IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Volume-5, Issue-2 E-ISSN: 2347-2693

Ligand Design- A Multiobjective Optimization Based Approach

Sreedhar D.[1*] and Chandrika S.[2]

[1*]Dept.of Computer Science, St. Thomas College of Engineering & Technology (Kerala University), Kerala, India
[2] Rajiv Gandhi Center for Bio-Technology, University of Kerala, Thiruvananthapuram, India

[*]*Corresponding Author: deepas267@gmail.com*

Available online at: www.ijcseonline.org

Received: 22/Jan/2017 Revised: 30/Jan/2017 Accepted: 17/Feb/2017 Published: 28/Feb/2017

***Abstract*— Excess accumulation or inadequate production of certain protein leads to diseases. A drug can play most important role in this scenario. A drug is an organic molecule that triggers or inhibits the function of a biomolecule such as a protein; this will be beneficial to the patient. In the most basic sense, drug design involves the design of small molecules that are complementary in shape and charge to the bio molecular target with which they interact and therefore will bind to it. The process of drug discovery by laboratory experiments is time consuming and very expensive. To reduce the time and cost of drug discovery process, computational techniques are incorporated to speed up the process. Initially dock the molecular fragments obtained by breaking the sigma bonds between the bioactive molecules against tuberculosis to the target site of the protein using docking software known as AutoDock. The score obtained from this is given as input to the program. Then, prioritize the fragments using Multiobjective Differential Evolution (MODE) algorithm with two objectives namely oral bioavailability and free energy. Next step is to design set of ligand molecules that can be represented as an array of fragments. Then analyses the performance of proposed approach by comparing it with another multiobjective optimization algorithm namely Archived Multiobjective Simulated Annealing (AMOSA).**

Keywords—*Drug Discovery, Drug Design, Multi-objective Differential Evolution, AMOSA*

I. INTRODUCTION

Drugs play a major role in the existence of humans now days. Classical drug discovery was solely based on observations of the natural phenomena and the consequences of consumption of materials that relieved distress. Drug discovery is mostly portrayed as a linear, intense, and lengthy consecutive process that starts with target identification which need to be validated before lead discovery process (an intense process proceed to find small drug like molecule), followed by lead optimization which will progress in to pre-clinical studies and if it is successful conduct experiments to determine if such compounds satisfy a number of pre-set criteria for initiating clinical development. After the successful finishing of clinical testing, the compounds (drugs) can be marketed as a drug. For the pharmaceutical industry, a number of years are taken to bring a drug from discovery to market. This is about 15 years and costing up to $1.4 billion dollars. Traditional method of drug discovery was by synthesizing compounds in a time-consuming multi-step processes against a battery of in vivo biological screens and further finding out the promising candidates for their pharmacological properties, metabolism and potential toxicity. Such a development process has resulted in high attrition rates with failures attributed to poor pharmacokinetics, lack of efficacy, animal toxicity, adverse effects in humans and various commercial and miscellaneous factors. Today, the process of drug discovery has been revolutionized with the advances in science and technologies.

Traditional Drug Screening:

- Random, trial and error
- Time consuming
- Very expensive
- Extremely low yield (1 in 100,000)

Computer-bsed Design:

- Target specific and structure-based
- Fast and automatic
- Very low cost
- High success rate

Drug design (rational drug design) is the innovative process of finding new drugs based on the knowledge of target protein. Sometimes, diseases are caused due to the variations in the production of certain protein (excess increase or insufficient production of certain proteins). The drug can activates or inhibits the function of a biomolecule such as a protein, thus it will helps to cure the diseases. The drug molecules obtained through drug design process are opposite in shape and charge to the biomolecular target with which they interact and therefore will bind to it [1]. Drug design depends on computer simulation (modeling) techniques is often referred to as computer-aided drug design (CADD). Basically, there are two categories of drug design approaches under CADD, namely, structure based and ligand based drug design. The former drug design relies on the knowledge of the three-dimensional structure of the biomolecular target. The latter one depends on the knowledge of other molecules (molecules that have already bind to biological target) that bind to the biological target of interest. A pharmacophore model that defines the minimum necessary structural characteristics a molecule must possess in order to bind to the target. Here, a model of the target is build based on the knowledge of what binds to it. Alternatively,

a quantitative structure-activity relationship (QSAR) is another approach under ligand based design. In QSAR model relationship between calculated properties of molecules and their experimentally determined biological activity may be predicted. De novo design, comes under structure based drug design strives to build novel small molecules with desired pharmacological properties. Generally two approaches are used for de novo design. The first method uses atoms to grow molecules and the other uses molecular fragments for the same. Fragment-based de novo design approaches generate novel structures by adding interacting molecular scaffolds to a fixed seed or primary scaffold. There are also several approaches that use atoms rather than fragments for ligand building which can produce more competitive results but process is more expensive and time consuming. Fragment-based de novo design approaches typically limit moiety addition to a library of fragments or chemical scaffolds which makes it more computationally powerful to find good ligands. Moreover, the scoring functions used to predict fitness of the designed ligand molecules make the algorithm faster. Though this approach is not as accurate as the comprehensive physics-based approaches but provides good approximation to probe further for drug discovery which renders more popularity to the approach. Fragment-based de novo design approaches use optimization techniques like genetic algorithm(GA)[2], simulated annealing (SA), and related methods for finding drug candidates (ligand). This approach with multiobjective optimization techniques are used in the present study to design ligand molecules against Tuberculosis (TB), which is the disease considered in this problem.

Goh and Foster [2] demonstrated a GA-based ligand design framework in which a tree structure representation is used for encoding ligands as chromosomes. In this interesting work, the nodes of the tree like structure (ligand) were filled in by fragments from a given library. The van der Waals potential was used as the fitness function to be minimized. However, the tree structure used in [2] was static irrespective of the protein target. This drawback was amended in a work of Bandyopadhyay et al. [3] where VGA [3] was proposed which addressed the need of altering the ligand size. Ecemis et.al. [4] implemented an evolutionary computation based approach for drug design. A set of compounds with better drug potential are identified before using this approach. The user selects a library of fragments for each components in order to create molecules (Ligands) with potentially better properties. Thus GA is applied to the resultant search space in order to seek the best solutions among the number of alternatives. Then computational models are determined to evaluate the suitability of GA created drug candidates. The user can adjust the optimal values of the model. The user can run the GA for a specific number of generations or indefinitely, in which case he should monitor its progress and stop it when the best population score exceeds a certain threshold or is no longer improving. When the GA stops, the feasible results obtained in the last population are presented to the user. The user evaluated the top 12 compounds in the population by providing negative or positive feedback. The feedback provided may strengthen or modify the direction in which the GA is heading. Ghosh et.al.[5] demonstrated an evolutionary approach to drug design using QBPSO algorithm. The target site considered here is the antiviral binding site of Human Rhinovirus strain 14. A tree like representation is employed for encoding ligands as fragments. The van der waals interaction energy of the drug protein complex is used as the objective function to be minimized. This strategy [5] made comparison of energy obtained by using fixed string length binary particle swarm optimization (BPSO) and variable string length BPSO (VBPSO). The Result was that the variable string length BPSO given minimum interaction energy, which means the better fitness value.

Again Ghosh et.al [6] implemented an evolutionary approach to drug design using NBGA. The main difference between GA and NBGA is that the latter uses ring topology to generate offspring and after crossover a trio selection is applied. The selection is based on the cost function and molecules with minimum cost function is selected. All the approaches mentioned above use scoring functions to evaluate the effectiveness of the ligands built. These scoring functions generally evaluate energy components and QSAR properties of the designed ligands to gauge their goodness. However, for a drug to be necessarily effective in an organism it must possess several additional properties other than being geometrically and chemically complementary to its target. Most importantly a proposed drug candidate must be synthesizable to qualify as a drug. A drug must possess high oral bioavailability so that it can be absorbed in the blood stream readily before it gets excreted. It does not participate in any nonspecific binding before it reaches the intended target. Optimization of all these properties is necessary to make sure that the in vivo behavior of the molecule is in a desired fashion [7]. In 2012, Sengupta et.al. [8] proposed MOO based approach for ligand design. This method employs a tree like structure to represent solutions and Archived Multiobjective Simulated Annealing as the optimization algorithm. Three objectives are considered to measure the goodness of the evolved molecule. Which are energy (Internal energy of a molecule and its interaction energy with the given target), chemical similarity to a reference scaffold and its oral bioavailability measure [8]. Low intra molecular energy ensures stable molecule and low-inter molecular interaction energy warrants stable complex formation between the designed molecule and the protein, implying better binding affinity [8].The multi objective optimization resulted in lower interaction energies and hence forms more stable complexes with the receptor.

In the present paper the drug/lead design problem has been posed in a multiobjective framework where the objectives are optimized separately and simultaneously. The MOO framework optimizes two objectives: energy component and oral bioavailability (OBA) measure. The OBA measure is a novel measure based on Lipinski's rule of five [8]. The rationale behind introducing this measure is to ensure design of small drug like molecules with better oral bioavailability. Apparently, it may seem that interaction energy will be minimized simultaneously as OBA measure gets maximized. But this is not the case, since large molecules will always have low interaction energy but its oral bioavailability a given inhibitor may not be ensured. Similarly, different derivatives of a single molecule may not have similar drug like properties or may not interact similarly to a given target. Thus, it can be concluded that these two properties are not related and needs to be optimized independently. The result of the proposed algorithm is a set of Pareto-optimal solutions which represents different tradeoffs in the objective space. The advantage of the proposed approach is that one can use multiple objectives of different types, which might be conflicting with each other, for simultaneous optimization. Although here the energy component, and a novel OBA measure are used as the objective functions in this work. This problem has also been modeled using another well-known MOO technique, AMOSA [9] for comparing the performance of MODE. The AMOSA and MODE are discussed in depth in section 2 and section 3 respectively. Section 4 describes ligand design using MODE and section 5 describes the result.

II. ARCHIVED MULTIOBJECTIVE SIMULATED ANEALING (AMOSA)

Simulated annealing is a probabilistic search technique based on the principle of annealing in metallurgy. It was developed by Kirkpatrick et. al. in 1983 for obtaining the global minima of a cost function which had several local minima. Like the annealing process, at each step SA strives to replace the existing solution by any other randomly generated solution. The new solution generated may either replace the previously existing solution or it may get discarded with a probability depending on the change in cost function value and temperature. Initially when the temperature is high the acceptance rate of poor solutions is more so that the search space can be well explored. As the temperature decreases the search space becomes more defined and the probability of acceptance of poor solutions gets reduced. If the cooling schedule is very slow, SA will take more time to find the optimal result. If it is too fast, then the algorithm might not be able to provide any good solution at all. If the cooling schedule and the stopping criteria are tuned to the problem at hand, then SA can provide a good approximation of optimal results quickly. Single objective SA is quite popular, but its application in multiobjective problems gets restricted due to its "search from a point" characteristic. To tackle this aspect of simulated annealing, Bandyopadhyay et al. have recently proposed an efficient multiobjective version of SA called AMOSA [9].

III. MULTIOBJECTIVEDIFERENTIAL EVOLUTION ALGORITHM (MODE)

Differential evolution (DE) is a recent optimization technique in the family of evolutionary computation. It is a simple, powerful, stochastic, and population based evolutionary algorithm for fast optimization. It was originated by Price and Storn in 1997 [10]. It is proposed as a variant of genetic algorithms to achieve the goals of robustness in optimization and faster convergence to a given problem. There exist some difference between DE and other evolutionary algorithms in their mutation and crossover phase. Also, it uses real numbers for representing each of the decision variables present in the chromosome. The approach proceeds by creating an initial population P of random individuals. Unlike some metaheuristic techniques such as genetic algorithms and evolutionary strategies, where perturbation occurs in accordance with a random quantity, DE uses weighted differences between solution vectors to perturb the population [10]. The approach proceeds by creating an initial population P of random individuals. Then create candidate solution (child) for each parent in the population. If the candidate dominates (is better than) the parent, replace the parent with new candidate. Otherwise the candidate is discarded. Mutate each parent in the population for creating candidates. For creating the child, select a main parent and three different parents randomly from the population and perform mutation. The mutation is done by adding the decision variable's value of one parent with the weighted difference of values of corresponding decision variable of other two parents [10]. Then apply objective function on main parent and newly created child to determine who will pass to the next generation. Obviously, the member having best cost is transferred to the next generation. This process continues until the maximum size of the population (NP) is reached.

It is difficult to handle multiple objectives with DE. Due to this reason an algorithm is needed that can optimize multiple objectives simultaneously. So DE is extended in order to attain multiobjective feature to it and is called Multiobjective Differential Evolution (MODE) algorithm. The result (non dominated solutions) obtained from MODE after evaluating the last generation is collectively called pareto optimal solutions (pareto front). To achieve the multiobjective optimization goals, the MODE methodology combines Pareto-dominance principles with DE and uses elitism in its evolution [11]. The MODE is a multiobjective optimization algorithm based on Differential Evolution (DE). The

proposed MODE methodology can be summarized by Algorithm 1.

Algorithm 1: MODE Algorithm

1. Input the required DE parameters like population size (NP), crossover constant (CR), scaling factor (F), maximum generation, number of objectives, bound constraints etc.
2. Initialize all the vectors randomly in the limit of bound constraints.
3. Set the generation counter, G = 0.
4. Perform mutation and crossover operations on all the population members.
    - For each parent, select three distinct vectors randomly from the current population. The selected vectors must not be the parent vector. These vectors combine to produce an offspring. So in DE, there are 3 parents that mutate to produce one offspring.
    - Calculate new mutation vector using the expression, $V_i(g) = X_{i3}(g) + F*(X_{i1}(g) – X_{i2}(g))$
    - Perform crossover using any crossover method.
5. Evaluate each member of the population and check if it is better or equal to the parents. Replace the parents with offspring in the next generation if the offspring is better or equal to the parents otherwise, the parents proceed to the next generation.
6. Increase the generation counter, G, by G+1. If G < GMAX, then go to step 4 and repeat mutation, crossover and selection.
    - If G = GMAX, then goto step 7.
7. Remove the dominated solutions in the last generation. A solution is dominated if there is another solution which is better than it in all the objectives.
8. Output the non-dominated solutions.

The population size, crossover constant, scaling factor, maximum generation and number of objectives are set to 100, .9, .2, 1000, and 2 respectively.

IV. LIGAND DESIGN USING MODE

Mycobacterium Tuberculosis is the causative agent of Tuberculosis (TB). TB is a bacterial infection that can spread through the lymph nodes and bloodstream to any organ in the body. It is most often found in the lungs. If TB bacteria are live in the body of patient in inactive form, the patient will never develop symptoms of TB. But if the immune system weakens, such as in people with HIV, Maleria, or elderly adults, TB bacteria can become active. TB bacteria in their active form can cause death of tissue in the organs they infect. About two billion people are infected in the world and 2 million die each year from the disease. Active TB disease can be fatal if left untreated. Since it is such a serious disease treatment should be given as quickly as possible.

RecA protein of Mycobacterium tuberculosis is the causative agent of TB and has been considered as the target protein for study. The intein splicing domain of RecA has been studied and its Adenosise triphosphate (ATP) binding site has been used as the active site to design small molecules. Primarily the bioassay results of the protein from PubChem have been studied to find molecules which are active against it [8]. The most frequent substructure obtained by analyzing the structure of these molecules are used here for further processing. This substructure is named as Seed. The Seed and active site information of the target protein is used for building ligand molecules by multiobjective ligand design approach.

A. Data Description

The proposed technique takes a seed, a fragment library, and the active site geometry of the target protein as inputs. A seed is a primary molecular scaffold supplied to the algorithm which is grown into different small molecules. The fragment library contains a number of molecular scaffolds which are used to grow the seed and derive it into plausible ligands [8]. The active site of the target guides the algorithm in determining the chemical property and the geometry of the molecules being designed. The input data are elaborately described as follows.

Fragment Library: The fragment library is created with the help of PubChem. PubChem contains bioactive molecules against mycobacterial proteins that can be downloaded before constructing the fragment library. The fragment library is formed by decomposing the sigma bonds between substantial atoms in the bioactive molecules such that no resultant fragment contains more than three rotatable bonds and it contains hundred such decomposed chemical frame work.

The Seed: The de novo design often requires a seed that might contribute greater chance of synthesizability to the designed molecule. The seed is the most frequent substructure in the bioactive molecules against mycobacterial proteins. Protein Data Bank (PDB) file 2IN9 contains the RecA protein [12]. The structure of the seed is shown in [8].

Active Site Processing: The biological functioning of the protein is determined by atoms that constitute the active site. Therefore, a ligand which is supposed to inhibit or alter the functionality of the target should be chemically and geometrically complementary to this site. The PDB file used to find the active site of the target is 2IN9 [12]. To help the algorithm build chemically relevant molecules preference matrix was employed. This matrix furnishes a preference measure of finding a fragment of the library in the vicinity of a particular atom on the active site. The matrix has n rows, where n is the number fragments. The preference measure of

finding fragment w in the neighborhood of active site atom is given in [8].The preference measure always retains a positive value. The NBI energy can be calculated with

NBI(w)= Van der waals energy + Electrostatic interaction energy (1)

The equations for Van der waals energy and electrostatic interaction energy are given in [8].

B. Proposed Method

This work aims to design drug molecules against tuberculosis by using Multiobjective Optimization (MOP) algorithm known as Multiobjective Differential Evolution (MODE) that optimizes more than one objective simultaneously. In first phase, the fragments are prioritised based on the free energy value of the fragments. The fragments are obtained by decomposing currently available bioactive molecules against tuberculosis. Then, select top 100 high priority fragments and apply multiobjective optimization algorithm such as MODE on them. For optimizing the fragments based on objectives, the properties of the fragments are evaluated by using docking software known as AutoDock. It will automatically dock molecular fragments to protein receptor target and returns the property values of the fragment namely LogP value, HBdonar (hydrogen bond donar), HBacc(Hydrogen Bond Acceptor), van der waals interaction energy and electrostatic interaction energy. Using these property values evaluate the following objective functions:

- Oral Bioavailability
- Interaction energy

Representation of the Solution: The solutions of MODE are called pareto optimal solutions. The set of fragments which constitute ligand molecules are the solutions in the present problem. A pareto optimal solutions in the present work is encoded using tree like representation and is stored as an array of positive integers. Numbers in all the odd position of the array representing a pareto optimal solution symbolize a molecular fragment which create the ligand molecule. The connectivity (the chemical bonds an atom can make with other atoms) that can be allotted for the fragments placing in odd positions is specified by succeeding even numbers. The connections are established between the heavy atoms by replacing the hydrogen atoms attached to them. Therefore, the number in each odd cell of the array is the node of the molecular tree and the value in the following even cell determines the number of children of that particular node on the tree.

Objective Function:

1) Oral Bioavailability

1. Initialize an array (A) of size four.
2. Check the number of hydrogen bond donors (HBD) of the molecular fragment under consideration.
3. A[0]=1, if HBD <= 5 else A[0]=1/(D+1) if HBD > 5 where, D= HBD-5.
4. Check the number of hydrogen bond acceptors (HBA) of the molecule under consideration.
5. A[1]= 1, if HBA <=10 else A[1]=1/(D+1) if HBA > 10 where, D=HBA-10.
6. Check the molecular weight (MW) of the molecule under consideration.
7. A[2]= 1, if MW <=500 or A[2]=1/(D+1) if MW >500 where, D=MW/500.
8. Check the partition coefficient (LogP) of the molecule under consideration.
9. A[3]= 1 , if LogP 5 or A[3]= 1/(D+1) if LogP > 5 where, D=LogP- 5.
10. OBA measure =$\sum_{i=0}$ A[i].

2) Interaction Energy

1. Initialise the variable E as zero.
2. E1=E_{vnd}
3. E2=E_{ele}
4. E=E1+E2

Here, Evnd is the van der waals interaction energy, Eele is the electrostatic interaction energy and E is the total interaction energy of the fragment. The two objectives are optimized using multi objective differential evolution algorithm.

V. RESULTS

RecA is a bacterial protein essential for the repair and maintenance of deoxyribonucleic acid (DNA). RecA has multiple functions related to DNA repair. It is responsible for the autocatalytic cleavage of the LexA repressor and the λ repressor in bacterial SOS response. It has a pivotal role in homologous recombination. This protein is essential for the existence of the pathogen and thus can be a drug target. It has several DNA binding sites and an ATP binding site. The ATP binding site of the target protein has been used for the present study. In the proposed method MODE is implemented to obtain the best solution in this problem it is molecule. In MODE the population size, number of generations, mutation and crossover probability are set to 100, 1000, .9, and .2, respectively. AMOSA is also implemented to compare the result obtained by applying MODE. The result shows that MODE performs better than AMOSA. The best result provided by MODE is shown in Figure 6 Table 2 shows the average maximum and best result of AMOSA and MODE.

Table 1: Average max. and Best of AMOSA and MODE

lgorithm	Average max.	Best
AMOSA	-8.6	-10.4
MODE	-9.1	-10.6

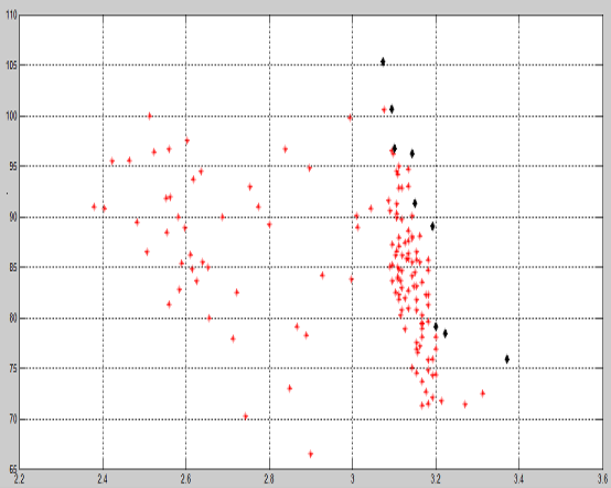

Figure 2: Best result of MODE

The black dots in the graph represent the set of non dominated solutions obtained by applying the algorithms.

CONCLUSION

TB is an infectious disease caused by the bacterium Mycobacterium Tuberculosis (MTB) that usually affects lungs. Compared with the other diseases caused by a single infectious agent, tuberculosis is the second biggest killer, globally. About two billion people are infected in the world and 2 million die each year from the disease. Traditional method of drug discovery is time consuming and costly. It may not produce effective drugs against TB. The computational techniques are incorporated to reduce the time taken for the drug discovery process. Many de novo design algorithms based on optimization techniques are developed earlier. The typical problems with such algorithms are that the molecules designed by them were often not synthesizable or drug like. In this proposed work, the problem is modeled using a multiobjective framework. MODE, a multiobjective optimization algorithm based on Differential Evolution, is used as the underlying optimization algorithm. So more desired result within less time, work and money.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to HOD and staff, Department of Computer Science and Engineering, St.Thomas College of Engineering and Technology for the constant encouragement throughout the duration of the work. A special gratitude I give to my guide Sivakumar K C, Information Officer, RGCB, Trivandrum, whose contribution in stimulating suggestions and encouragement helped me to complete this work.

REFERENCES

[1] Bandyopadhyay S, Saha S, Maulik U, and Deb K, "*A Simulated-Annealing Based Multiobjective Optimization Algorithm:AMOSA,"* IEEE Trans. Evolutionary Computing, vol. 12,no. 3, pp. 269-283, June 2008.

[2] Goha G, and Foster A*," Evolving Molecules for Drug Design Using Genetic Algorithms"* Proc. Int. Con. On Genetic and Evol. Computing, Morgan Kaufmann ,pp.27-33,2000.

[3] Bandyopadhyay S*,"Active Site Driven Ligand Design: An Evolutionary Approach",* Journal of Bioinformatics and Computational Biology Vol. 3, No. 5, pp. 1053–1070, 2005.

[4] Ecemis I¸ Wikel H, Bingham C, and Bonabeau E*, " A Drug Candidate Design Environment Using Evolutionary Computation",* Presented at IEEE Trans. Evolutionary Computation,Vol.12, pp. 591-603, October-2008.

[5] Ghosh A, Ghosh A, Chowdhury A, and Hazra J, **"***An Evolutionary Approach to Drug-Design Using Quantum Binary Particle Swarm Optimization Algorithm***"** ,in IEEE students' Conference on Electrical Electronics and Computer Science(SCEEC) pp.1-4, March-2012.

[6] Ghosh A, Ghosh A, Chowdhury A, and Konar A, "*An Evolutionary Approach to Drug-Design Using a Novel Neighbourhood Based Genetic Algorithm*", Journal .presented at CoRR, 2012.

[7] Roey V, Pereira B, Li Z, Hiraga K, Belfort M, and Derbyshire V, "*Crystallographic and Mutational Studies of Mycobacterium Tuberculosis Reca Mini-Inteins Suggest a Pivotal Role for a Highly Conserved Aspartate Residue,"* Journal Molecular Bilogy,vol. 367, pp. 162-73, 2007.

[8] Sengupta S and Bandyopadhyay S, "*De novo design of potential RecA inhibitors using Multi Objective Optimization*",IEEE/ACM Transactions on Computational Biology and Bioinformatics,Vol. 9,pp.1139-1154, July 2012.

[9] Bandyopadhyay S, Saha S, Maulik U, and Deb K, "*A Simulated-Annealing Based Multiobjective Optimization Algorithm:AMOSA,"* IEEE Trans. Evolutionary Computing, vol. 12,no. 3, pp. 269-283, June 2008.

[10] R. Storn and K. Price. "*Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces*",Journal Global Optimization,vol.11(4),pp.341- 359, 1997.

[11] S. Ekins, J.D. Honeycutt, and J.T. Metz, "*Evolving Molecules Using Multi-Objective Optimization: Applying to ADME/Tox,"* Journal Drug Discovery Today, vol. 15, pp. 451-460, 2010.

[12] Roey V, Pereira B, Li Z,Hiraga K, Belfort M, and Derbyshire V, "*Crystallographic and Mutational Studies of MycobacteriumTuberculosis Reca Mini-Inteins Suggest a Pivotal Role for a Highly Conserved Aspartate Residue*," Journal Molecular Bilogy,vol. 367, pp. 162-73, 2007.

[13] Little J, "*Autodigestion of lexA and phage lambda repressors*", Proc. National Academy of Sciences USA, vol. 81, no. 5, pp. 1375-1379,1984.

[14] Roca I and Cox M, "*The RecA Protein: Structure and Function",* Journal Biochemistry and Molecular Biology,vol.25, no. 6, pp.415-456,1990.

[15] Dey F and Caflisch A, "*Fragment-Based de Novo Ligand Design by Multiobjective Evolutionary Optimization*," Journal Chemical Information Modelling, vol. 48, pp. 679-690, 2008.

[16] Ferreira L, Recardo N,Santos D, Oliva G and Audricopulo A"*M*olecular Docking and Structure Based Drug Design

Strategies," Journal Molecules,Vol 20,pp. 13384-13421,July 2015.

[17] UmaDevi B,Sundar D,Alli P,"*Enhancement of the portfolio determination using Multiobjective Optimization",*IJCSE,vol 2,pp. 67-75, 2014.

Authors Profile

Mrs. Deepa Sreedhar pursed Bachelor of Technology from TKM Institute Of Technology, Karuvelil in 2009 and Master of Technology from SCT College of Engineering and Technology in year 2013.She is presently working as Assistant Professor in Department of Computer Science, St.Thomas College of Engineering and Technology.She has 4 years of teaching experience.

Sivakumar KC, is a computational biologist at Rajiv Gandhi Centre for Biotechnology, India. He is an expert in the area of computer-aided drug design, systems biology and molecular algorithms. He specialises in setting pipelines of Next Generation Sequence analysis. He has co-authored over 30 peer-reviewed manuscripts. His work has led to over 100 presentations at national and international scientific meetings

