



A GENETIC ALGORITHM BASED APPROACH FOR COMPARATIVE DOCKING ANALYSIS OF BREAST CANCER SUSCEPTIBILITY PROTEIN WITH CURCUMIN

V. PREMALATHA^{*1} AND P.VENKATESAN²

¹ Department of Mathematics, St. Joseph's College of Engineering, Chennai-119.

² National Institute for Research in Tuberculosis, Indian Council of Medical Research, Chetpet, Chennai-31.

ABSTRACT

In recent years, the demands on drug discovery process have increased dramatically, partly because of the necessity to recognize novel target that are both pertinent to disease and chemically tractable. The emergence of bioinformatics gives room to investigate diseases at the molecular level using computational techniques. Curcumin, is one of the biologically active phytochemical compounds which act a target, for treatment of number of diseases. The key protein breast cancer type 1 susceptibility protein (BRCA1) which upon mutation plays major roles in cyst formation and carriers have a 4-fold increased risk of colon cancer and increased chance of bilateral cancers (breast-ovarian cancer). In this study, docking scores and protein-ligand interactions are obtained using genetic algorithm, local search algorithm and simulated annealing algorithm for molecular docking with curcumin. Among them the docking scores and interactions obtained using genetic algorithm indicates curcumin as potential and natural therapeutic agent to combat breast cancer than the other algorithms. The efficiency of genetic algorithm relative to simulated annealing algorithm and local search algorithm is proved by docking cyclooxygenase 2 with aspirin and ibuprofen.

KEYWORDS: Genetic algorithm, Simulated annealing algorithm, Local search algorithm, Molecular Docking, Breast cancer, Curcumin.



V. PREMALATHA

Department of Mathematics, St. Joseph's College of
Engineering, Chennai- 119

**Corresponding author*

INTRODUCTION

Docking is a process by which one can predict the significant orientation of one molecule to a second when bound to each other to form a stable complex^{1,2}. Docking is mostly used for finding the binding between the ligand and the receptor. Hence in drug designing docking plays a vital role³. Between the two molecules, the binding affinities strength is predicted using the preferred orientation. For docking we require 3D structure of the protein and ligands as the input, for which the bound conformation of the ligand with that of the protein active site is predicted^{4,5}. We need a search algorithm which evolves new low energy conformations of the macromolecule with that of the micro among all the possible orientations. Genetic algorithm is one such algorithm. Large amount of ligands are there in the data bank, because of the exponential time complexity wet lab is not preferred first. Computer assisted drug design is preferred, and then go for wet lab only if there is good interaction between the ligands and the receptor. The docking job is done in dry lab using several commercial docking programs⁶. AutoDock is a suite of programmed docking tools. It is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure⁷. Local search algorithm, simulated annealing algorithm and genetic algorithms are some of the important algorithms used in docking analysis. Genetic algorithms are adaptive randomized search optimization technique based on the evolutionary ideas of natural selection with stochastic operators⁸⁻⁹. It was developed by John Holland along with his students and colleagues in 1970's, in an idea to import mechanisms of natural adaptation into computer system¹⁰. Genetic algorithm deals with pool of solution (i.e., population of strings) called the chromosomes. Genetic algorithm operates on encoded representation of the candidate solutions to an optimization problem, which is equivalent to those chromosomes of the individuals in nature. For evaluating each chromosome a fitness function should be

assigned, then the fittest chromosomes are selected to undergo the operations crossover and mutation to produce a new pool of solutions. The process will be continued till we get more highly fittest chromosomes in the pool of solutions. Other stopping criteria are number of generations, almost stable value of the best fitting individual and almost stable value of the average fitness of the population¹¹⁻¹². Autodock efforts to locate the best possible docked position of the small molecule inside a 3D grid box specified in the active site of the protein. Genetic algorithm works with a pool of solution where each individual in it is considered with its genetic code. In autodock, the particular position of the ligand with respect to the protein is treated as an individual in Genetic algorithm. The fitness of the individual is the total energy required for the particular orientation of the ligand inside the receptor. If the energy is lower, then it is the better one. The individual are selected according to their fitness value and they undergo crossover and mutation to produce a new generation. This process is continued until maximum number of evaluations or generations reached¹³.

Local algorithm is a meta heuristic method, applied for solving computationally hard optimization problem¹⁴. It starts with a candidate solution and iteratively moves to the neighboring solution. For every candidate solution there will be more than one neighborhood. So the choice of which neighboring solution to take into account depends only upon the information available about the solution. If it explores better solution in the neighborhood then it will replace the current solution. The process is continued until all better neighborhood solutions are exhausted. Simulated annealing is a probabilistic meta heuristic method used for solving optimization problems which has a large search space¹⁵. Iteratively it moves from one solution to another nearby solution, resulting in an optimum solution. In docking, if the new state's energy is lower than the previous one then the new state

is accepted instantly. But if the new state's energy is higher than the previous one then it is acknowledged probabilistically. The docking parameters for obtaining the final docked

structures, using genetic algorithm, local search algorithm and simulated annealing are shown in Fig. 1.a, Fig. 1.b and Fig. 1.c respectively.

Genetic Algorithm Parameters		Local Search Parameters		Simulated Annealing Parameters	
Number of GA Runs:	10	Number of LS Runs:	50	NUMBER OF:	
Population Size:	50	Maximum Number of iterations:	300	Runs:	10
Maximum Number of energy evaluations:	250000	Maximum Number of successes in a row before changing rho:	4	Cycles:	50
Maximum Number of generations:	27000	Maximum Number of failures in a row before changing rho:	4	Accepted steps/cycle:	100
Maximum Number of top individuals that automatically survive:	1	Soils&Wets parameter defining initial variance and size of local space to sample (rho):	1.0	Rejected steps/cycle:	100
Rate of Gene Mutation:	0.02	Lower bound on rho:	0.01	TO BEGIN NEXT CYCLE, USE:	
Rate of Crossover:	0.8	Probability of any particular phenotype being subjected to local search:	0.06	minimum state: <input type="text"/>	last state: <input type="text"/>
Mean of Cauchy distribution for gene mutation:	0.0	FOR LOCAL SEARCH, USE:		REDUCTION SCHEDULE TYPE:	
Variance of Cauchy distribution for gene mutation:	1.0	Soils & Wets with uniform variances: <input type="text"/>		<input checked="" type="radio"/> Linear	<input type="radio"/> Geometric
Number of generations for picking worst individual:	10	pseudo-Soils & Wets with relative variances: <input type="text"/>		REDUCTION FACTORS PER CYCLE:	
Accept	Close	Accept	Close	Translation: 1.0	Quaternion: 1.0
				Dihedral: 1.0	Temperature: 0.95
				Initial Temperature(Degrees): 1000.0	
				Accept	Close

Figure 1

Breast cancer is a collection of cancer cells which arises from the cells of the breast and it mostly occurs in women. The easiest way to understand how the inside of the breast is formed is by comparing it to an upturned bush¹⁶. Cancer cells that stay restricted to the lobule and the ducts are called 'non-invasive'. They are also occasionally referred to as pre-cancers in detection of the fact that these cells have not yet achieved the capacity to spread to other parts of the body, which is the characteristic that most people associate with cancer. If the cancer cells move outside the ducts and lobules into the surrounding breast tissue then it is said to be an invasive cancer. BRCA1 gene belongs to the class

of tumor suppressor genes. Like many other tumor suppressors, the protein produced from the BRCA1¹⁷, gene helps prevent cells from growing and dividing too rapidly or in an uncontrolled way. BRCA1 mutations alter single protein building blocks (amino acids) in the protein or remove large segments of DNA from the BRCA1 gene. Mutations consist of replacements, duplications, deletions or shortenings in the BRCA1 gene. This may result in proteins that may not function, work less effectively, are more quickly degraded, or present in inadequate numbers. Location of BRCA1 in chromosome 17 at position 21 from the base pair 41,196,311 to base pair 41,277,499 is given below¹⁸.

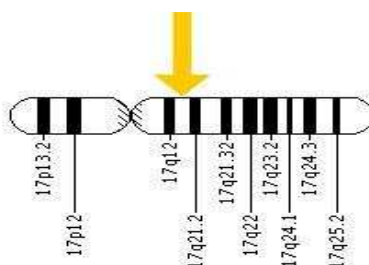


Figure 2

The BRCA1 gene is located on the long (q) arm of chromosome 17 at position 21.

Researchers have recognized above 1,000 mutations in the BRCA1 gene, many of which are associated with an increased jeopardy of cancer, particularly breast cancer in women. The majority of these mutations escorts to the production of a peculiarly short version of the BRCA1 protein, or prevent any protein from being made from one copy of the gene. Other BRCA1 mutations alter single protein building blocks in the protein or delete large sections of DNA from the BRCA1 gene. Researchers believe that a faulty or missing BRCA1 protein is not capable to help repair damaged DNA or fix mutations that take place in other genes. As these faults mount up, they can let cells to grow and divide uncontrollably and form a tumor leading to the risk of several other types of cancer including fallopian tube cancer, male breast cancer and pancreatic cancer. The mutated BRCA1 structure is already available in PDB and it was docked with curcumin using bioinformatics tools. In Asian and Chinese medicines, curcumin has been used as a dietary supplement as well as therapeutic agent¹⁹. Also it is employed as an anti-inflammatory representative for centuries in traditional Indian Ayurvedic medicine²⁰. Curcumin can adapt multiple signaling pathways and interact with many molecular targets²¹. Hence, it is believed to have the capability to act against a large number of cancers²². Curcumin induced DNA damage in triple negative breast cancer cells and regulated BRCA1 protein expression and modification²³. Rowe et al describes curcumin as the anti-cancer agent, which can also improve the survival of the patients with TNBCs by providing a non-toxic therapy. Curcumin in the future combinatorial therapy will give another dimension for remedy of cancer patients²⁴. In this work, the docking scores and H-bond formation defines the best interaction between the curcumin and the BRCA1. To prove genetic algorithms competency, we dock already well

documented protein- ligand pair, which is having abundant literature proof. So cyclooxygenase 2 - aspirin pair and cyclooxygenase 2 – ibuprofen pair is taken into account. The interactions formed in genetic algorithm are paramount when compared to other two algorithms.

MATERIALS AND METHODS

(i) Retrieval of Target Protein Sequence

The protein sequence of **breast cancer type 1 susceptibility protein**, was obtained from the protein sequence database of Uniprot (Accession No: P38398)²⁵. The 3D model of BRCA1 was already available in PDB structure was visualized in RASMOL²⁶ (**Fig. 3.a**). The three dimensional structure provides valuable insight in to molecular functions and also enables the analysis of its interactions with suitable inhibitors. The mutated three dimensional structure of breast cancer type 1 susceptibility protein is already available in PDB database (PDB ID: 1JNX)²⁷; hence an attempt has been made in the present study to determine the mutated structure BRCA1. The functional analysis of breast cancer type 1 susceptibility protein is predicted using Pfam Database²⁸. The functional regions of the breast cancer type 1 susceptibility protein is predicted and found to have single domain region -Zinc finger, C3HC4 type (RING finger). To further validate the efficacy of genetic algorithms in molecular docking, an attempt was made to dock aspirin and ibuprofen to cyclooxygenase 2. The protein sequence of cyclooxygenase 2 from *Mus musculus* (pdb id - 6COX) which is of identical length (604 residues) shares about 90% similarity with its human counterpart. The 3D model of cyclooxygenase 2 was already available in PDB structure was visualized in RASMOL (**Fig. 3.b**).



(a) The 3D structure of BRCA1



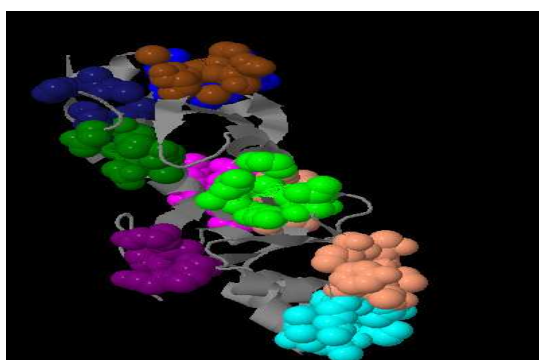
(b) The 3D structure of Cyclooxygenase 2

Figure 3

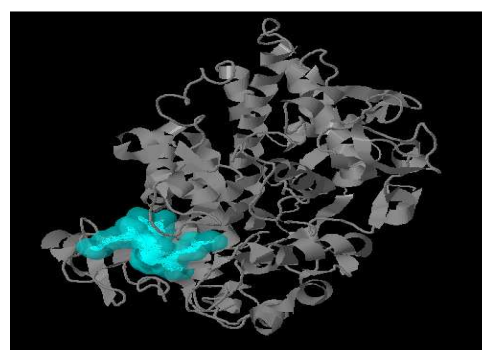
(ii) Active Site Prediction

The possible binding sites of breast cancer type 1 susceptibility protein were searched using Q-SiteFinder²⁹. Ten binding sites were obtained as active sites of breast cancer type 1 susceptibility protein. Among the ten binding sites obtained from Q-SiteFinder, site 1 is highly conserved. The residues in site 1 are YSX1667, PHEX1668, LYSX1671, HISX1672, SERX1722, ILEX1723, LYSX1724, ARGX1726, LYSX1727, METX1728. The residues forming the binding pocket are shown in Fig. 4.a.

The possible binding sites of cyclooxygenase 2 protein were searched using Q-SiteFinder. The residues forming the binding pocket are shown in Fig. 4.b. The residues are ASN34, CYS36, CYS37, ASN39, PRO40, CYS41, GLN42, ASN43, ARG44, GLY45, GLU46, CYS47, MET48, THR50, LYS56, ASP58, TYR122, LEU123, ASP125, THR129, TYR130, ASN131, VAL132, HIS133, TYR134, GLY135, TYR136, LYS137, ARG150, ALA151, LEU152, PRO153, PRO154, VAL155, ALA156, ASP157, CYS159, GLY164, GLN461, GLU465, LYS468, ARG469, PHE470, SER471.



(a) Active site prediction of BRCA 1 using Q-site finder



(b) Active site prediction of Cyclooxygenase 2 using Q-site finder

Figure 4

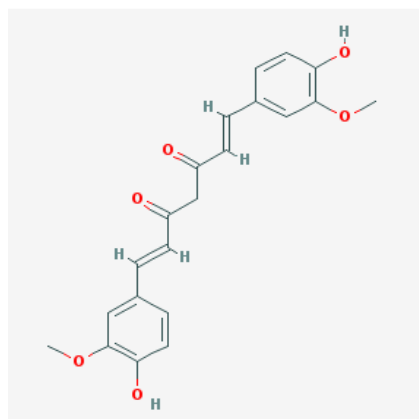
(iii) Ligand Preparation

The structure & properties of small molecule/inhibitor (curcumin) has been obtained from Pubchem Database^{30,31}. If the small molecule violates more than one of the Lipinski's rule of five, then the problem of bioavailability arises. The properties of the small molecule can be determined by

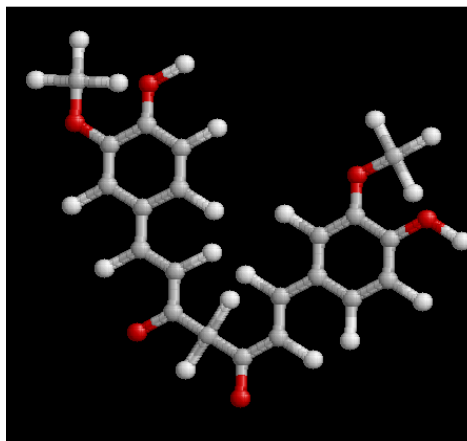
Molinspiration property calculator tool³². Curcumin's molecular weight is 368.3799(g/mol), XLogP3-AA is 3.2, the number of H-Bond donor is 2 and the number of H-Bond acceptor is 6. As curcumin abided by Lipinski's rule of five³³, it has a high probability of drug likeliness. The 2D structure of small molecule obtained from the Pubchem

Database and the 3D structure visualized using Rasmol tool was shown in Fig. 5.a and Fig. 5.b. In autodock, we need both ligand and the protein in PDB format. So the structure

downloaded in SDF format and was then converted to PDB format using OPEN BABEL 2.3.1³⁴.



(a) 2D structure of Curcumin

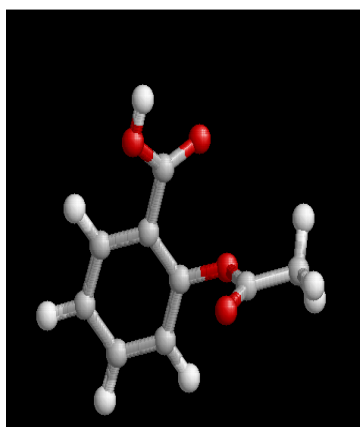


(b) 3D structure of Curcumin

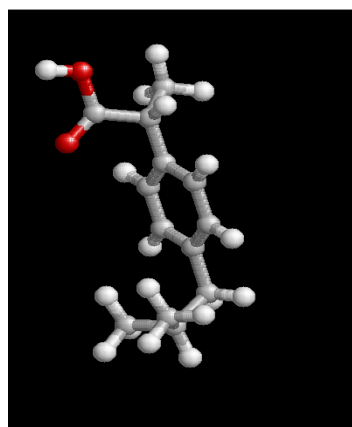
Figure 5

The analgesic, antipyretic and anti-inflammatory results of acetylsalicylic acid (aspirin) are due to actions by both the acetyl and the salicylate portions of the intact molecule as well as by the active salicylate metabolite. Aspirin directly and irreversibly inhibits the activity of both types of cyclooxygenase (COX-1 and COX-2) to decrease the formation of precursors of

prostaglandins and thromboxanes from arachidonic acid³⁵. Like aspirin, ibuprofen is a nonselective COX inhibitor, in that it inhibits two isoforms of cyclooxygenase, COX-1 and COX-2³⁶. The structure of aspirin and ibuprofen were retrieved from pubchem database and were converted into pdb format by using open babel. The 3D structure of aspirin and ibuprofen were shown in Fig. 6.a and Fig. 6.b.



(a) 3D structure of Aspirin



(b) 3D structure of Ibuprofen

Figure 6

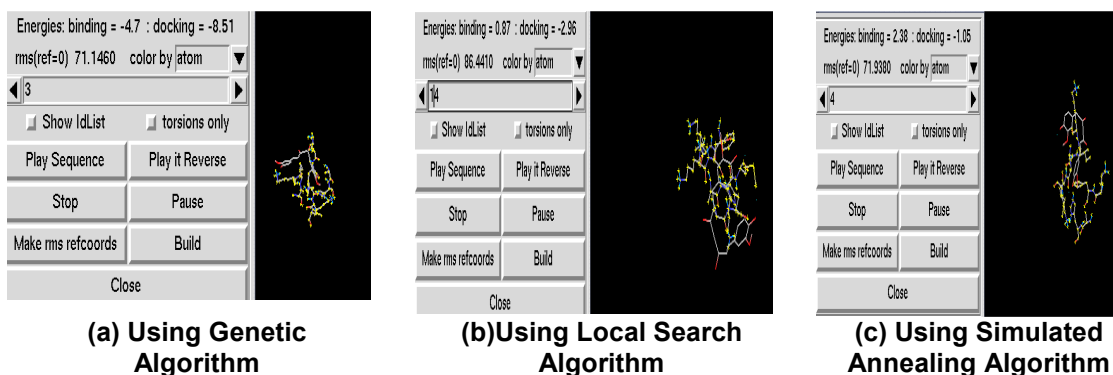
(iv) Docking of Proteins with Potential Inhibitors using three different algorithms

The curcumin have aroused considerable interest recently because of its potential beneficial effects on human health. It has been reported to have antiviral, antiallergic, antiplatelet and anti-inflammatory effect. Curcumin is docked with BRCA1 using the genetic algorithm, local search algorithm and simulated annealing algorithm. Also Cyclooxygenase 2 is docked with patented drugs like Aspirin and Ibuprofen.

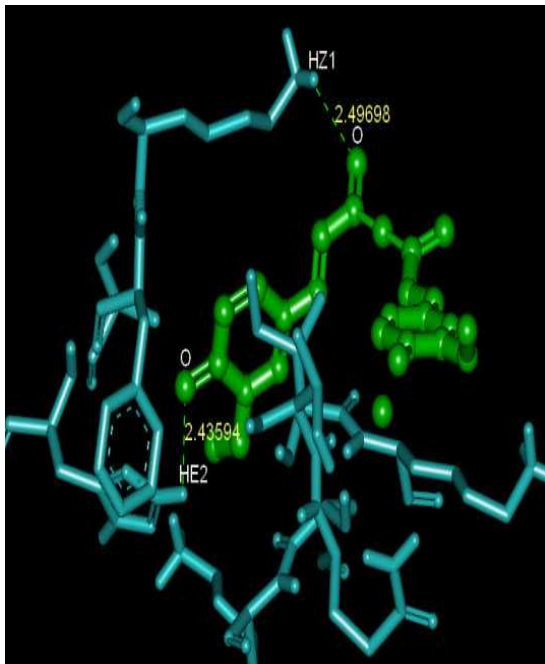
RESULTS AND DISCUSSION**(i) Final Docked Conformations of BRCA1 with Curcumin using different Algorithms**

Docking of breast cancer type 1 susceptibility protein was performed with curcumin inhibitor using three different algorithms and the final docked conformations (Fig. 7.a, 7.b & 7.c) obtained for the inhibitor in 3 algorithms were evaluated based on the number of hydrogen bonds formed. Based on the lowest energy and minimal solvent accessibility of the ligand, the most excellent ligand-protein complex was chosen from the docked structures for the 3 algorithms. The final conformations are visualised using Accelrys discovery studio visualisation tool are shown in Fig. 8.(a,b,c,d,e) and compared in table1 for the three algorithms.

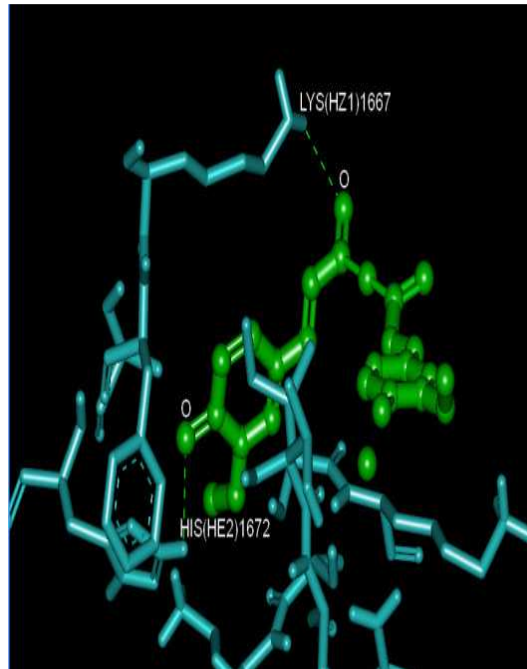
Figure 7
Docking Score of BRCA1 against Curcumin



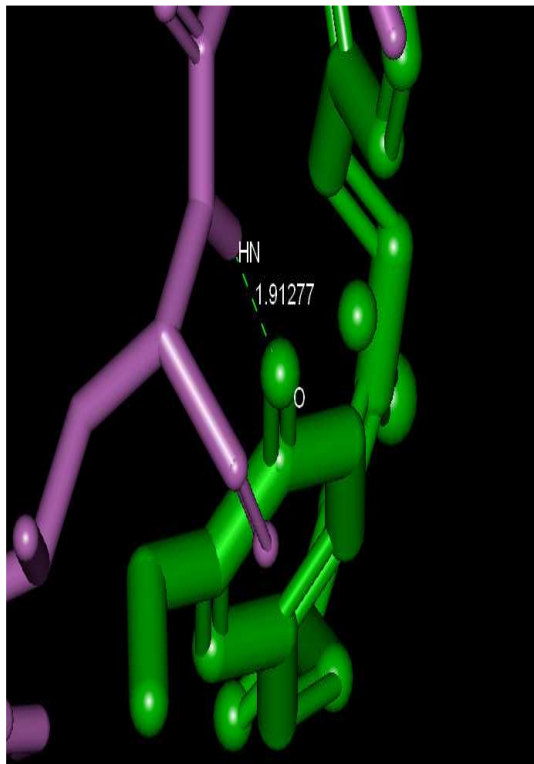
Docking conformations of BRCA1 against Curcumin



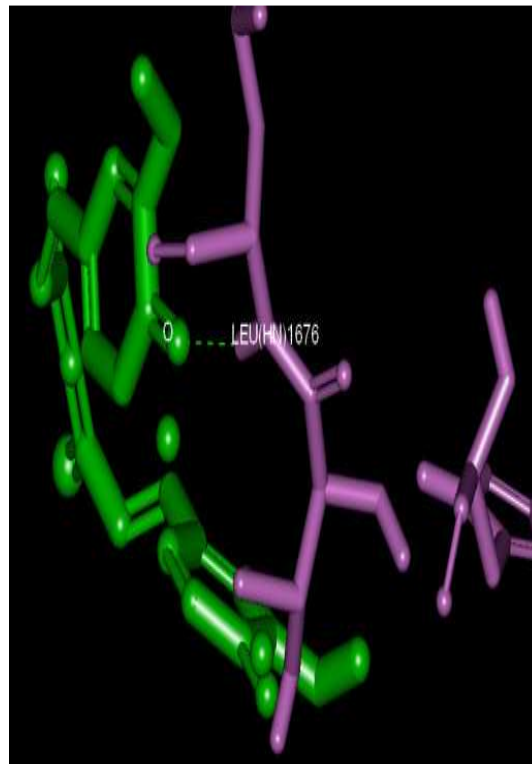
(a) H-Bond distance



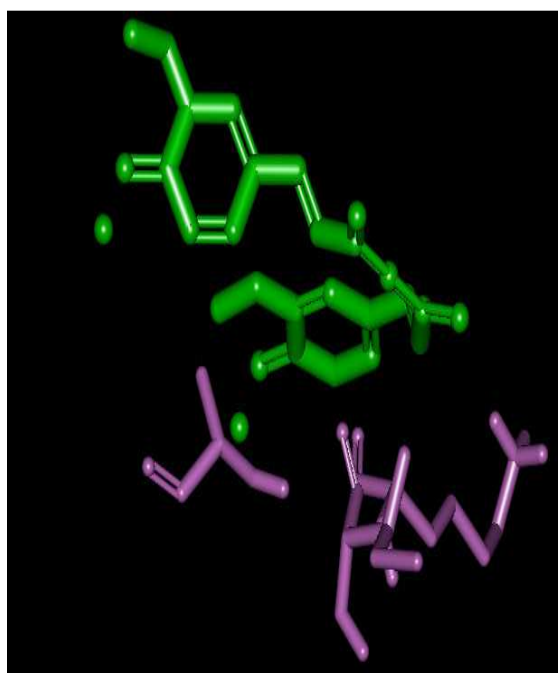
(b) Residue attributes



(c) H-Bond distance



(d) Residue attributes



(e) No Interaction

Figure 8

(a),(b) Docked conformation with its attributes using Genetic Algorithm,
 (c),(d) Docked conformation with its attributes using Local Search Algorithm,
 (e) Docked conformation obtained in Simulated Annealing Algorithm

Table 1**Comparison of Docked interactions using Algorithms**

Algorithms	Docked Interactions					
	BRCA1		Curcumin	Distance (Å ⁰)	H-BONDS	Docked energy (kcal/mol)
	Residue	Atom				
Genetic algorithms	HIS1672	HE2	O	2.43594	2	-8.51
	LYS1667	HZ1	O	2.49698		
Local search algorithm	LEU1676	HN	O	1.91277	1	-2.96
Simulated annealing algorithm	-	-	-	-	0	-

Breast cancers with BRCA1 illustrate a high incidence of chromosomal irregularities. In this work, 3D model of BRCA1 was already available in PDB structure was mutated and the structure was visualized in RASMOL. The BRCA1 protein has Zinc finger domain from 22-66 was predicted using PFam. Ten active site regions of BRCA1 were analyzed using Q-

Sitefinder. Among them, site 1 was considered to be the best. Using autodock, the docking between the macromolecule and the ligand was conducted. The docking scores and H-bond formation defines the best interaction between Curcumin and BRCA1. The interaction formed in genetic algorithm having 2 H-Bonds and energy -8.51 kcal/mol. For

local search algorithm the interactions formed is having 1 H-Bond and energy -2.96 kcal/mol. But there is no interaction formed in simulated annealing algorithm.

(ii) Final Docked Conformations of Cyclooxygenase 2 with Aspirin and Ibuprofen using different Algorithms

Docking of cyclooxygenase 2 protein was performed with inhibitors aspirin and ibuprofen

using three different algorithms. Based on the lowest energy and minimal solvent accessibility of the ligand, the most excellent ligand-protein complex was chosen from the docked structures for the 3 algorithms. The final conformations of Cyclooxygenase 2 with Aspirin and Cyclooxygenase 2 with Ibuprofen are visualised using Accelrys discovery studio visualisation tool are shown in Fig. 9.(a,b,c) and Fig. 10.(a,b,c) respectively.

Docking conformations of Cyclooxygenase 2 against Aspirin

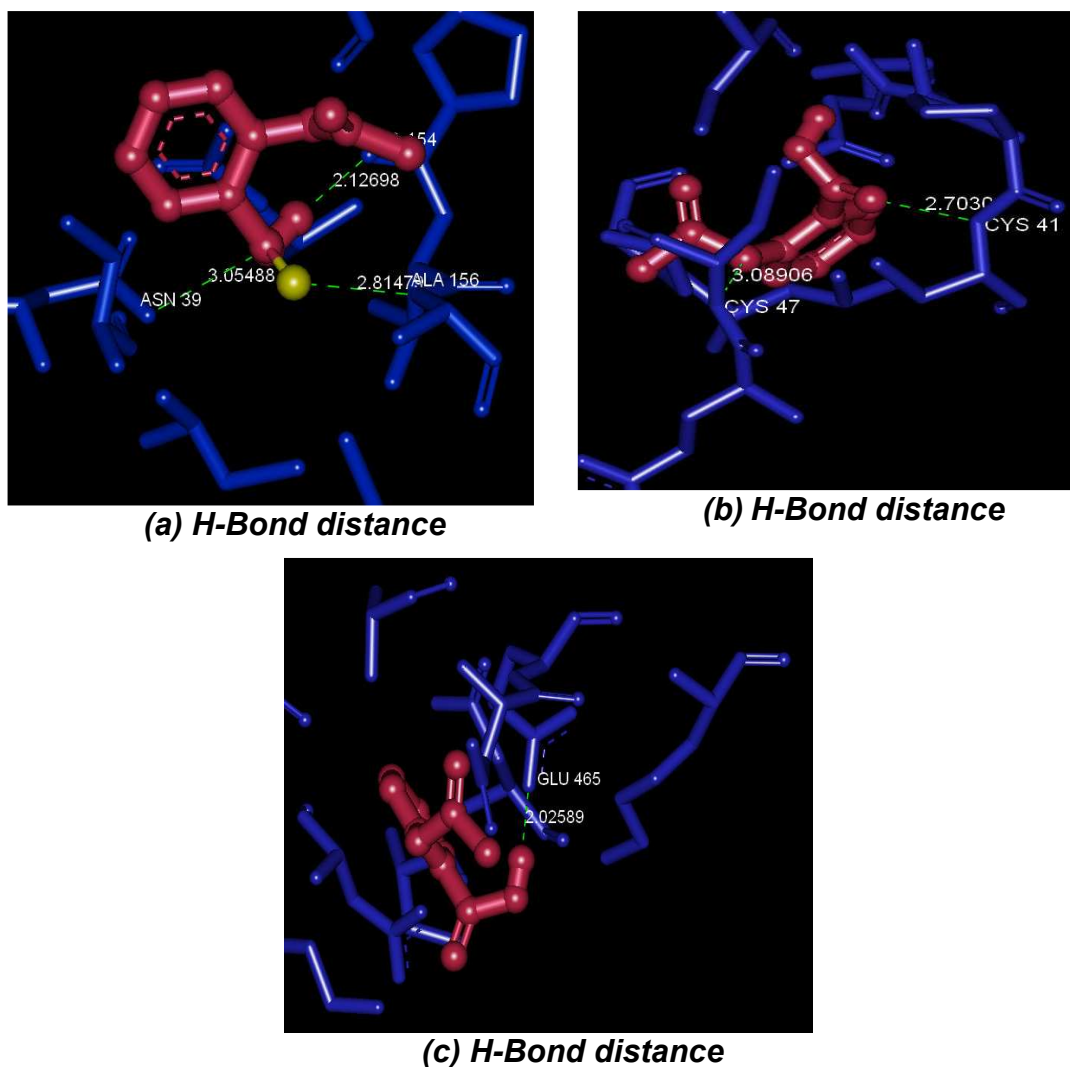


Figure 9
Docked conformation with its attributes using (a) Genetic algorithms (b) Simulated annealing algorithm (c) Local search algorithm

Docking conformations of Cyclooxygenase 2 against Ibuprofen

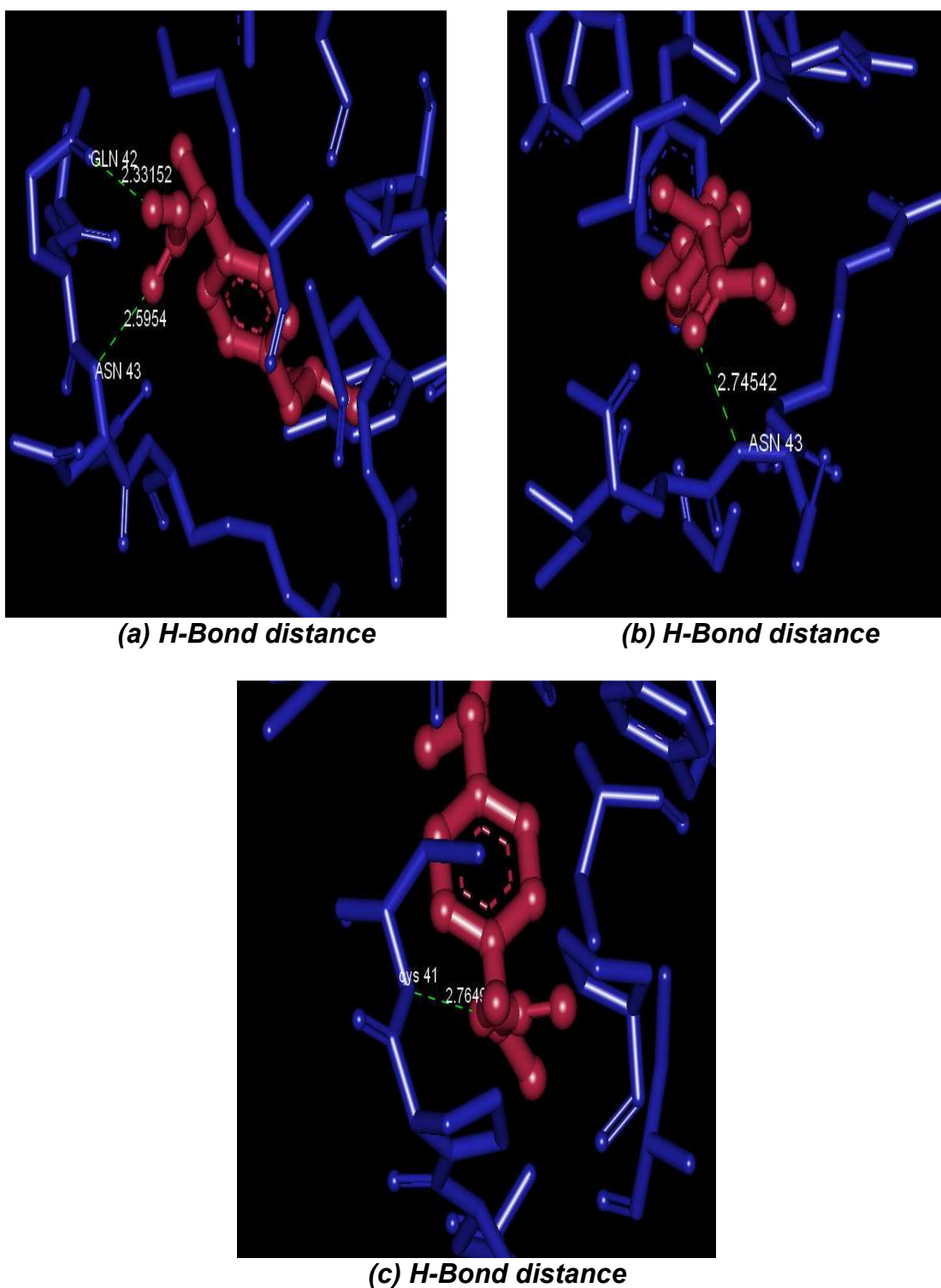


Figure 10
Docked conformation with its attributes using (a) Genetic algorithms (b) Simulated annealing algorithm (c) Local search algorithm

Table 2
Comparison of Docking Scores and Bonds for the three algorithms

MACRO MOLECULE	INHIBITOR MOLECULE	Genetic Algorithm		Stimulated Annealing Algorithm		Local Search Algorithm	
		Energy	H-Bond	Energy	H-Bond	Energy	H-Bond
BRCA1	Curcumin	-8.51	2	-1.05	0	-2.96	1
Cyclooxygenase2	Aspirin	-6.45	3	-6.17	2	-5.58	1
	Ibuprofen	-7.98	2	-7.74	1	-7.22	1

Based on the above results shown in Table 2, it has been proved that the interactions formed using genetic algorithm shows better result than the other two algorithms.

CONCLUSION

The interactions between BRCA1 and the natural inhibitor curcumin, Insilico studies are useful to understand the underlying mechanism of enzyme inhibition. Favorable results of docking analysis shows curcumin as best therapeutic drug. Docking scores results indicate the application of Curcumin as Potential and Natural Therapeutic agents to treat Breast Cancer. Among the three

algorithms the interactions formed in genetic algorithm shows the best result when compared to other two algorithms. Also the efficacy of genetic algorithm is authenticated in molecular docking by an attempt to dock aspirin and ibuprofen to cyclooxygenase 2. Further studies are needed to confirm these findings.

REFERENCES

1. Lengauer T, Rarey M, Computational methods for biomolecular docking. Curr. Opin. Struct. Biol, 6 (3): 402–6, (1996).
2. Jorgensen WL, Rusting of the lock and key model for protein-ligand binding. Science, 254 (5034): 954–5, (1991).
3. Kitchen DB, Decornez H, Furr JR, Bajorath J, Docking and scoring in virtual screening for drug discovery: methods and applications. Nature reviews/Drug discovery 3(11): 935–49, (2004).
4. Venkatachalam CM, Jiang X, Oldfield T, Waldman M, LigandFit: A novel method for the shape-directed rapid docking of ligands to protein active sites. J Mol Graph Model, 21 (4) : 289-307, (2003).
5. Rarey M, Kramer B, Lengauer T, Time-efficient docking of flexible ligands into active sites of proteins. Proc Int Conf Intell Syst Mol Biol, 3 : 300-8, (1995).
6. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS, Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem, 47(7) :1739-49, (2004).
7. <http://autodock.scripps.edu/downloads> accessed on 5.04.12..
8. Holland J, Adaptation in Natural and Artificial Systems. The University of Michigan, (1975).
9. Goldberg DE, Genetic algorithms in Search, Optimization and Machine Learning.

- Reading, MA: Addison-Wesley, (1989).
10. Davis L, Handbook of genetic algorithm. San Nostrand Reinhold, New York, (1991).
 11. Mitchell M, An introduction to genetic algorithms, the MIT Press, Cambridge, Massachusetts, (1997).
 12. D.Beasley, D.R. Bull, and R.R. Martin, An overview of genetic algorithms: Part 1- Fundamentals. Univ. comput., 15(2): 58-69, (1993).
 13. Michael. D. Vose, The Simple Genetic Algorithm-Foundations and theory. MIT Press, (1999).
 14. Battiti Roberto, Mauro Brunato, Franco Mascia, Reactive Search and Intelligent Optimization. Springer Verlag, (2008).
 15. Kirkpatrick. S, Gelatt. C.D, Vecchis. M.P, Optimization by Simulated Annealing. Science 220 (4598): 671– 680, (1983).
 16. American Cancer Society (2009). What are the risk factors for breast cancer? Detailed guide: Breastcancer. Available online: http://www.cancer.org/docroot/CRI/content/CRI_2_4_2X_What_are_the_risk_factors_for_breast_cancer_5.asp, accessed on 5.04.12.
 17. National Cancer Institute (2006). Probability of breast cancer in American women. National Cancer Institute Fact Sheet. Available online: <http://www.cancer.gov/cancertopics/factsheet/Detection/probability-breast-cancer>, accessed on 5.04.12..
 18. <http://ghr.nlm.nih.gov/gene/BRCA1> accessed on 26.04.12.
 19. Aggarwal.B.B, Sundaram.C, Malani.N, Ichikawa.H, Curcumin: the Indian solid gold. Adv Exp Med Biol, 595:1-75, (2007).
 20. Strimpakos AS, Sharma RA, Curcumin: preventive and therapeutic properties in laboratory studies and clinical trials. Antioxid Redox Signal, 10:511–45, (2008).
 21. Itokawa H, Shi Q, Akiyama T, Morris-Natschke SL, Lee KH, Recent advances in the investigation of curcuminoids. Chin Med, 3:11, (2008).
 22. Preetha Anand, Chitra Sundaram, Sonia Jhurani, Ajaikumar B. Kunnumakkara, Bharat B. Aggarwal, Curcumin and cancer: An “old-age” disease with an “age-old” solution. Cancer Letters-267, 133–164, (2008).
 23. Rowe. DL, Ozbay. T, O’ Regan.RM, Nahta R, Modulation of the BRCA1 Protein and Induction of Apoptosis in Triple Negative Breast Cancer Cell Lines by the Polyphenolic Compound Curcumin. Breast Cancer: Basic and Clinical Research.61-75, (2009).
 24. Zhenyu Ji, Testing DNA Damage and repair by Curcumin. Breast Cancer: Basic and Clinical Research. 1-3, (2010).
 25. <http://www.uniprot.org/uniprot/P38398> accessed on 26.04.12.
 26. Herbert J. Bernstein, Recent changes to RasMol, recombining the variants. Trends in Biochemical Sciences (TIBS), 25 (9): 453-455, (2000).
 27. Williams RS, Green R, Glover JN, Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nat Struct Biol., 8(10), 838-42, (2001).
 28. <http://pfam.sanger.ac.uk/>, accessed on 26.04.12.
 29. <http://www.modelling.leeds.ac.uk/qsitefinder/>, accessed on 26.04.12.
 30. <http://pubchem.ncbi.nlm.nih.gov/>, accessed on 26.04.12.
 31. Bolton EE, Wang Y, Thiessen PA, Bryant SH. PubChem: Integrated platform of small molecules and biological activities. Annu. Rep. Comput. Chem. 4:217–241. Chapter 12, (2008).
 32. <http://www.molinspiration.com/cgi-bin/properties> accessed on 26.04.12.
 33. Lipinski. C.A, Lombardo. F, Dominy. B.W, Feeney. P.J, Experimental and computational

approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv Rev*, 46(1-3): 3-26, (2001).

34. Werner J, Geldenhuys, Kevin E.Gaasch, Mark Watson, David D.Allen and Cornelis J.Vander Schyf, Optimizing the use of opensource software applications in drug discovery, *Drug Discovery Today*. 11(3-4)127-132 (2006).
35. <http://www.drugbank.ca/drugs/DB00945> accessed on 18.09.12.
36. <http://www.drugbank.ca/drugs/DB01050> accessed on 18.09.12.