, N.J., , p 581, (1995).

[11] Karaboga, D., Yapay Zeka Optimizasyon Algoritmaları, ***Atlas Yayınevi***, Yayın No: 38, Istanbul, 79p, (2004).

[12] Mitchell, M. An Introduction to Genetic Algorithms, ***MIT Press***, Massachusetts, p 205, (1996).

[13] Pugh W. J. and Hadgraft J., "Ab initio prediction of human skin permeability coefficients", ***Int. J. Pharm.***, 103: 163-178, (1994).

[14] Degim I. Pugh, W.J. Hadgraft, J. "Skin permeability: Anomalous results", ***Int. J. Pharm***. 170:129 – 133, (1998).