Rational Drug Design using Genetic Algorithm : Case of Malaria Disease

[1] Hassen Mohammed Alsafi, [2] Imad Fakhri Al-Shaikhli

[1, 2] Faculties of Information & Communication Technology, Department of Computer Science

Kuala Lumpur, Malaysia

[1] abosafi87@gmail.com, [2] imadf@iium.edu.my

ABSTRACT

With the rapid development in the amount of molecular biological structures, computational molecular docking (CMD) approaches become one of the crucial tools in rational drug design (RDD). Currently, number of researchers are working in this filed to overcome the recent issues of docking by using genetic algorithm approach. Moreover, Genetic Algorithm facilities the researchers and scientists in molecular docking experiments. Since conducting the experiment in the laboratory considered as time consuming and costly, the scientists determined to use the computational techniques to simulate their experiments. In this paper, auto dock 4.2, well known docking simulation has been used to perform the experiment in specific disease called malaria. The genetic algorithm (GA) approach in the autodock4.2 has been used to search for the potential candidate drug in the twenty drugs. It shows the great impacts in the results obtained from the CMD simulation. In the experiment, we used falcipain-2 as our target protein (2GHU.pdb) obtained from the protein data bank and docked with twenty different available anti malaria drugs in order to find the effective and efficient drugs. Drug Diocopeltine A was found as the best lowest binding energy with the value of -8.64 Kcal/mol. Thus, it can be selected as the anti malaria drug candidate.

Keywords: *Genetic algorithm , drug candidte , rational drug design , molecular docking , docking , protein , ligand , computer simulation algorithm and computational molecular docking.*

1. INTRODUCTION

Rational drug design (RDD) is a central paradigm in pharmaceutical research. It begins with predefined knowledge about specific compound or drug and then studies its behavior towards the disease [1]. Drug is known as any chemical substance that, when absorbed into body of a living organism, alters normal bodily function. Moreover, the drug design is known as the procedure that used to find drugs by design through several stages based on their biological targets. The   development of drug design and discovery starts by discovering the drug(s) then register it for the marketing and commercialization purpose [2]. The drug discovery and development is a costly process that includes extensive clinical testing. The discovery and designing of new drugs involves several phases, namely as identify the disease, isolate the protein target causing the disease, finding the drug, pre-clinical test, clinical phases 1,2 and 3 and finally FDA[1] approval[3]. Fig. 1 summarizes the RDD process:
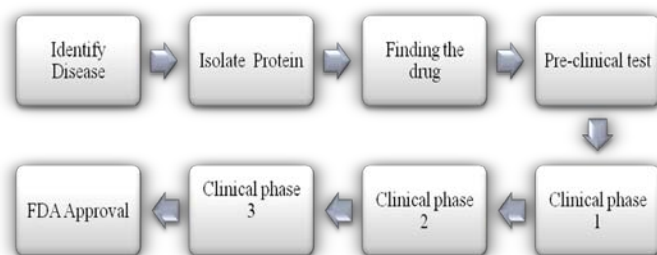

Fig 1: Drug Design and Discovery Process

2. GENETIC ALGORIHM

Genetic Algorithm is a search heuristic that imitate the process of natural evolution. Genetic algorithms (GAs) are inspired in Darwin's theory of evaluation by natural selection and are powerful tool in difficult search and optimization problems [4]. Genetic algorithm model the solution of the problem into data structure called chromosome or genotype which represents the possible solutions called individuals or phenotypes. A sequence of genetic operators is applied to these chromosomes in order to achieve a high optimization of the problem [5]. Two components play a significant role in the GA method: the problem codification and the evaluation function. The problem codification is the mapping that is made between the chromosomes and the individuals. Typically, the individuals are mapped into a string of 1's and 0's indicating the presence, or not of some feature or characteristic as shown in Fig.2. The evaluation function takes one individual and calculates its fitness. Usually, the fitness is a performance measure of the individual as a solution to the problem. However, Fig. 3 describes the pseudo-code of general genetic algorithm.

f0	f1	f2	F3	F4	F5	F6	F7	F8	F9
1	1	1	0	0	0	0	0	1	0

Fig 2: Chromosome representation in GA

---

[1] Stands for Food Drug Administration.

Algorithm 1 A Genetic Algorithm Pseudo-Code
1: Choose an initial random population of individuals
2: Evaluate the fitness of the individuals
3: **repeat**
4:     Select the *best* individuals to be used by the genetic operators
5:     Generate new individuals using crossover and mutation
6:     Evaluate the fitness of the new individuals
7:     Replace the *worst* individuals of the population by the best new individuals
8: **until** some stop criteria

Fig 3: Genetic Algorithm Pseudo-Code

Fig. 4 shows the genetic algorithm flowchart. In order to evaluate the optimized solution, it must goes through certain specific as stated in the flowchart.

Fig 4: Genetic Algorithm Flowchart

3. DRUG DESIGN

Drug design is known as approach that uses specific tools to explore and search   for candidate drugs by design. It is based on their biological targets such as the target protein and receptor. Naturally, a drug target is a key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition [6]. A number of approaches try to stop the functioning of the pathway in the diseased state by causing a key molecule to stop functioning. Drugs may be designed that bind to the active region and inhibit this key molecule. However these drugs would also have to be designed in such a way as not to affect any other important molecules that may be similar in appearance to the key molecules. The structure of the drug molecule that can specifically interact with the bimolecular (protein) can be modeled using computational tools such as Auto dock simulation. These tools can allow a drug molecule to be constructed within the bimolecular using knowledge of its structure and the nature of its active site. Construction of the drug molecule should be made inside the programs by combing the two files together in the same directory [7].

4. MOLECULAR DOCKING

Molecular docking is a computer simulation or computational tool that uses specific procedures to predicate the conformation of a receptor-ligand complex, where the receptor is usually a protein and the ligand is either a small molecule (drug) or another protein [8]. It is consider as the key step in identifying potential drug candidates [9]. In addition, it is known as the automated computer algorithm that determines how candidate ligands will be docked into the active site of the target protein. It involves several processes such as determination of orientation, conformational geometry and the scoring function. The real task of docking algorithm is to locate the ligand in many different orientations and conformations in the active site of the selected protein and then calculate the score for each ligand docked in the protein. Moreover, docking consider as the common tool used in the field of computational drug design. It predicates if a selected drug will be a good inhibitor of target protein. As a result, pharmaceutical companies evaluate and examine a huge numbers of ligands using docking technique [10]. The next phase after analyze the drugs, the best docking score will be selected as the synthesized and it will be test in the laboratory.
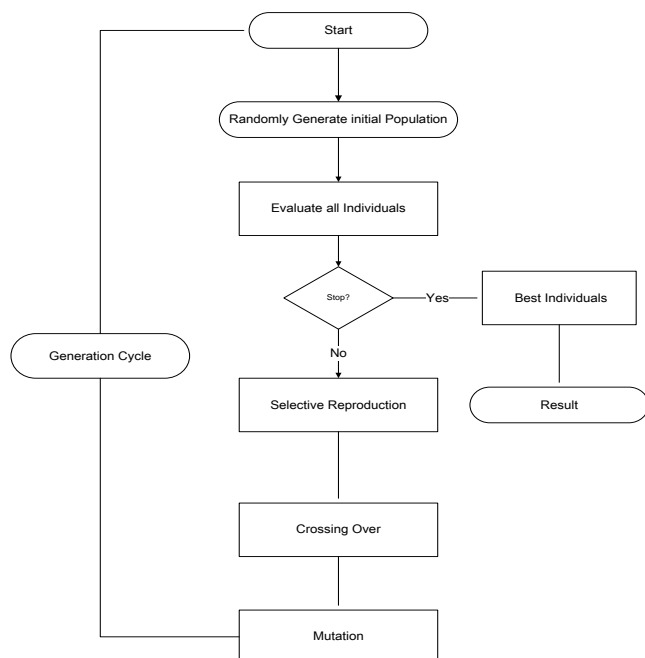
5. MOLECULAR DOCKING PROCESS

The process of searching for a ligand that fit to protein binding site is called molecular docking [11].It starts with finding the proper docking conformation of the drug docked into the target protein. It will be elaborated more in the techniques used in docking.   The binding mode of the ligand with the target receptor can be distinctively defined based on its state variables. It encompass of its position(x-, y- and z-translation) in the 3D structure, orientation (Euler angles, axis-angle, or a quaternion), and if the ligand is flexible, its conformation (torsion angles for each rotatable bond). Each of these state variables describes one degree of freedom (DOF) in a multidimensional search space, and their bounds describe the extent of the search. Generally, the rigid docking is faster than the flexible ligand due to the size of the search space is smaller than the flexible.

6. METHODOLOGY

In this paper we have conducted the experiment on

1094

20 drugs to select the best drug candidate. Beside that all software(s) used for the analysis is free for academic used. In the following we summarize all the sites/software used in this project:

a. Protein Data Bank (PDB)

✓ Protein Data Bank (PDB) is a worldwide repository for processing macromolecular structure data.
✓ It has more than 25000 structures determined by x-ray diffraction.
✓ Protein structure downloaded from the site with particular keywords or numbers.

b. PyMol

✓ PyMol is Python-based visualization software.
✓ Used to visualize the PDB files obtain from RCSB database.
✓ Measure the distance between atoms.

c. AutoDock 4.2

✓ Drug discovery tool uses genetic algorithm.
✓ It is used to generate all the necessary input files and docking grid.
✓ Used to perform the docking experiments in the selected drugs and proteins.
✓ It is automated docking tools designed to predict how small molecules/ligands such as drug candidate bind to receptor or protein of known 3D structure.
✓ Allows users to interact with molecular and perform computational experiments easily.
✓ Auto dock consists of two separate programs namely as Auto Dock and Auto Grid

▪ **Auto dock:** used to perform the docking of the ligand (drug) to a set of grids describing the target protein.

▪ **Auto Grid:** used to pre calaculates these grids.

7. TECHNIQUES USED IN DOCKING

The molecular docking process can be divided in to two phases. One is the searching algorithm, which finds possible binding geometries of the protein and its ligand [12]. The other is the scoring function, which ranks the searching results and selects out the best binding geometry based on the energies of the complexes. The docking software name Autodock4.0 is used for docking [13]. A typical docking study requires three computational steps before running the docking program which is as the following [14]:

1. Preparation of receptor & ligand files.

2. Calculation of affinity maps by using a 3D grid around the receptor & ligand.

3. Defining the docking parameters and running the docking simulation.

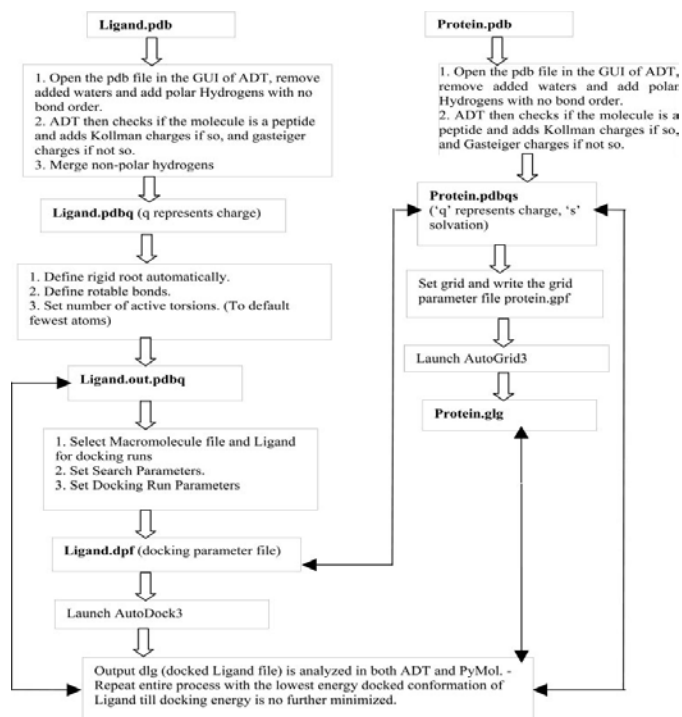The following diagram shows the common steps involved in the docking process:

Fig 5: Docking Procedure Chart

Fig. 5 shows the preparation procedures that perform in order to dock the ligands or drugs to the protein active site. Now, we will explain the process that involved in the docking mechanism starting from downloading the target protein till the end of docking and getting the output file in the .dlg format which is docked ligand file contains all the required information about the ligand such as binding energy value, intermolecular value and so on.

8. METHODS AND MATERIALS

First of all we have to identify the potential target disease which is in case of Malaria. The next step is to find out the 3D structure for this disease which is 2GHU.pdb; and it was retrieved from RCSB database.

Table 1: Protein Selection and Identification

Target Disease	Target Protein (PDB)
Malaria	2GHU.pdb

a. Target Selection and Identification

The search for effective, safe, and affordable drugs for malaria treatment is an urgent need worldwide [15]. Parasite Plasmodium Falciparum develops a special organelle comprising a set of proteases to utilize cytosolic proteins of the host cell as a food source during its life cycle within host red blood cell. There are several target proteins that play a significant role in life cycle of Plasmodium Falciparum. These proteins target have been well studied and then validated for their significance survival as well as their "drug-ability" [16] .Moreover, with the sequencing of the Plasmodium genome several new potential targets came into light. These potential antimalarial drug targets including targets involved in hemoglobin degradation. This protease contains plasmepsins and falcipains and they have been treated as potential drug targets [17].

b. Ligand (Drug) Identification

Once the therapeutic target has been identified, we must then find one or more Drugs (e.g., chemical compounds or molecules) that interact with the therapeutic target so as to induce the desired therapeutic effect, e.g., through antiviral or antibacterial activity. In order to discover the ligands (Drugs) whose pharmacological properties are likely to have the required therapeutic effects, researchers must test a large variety of them on one or more targets. However, in this paper we retrieved the ligands from pubchem chemical database. All the compounds were in .sdf format and then we convert it into .pdb format for ligands. The conversion is done by using openbable software which is open source software that helps us to convert from one format to another easily.

c. Pubchem

Pubchem is a free chemical database structures of small organic molecules and information on their biological activities hosted in NCBI web server. PubChem is organized as three linked databases within the NCBI's Entrez information retrieval system. These are PubChem Substance, PubChem Compound, and PubChem BioAssay. PubChem also provides a fast.

9. GENETIC ALGORITHM IN AUTODOCK

Genetic algorithms (GAs) are inspired in Darwin's theory of evaluation by natural selection and are powerful tool in difficult search and optimization problems [5]. Genetic Algorithm works with a population of individuals where each individual represents a possible solution for the problem .In the ligand-protein docking, a candidate solution specifies the position of the ligand with respect to the protein. Autodock uses a Lamarckian Genetic Algorithm to search for the optimal docking position of a ligand inside the

protein [18]. Auto Dock represents the chromosome as a vector of real numbers as shown in figure 4. Each component in the vector stands for a gene. The individual chromosome has three genes representing the ligand translation, four genes representing the ligand orientation and the other genes represent the ligand conformation [19].

❖ Solution Space

Auto Dock program intend to discover the optimal and best docked position of the ligand inside a 3D grid box specified in the active site of the receptor. The 3D grid box represents the solution space. The box is a continuous space and thus the total number of possible solutions is unlimited. Each 3D point inside this grid box is a potential candidate where the ligand can centered.

❖ Genetic Code (Chromosome)

Genetic algorithm usually identifies the notion of an individual and its genetic code (chromosome) with respect to the problems solution space [20]. The genetic code must uniquely identify a candidate solution. Auto Dock treats a particular orientation of ligand inside the protein as an "individual" in its GA. Hence, Auto Dock's genetic code uniquely identifies the orientation of a ligand. Auto Dock represents the chromosome as a vector of real numbers. Each element in the vector stands for a gene. Pictorial representation of the chromosome can be found in the following figure. Following are used as genes contained in an individual's chromosome. The following figure shows the representation of GA chromosome in the Auto Dock.

T_x	T_y	T_z	Q_x	Q_y	Q_z	Q_w	R_1	R_n

Fig 6: GA representation in AutoDock

The Fig. 6 shows the parameters that used as genes contained in an individual's chromosome. Translation genes (T_x, T_y, T_z) specifying the center of ligand inside the 3D grid [18]. On the other hand, Quaternion genes (Q_x, Q_y, Q_z, Q_w) specifying the angles of' N' rotatable bounds in the ligand.

❖ Genetic operation

Autodock performs all the genetic operators namely, selection, crossover and mutation on the genetic code.

❖ Fitness Evaluation(Energy Evaluation)

Fitness evaluation (FE) evaluates the total energy of a particular orientation on the ligand inside the protein. Lower is better [18]. The energy evolution is used frequently

1096

in AutoDock for every generation to recalculate the binding energy [19].

10.   PROTEIN PREPARATION IN ADT

The crystal structure of falcipan-2 was retrieved from Protein Data Bank. The 3D coordinates used in this project are 2GHU.pdb. After retrieving the protein, save in the right directory with the collection of drugs. The next step is to read the molecular from the AutoDockTools (ADT) program and then remove all the water molecules and add the hydrogens to it. In addition, we have to make sure that the Kollman charges are added to the same protein.

Fig 7: 3D structure of the target protein

11.   LIGAND PREPARATION IN ADT

The candidate antimarlais were downloaded from pubchem database. They were obtained from a comprehensive literature survey. The table below contains the pubChem ID, molecular weight, Hydrogen Bond Donor and Hydrogen Bond Acceptor. Basically, the format of the drugs in this database is on .sdf so, we have to convert it to .pdb. In order to do so, we used the openbabel application which helps us to reformat the drugs.

a. Open babl

It is free software, a chemical expert system mainly used for converting chemical file formats. Due to the strong relationship to informatics this program belongs more to the category cheminformatics than molecular modeling. It is a viable for windows, UNIX and Mac Os. [21]It is mainly used to convert from one format to another. For example, in our case, I want to convert from .sdf to .pdb. The following commend line should be executed in order to perform such conversion:

```
babel -i input_file -o output_format output_file
```

In this case we chose the .pdb as our extension instead of .sdf which is obtained from pubchem database. After typing the above commend in Linux terminal, it will convert the sdf format to pdb which is readable in the autodock simulation.

b. Grid File parementers in ADT

After we have done the preparation for the target protein and the ligand now the task is to precalaculate these grids. In order to create the 3D box (grid) in which the protein molecule is enclosed we need to calculate the interaction of docking energy correctly and accurately. The grid box should be large enough to allow the drug to rotate freely inside the grid box, even when the ligand is in its most fully extended conformation. The parameters required to create such a grid are stored in the grid parameter file.gpf. In this experiments we used the 60X60X60 coordinates grids. Then, after that the autogrid4 run in the terminal to create a map for every atom type in the drug and create the corresponding macromolecular file with the following extension .glg

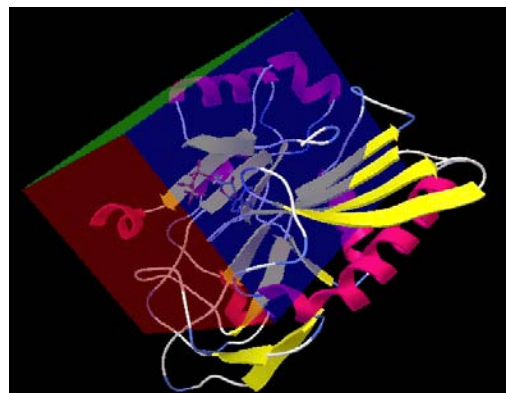**autogrid4 –p file name.gpf –l filename.glg**

Fig 8: Grid Box used in the simulation

12.   DOCKING FILE PARAMETER IN ADT

The primary goal of using docking parameter is to instruct Autodock about the ligand to move, the map files to use and some other properties that has been defined in the autogrid process. In the search methods about autodock, we select to work on genetic algorithm and we have set different value of crossover, mutation and population size to observe the results. Finally, we run the autodock4 in the terminal and it docked ligand files with the extension of .dlg which is docked ligand file.

autodock4 –p filename.dpf –l filename.dlg

13.   EVALUATING DOCKING RESULTS

Regardless of drug-protein docking programs has been used to conduct the simulation[22], docking algorithm parameters should be selected wisely in order to get the

appropriate dock to the ligand in the appropriate conformation in the active site. In order to validate the docking we checked the binding energy result which obtains by the following equation:

$$\Delta G = \Delta G_{vdw} \sum_{i,j}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + \Delta G_{hbond}\sum_{i,j}E(t)\left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{rj}^{10}}\right) + \Delta G_{ele}\sum_{i,j}\frac{q_{ij}}{\varepsilon(r)r_{ij}} +$$

$$\Delta G_{tor}N_{tor} + \Delta G_{sol}\sum_{i,j}\left(S_iV_j + S_jV_i\right)e^{\left(\frac{r_{ij}^2}{2\delta^2}\right)}$$

This equation used to calculate the binding energy between protein and ligand in Autodock 4.0. The total binding energy is composed of

- ΔGvdw : van der Waals
- ΔGH-Bound: Hydrogen Bonding
- ΔGelec: Electrostatics
- ΔGdesolv: Desolvation
- ΔGtor = Wtor Ntor

So the above are the weight for the respective terms.

- Aij, Bij, Cij, Dij are atom type pair specific terms.
- rij: is the interatomic distance.
- E (t): correspond to a directional weight based on the angle t between the probe and the target atom. (That gives H-Bond term the character of "directional").
- Ntor : is the number of rotatable bonds.
- Si and VI are salvation parameters.
- σ is a distance weighting function set to 3.5 A
- qi and qj are point charges on atoms i and j.
- ε(r) is dielectric constant.

14. EXPERIMENTAL RESULTS

The molecular docking study was performed on protein target name falcipain-2 (2GHU.pdb) and a set of 20 Drugs retrieved from pubchem database. The experiment was conducted in three different cases with three different workstations. In the first case we used the default parameters then in the next case we used the parametric study proposed in [23] to select the best drug candidate. Finally, we conducted the experiment to calculate the required docking time to finish one simulation. The results shows drug Dioncopeltine A has the lowest binding energy which is -8.64 Kcal/mol. Thus, it selected as the best candidate drug design in the experiment.

- Lipinski's rule of five

The following properties are essential for good permeation

1) The molecule has less than five proton-donors. (HBD<5).
2) The molecular weight is smaller than 500 Dalton (MW<500).
3) log P smaller than 5(log p<5).
4) The molecule has less acceptors than 10 (HBA<10).

CASE 1:

In the first case we performed the simulation [24] in all the compounds (drugs) with the default parameters as in table 2.

Table 2: Auto Dock parameters and processing data

Parameters	Value
Maximum Number of GA runs	10
Population size	150
Maximum number of evaluation	250000
Rate of Gene mutation	0.02
Rate of Crossover	0.8

We performed the experiment in single workstation AMD phenom™ 9600B Quad-core processor (2.3 GHz) with 2 GB memory size. The implementation is done using AutoDock simulation tool. Also, we concern about the time consuming in running of each simulation. While we conducting the experiments we got different time based on the parameters chosen. Table 3 table contains the binding energy for each drug. As we said early, the best binding energy is the lowest.

Table 3: Binding Energy results

Drug Name	Pubchem ID	Binding Energy[2]
Amodiaquine	2165	-8.08
Artemether	456408	-7.68
Artemisinine	68827	-7.71
Artesunate	65664	-7.12
Atovaquone	74989	-7.68
Chloroquine	2719	-7.53
Dapsone	2955	-6.07
dihydroquinghaosu	107770	-7.63
Dioncopeltine A	185971	-8.64

---

[2] Since the value of binding energy is negative so the graph is appear like that.

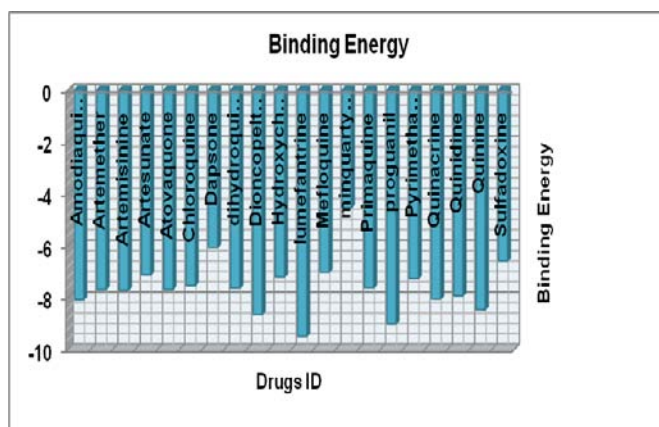| Hydroxychloroquine	3652	-7.2
lumefantrine	6437380	-9.48
Mefloquine	4046	-7.01
minquartynoic acid	183614	-4.53
Primaquine	4908	-7.62
proguanil	6178111	-9
Pyrimethamine	4993	-7.26
Quinacrine	237	-8.05
Quinidine	441074	-7.93
Quinine	8549	-8.45
Sulfadoxine	17134	-6.58

Fig 9: Binding Energy Diagram.

Based on fig.9, lumefantrine drug has the highest binding energy which is consider as the lowest binding energy. But based on Lipinski's rule of five, it cannot be selected as drug candidate for our experiment because log p value is 8.7 which is higher than 5 as stated in the rule. So, we have to search for another drug that satisfies the conditions.
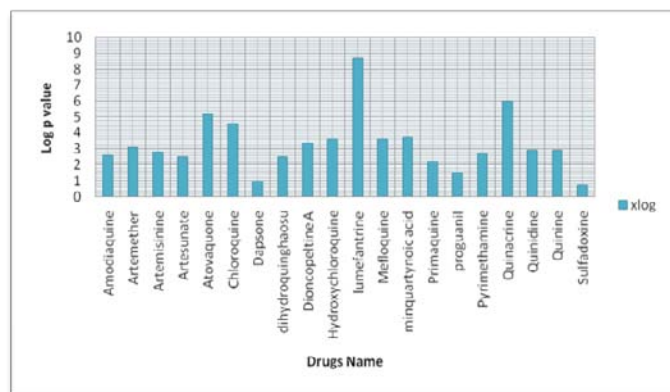

Fig 10: log p values for various drugs

As we can see from fig.10, the drug Dioncopeltine A satisfies the requirements. The lowest binding energy was around the -8.64 Kcal/mol with log p value is 3.4 which is less than 5 .hence; it satisfies the Lipinski's rule of five.

CASE 2:

In this case, we preformed the method which proposed in [23]. The main aims of this method were to study the performance of the genetic algorithm (GA) in the selected drugs. Since we have 20 selected drugs for the malaria with protein target, the parametric study used to find the suitable set of parameters. We conducted experiment of 480 samples with different probabilities of population size, crossover rate and mutation rate. The population size set from 50,100,150 and probability of crossover set from 0.2, 0.4, 0.6, and 0.8. Finally, the mutation rate set to 0.01 and 0.02. Table 4 show the results obtained from the experiment.

Table 4: Results of Parametric study

| Drug Name	ID	MW	Log p	HB A	HB D	Pop size	Crossover Rate	Mutation Rate	Binding energy
Amodiaquine	2165	355.8611	2.6	4	2	100	0.2	0.02	-8.41
Artemether	456408	298.3746	3.1	5	0	50	0.4	0.01	-7.70
Artemisinine	68827	282.3322	2.8	5	0	50	0.4	0.01	-7.69
Artesunate	65664	384.4208	2.5	8	1	150	0.2	0.02	-7.11
Atovaquone	74989	366.8374	5.2	3	1	150	0.8	0.01	-7.68
Chloroquine	2719	319.8721	4.6	3	1	150	0.2	0.01	-7.53
Dapsone	2955	248.3008	1	4	2	150	0.8	0.01	-6.07
Dihydroquinghaosu	107770	284.348	2.5	5	1	50	0.8	0.01	-7.63
Dioncopeltine A	185971	379.4489	3.4	5	4	100	0.2	0.01	-8.64
Hydroxychloroquine	3652	335.8715	3.6	4	2	150	0.6	0.02	-7.32
lumefantrine	6437380	528.9401	8.7	2	1	150	0.4	0.01	-9.48
Mefloquine	4046	378.3121	3.6	9	2	150	0.2	0.02	-7.0
minquartynoic acid	183614	284.3496	3.7	3	2	100	0.8	0.01	-4.53
Primaquine	4908	259.3467	2.2	4	2	150	0.2	0.01	-7.63
proguanil	6178111	253.7312	1.5	1	3	50	0.8	0.02	-8.63
Pyrimethamine	4993	248.7114	2.7	0	2	50	0.8	0.01	-7.63
Quinacrine	237	399.9568	6	4	1	50	0.8	0.01	-8.05
Quinidine	441074	324.4167	2.9	4	1	150	0.2	0.01	-7.93
Quinine	8549	324.4167	2.9	4	1	50	0.4	0.02	-8.44
Sulfadoxine	17134	310.3289	0.7	5	2	150	0.8	0.01	-6.65
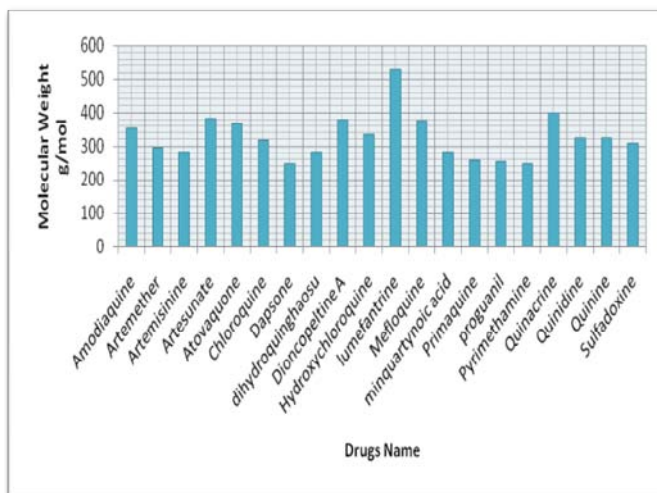
Fig 11: molecular Weight values for drug candidates

Based fig.11, we found that, drug lumefantrine has the height molecular weight among the drug candidates.

From table 4, we conclude that the drug Dioncopeltine A has the lowest binding energy from the docking simulation which is -8.64Kcal/mol. The result is shown in table 4 which is partial result. Among the 480 samples we selected the best drugs with lowest binding energy. In table 4, we summarize the results and it shows that the best binding energy was with population size of 100, crossover rate 0.2 and mutation rate is 0.01.

The expected output is illustrated in fig.11. However, the fig.9 shows the lowest binding energy based in the first run, but since the lumefantrine drug does not follow the Lipinski's rule of five with the value of log p 8.7 which is greater than 5. So, we have to search for the alternative candidate which is Dioncopeltine A with the value of -8.64 Kcal/mol.

Fig 12: docking file result

Fig.12 demonstrates the partial results obtained from the docking simulation. The first column is the cluster rank which shows the point where the lignd docked into the best conformation with the protein active site. The second column is lowest binding energy. We took this parameter as standard criteria to show the best drug fit with the protein. Usually, the lowest binding energy is the best candidate drug. The mean binding energy is the parameter used to find the critical point between two values. Number in cluster showed how many clusters are in this conformation and finally the histogram column which shows the visualize the experiments.

CASE 3:

In this case, we attempted to calculate the time required to perform the docking for each drugs (ligand). We investigate the central processing unit (CPU) time required for completing a docking attempt on docking simulation. Table 10 shows the docking time spent to perform single docking simulation:

Table 5: Docking times for ligand

| Drug Name	Pubchem ID	Docking time (Minutes)
Amodiaquine	2165	10
Artemether	456408	5
Artemisinine	68827	4
Artesunate	65664	9
Atovaquone	74989	9
Chloroquine	2719	9
Dapsone	2955	7
dihydroquinghaosu	107770	5
Dioncopeltine A	185971	12
Hydroxychloroquine	3652	10
lumefantrine	6437380	14
Mefloquine	4046	10
minquartynoic acid	183614	10
Primaquine	4908	8
proguanil	6178111	7
Pyrimethamine	4993	5
Quinacrine	237	11
Quinidine	441074	9
Quinine	8549	8
Sulfadoxine	17134	9
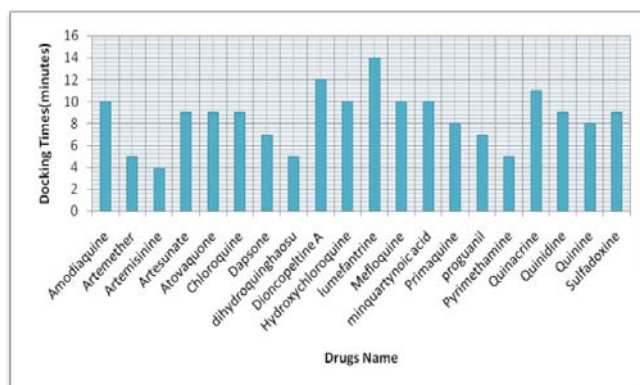
Fig 13: The docking time for drugs

From fig.13, we can conclude that, drug lumefantrine required some times to perform the docking simulation due to the molecular weight. So there a great impact of the molecular weight in the molecular docking simulation. The time required to do this docking is approximately 14 minutes. To overcome such delay in the experiment some researcher proposed some mechanism as discussed in [25]. They proposed the use of grid computing in this simulation that will help to perform more than one simulation in same computers and in parallel.

15. CONCLUSION AND FUTURE WORK

In conclusion, designing drugs through the rational drug design (RDD) reduce the time require to conduct the experiment in the laboratory. In addition, computational molecular docking. In fact, this project discussed the usage of genetic algorithm (GA) as crucial tool in the rational drug design. It is clear that using the computational techniques in molecular docking helps the scientists and the researchers to simulate their experiments easily and rapidly. Moreover, we found that, from the experiment that there are so many

antimalaria drugs are in the market but choosing the effective and efficient among them is still the problems. By using autodock docking simulation we hope that will overcome such issues. The Protein-Ligand interaction plays a significant role in structural based drug designing. In this project, we performed the experiment in Malaria as our target disease using the Genetic Algorithm tools as main approach. The receptor Plasmodium falciparum which is the most dangerous and widespread disease-causing species has been selected for further investigation. In addition, falcipain-2(FP-2) of p. Flaciparum is a protein-family that plays an important role in the parasite life cycle. Also, I identified the drugs that were used against Malaria Disease. When the receptor (2GHU) was docked with the 20 selected drugs, the results from molecular docking simulation indicate that drug Diocopeltine A was found as the best lowest binding energy with the value of -8.64 Kcal/mol and crossover rate was around 0.2, mutation rate was 0.01 and population size was around 100. Thus, it can be selected as the anti marlia drug candidate.

Finally, for the future work, further investigations on selected drugs and in vivo studies are necessary to develop potential chemical entities with the help of genetic algorithm tools in the computational molecular docking. In addition, I do propose the use of serial processing which sequentially solve each piece of a problem on block beside use the grid computing in the future work.

REFERENCE

[1] C. Ramu, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, and J.D. Thompson, "Multiple sequence alignment with the Clustal series of programs", presented at Nucleic Acids Research, 2003, pp.3497-3500.

[2] Young, David C. (2009) "Computational Drug Design, a guide for computational and medicinal chemists", John Wiley R.D. Taylor, P.J. Jewsbury & J.W. Essex: A review of protein-small molecule docking methods, Journal of Computer-Aided Molecular Design, 16: 151-166, 2002

[3] Young, David C. (2009)"Computational Drug Design, a guide for computational and medicinal chemists" Wiley-Interscience John Wiley & Sons, USA, 2009pp.230.

[4] Holland, J. H., "Adaptation in Natural and Artificial Systems", MIT Press, Cambridge, MA, 1975.

[5] Holland, J. H., "Adaptation in Natural and Artificial Systems", MIT Press 1992.

[6] R. Chenna, "Multiple sequence alignment with the Clustal series of programs," Nucleic Acids Research, vol. 31, no. 13, pp. 3497-3500, Jul. 2003.

[7] M. A. Jamakhani, "IN-SILICO DESIGNING AND DOCKING ANALYSIS ON INSULIN-LIKE GROWTH FACTOR 1 RECEPTOR," Bioinformatics, vol. 1, no. 1, pp. 69-83, 2010.

[8] Michalewicz, Z., Genetic Algorithms Data structures = Evolution program.3rd ed.1996, London, UK: Springer-Verlag.387.

[9] H. M. Berman et al., "The Protein Data Bank.," Nucleic acids research, vol. 28, no. 1, pp. 235-42, Jan. 2000.

[10] Young, David C. (2009)"Computational Drug Design, a guide for computational and medicinal chemists, Wiley-Interscience John Wiley & Sons, USA, 2009pp.128.

[11] Teodoro, M.L., Phillips, G.N., Jr. and Kavraki, L.E. (2003). "Understanding protein flexibility through dimensionality reduction". J. Comput. Biol. 10: 617-634.

[12] Smith, D.M.; Daniel, K.G.; Wang, Z.; Guida, W.C.; Chan, T.H.; Dou, Q.P. "Docking studies and model development of tea polyphenol proteasome inhibitors: Applications to rational drug design. Proteins" 2004, 54, 58–70.

[13] J. H. Wikel, C. Bingham, and E. Bonabeau, "A Drug Candidate Design Environment Using Evolutionary Computation," Mobius, pp. 1-13.

[14] A. Dhanik, J. S. Mcmurray, and L. Kavraki, "On modeling peptidomimetics in complex with the SH2 domain of Stat3," Biomedicine, vol. 3, pp. 3-6, 2011.

[15] Philip J. Rosenthal: "Review Antimalarial drug discovery: old and new approaches". The Journal of Experimental Biology 206, 3735-3744.

[16] Breman, J. G.; Alilio, M. S.; Mills, A. "Conquering the intolerable burden of malaria: what's new, what's needed: a summary". Am. J. Trop. Med. Hyg. 2004, 71S, 1-15.

[17] Pattanaik, P.; Raman, J.; Balaram, H. "Prospective in drug design against malaria". Current Topics in Medicinal Chemistry". 2002, 2, 483-505.

[18] E. Atilgan, "Efficient Protein-Ligand Docking Using Sustainable Evolutionary Algorithms," Evolutionary Computation, pp. 113-118, 2010.

[19] S. Kannan and R. Ganji, "Porting Autodock to CUDA," in Proc. 2010. IEEE Congress on Evolutionary Computation, Jul. 2010, pp. 1–8.

[20] AlShaikhli, Imad Fakhri Taha (2011) "A practical method for the analysis of genetic algorithms". Journal of Advanced Computer Science and Technology Research(JACSTR), vol.1. pp. 1-9. ISSN 2231-8852

[21] Smith, D.M.; Daniel, K.G.; Wang, Z.; Guida, W.C.; Chan, T.H.; Dou, Q.P. "Docking studies and model development of tea polyphenol proteasome inhibitors: Applications to rational drug design. Proteins" 2004, 54, 58–70.

[22] R. Zheng, T.-sheng Chen, and T. Lu, "A Comparative Reverse Docking Strategy to Identify Potential Antineoplastic Targets of Tea Functional Components and Binding Mode," International Journal of Molecular Sciences, vol. 12, no. 8, pp. 5200-5212, Aug. 2011.

[23] I. Fakhri, A. Shaikhli, and I. Khalil, "An Improved Genetic Algorithm For Solving The Multiprocessor Scheduling Problem," Science, vol. 5, no. 12, pp. 947-951, 2011.

[24] C. L. Gupta, S. Singh, M. Kalim, and A. Khan, "Molecular docking studies of antimalarial drug and its analogues against falcipain-2 protein," Methodology, vol. 1, pp. 14-22, 2011

[25] S. Park and J.-in Kim, "International Conference on Convergence and Hybrid Information Technology 2009 A Molecular Docking System using CUDA," Performance Improvement, pp. 0-5, 2009.