Available online at www.scholarsresearchlibrary.com

Scholars Research Library

J. Comput. Method. Mol. Design, 2011, 1 (3):1-10
(http://scholarsresearchlibrary.com/archive.html)

ISSN : 2231- 3176
CODEN (USA): JCMMDA

Accurate prediction of breast cancer gene BRCA1 domain ligands with potential drugs via Homology modeling

A.G. Murugesan*, S. Sasi premila and K. Bala Amutha

Manonmaniam Sundaranar University, Sri Paramakalyani Centre of Excellence in Environmental Sciences, Alwarkurichi - 627 412. Tamil Nadu, India

---

ABSTRACT

Breast cancer is the most common malignancy in women. The role of BRCA1 genes which modulate or promote cancer has been fully understood. The current status in rational drug design using homology modeling is discussed in this study which focuses on template selection, model building, model verification and strategies for drug design based on the model structures. A novel approach to identify the unique binding site for selected drugs like Anastrozole, Extremestane, Epirubicin and letrozole, based on homology and docking is described.

Key words: Breast cancer, BRCA1 genes, homology modeling, and interacting drugs

---

INTRODUCTION

Breast cancer is one of the most frequently diagnosed malignancies and is the second leading causes for women death. Wide ranges of carcinogens are responsible for carcinogenicity [1, 2]. Radiation is well documented risk factor for breast cancer and its exposure induces the formation of free radicals [2]. Several improvements in diagnostic protocols enhanced the ability for earlier breast cancer detection with improved therapeutic outcome and survival rate. Breast cancer type 1 susceptibility protein (BRCA1) has been identified as a novel marker for early cancer detection. The multifactorial BRCA1 gene product is involved in DNA repair mechanism, ubiquitination, transcriptional regulation and other functions. Women with an altered BRCA1 or BRCA2 gene are 3 to 7 times more likely to develop breast cancer than other women [3].

Nowadays, rational drug design is an important concept in pharmaceutical research. The goal is to identify a key drug target based on meticulous understanding of regulatory networks and metabolic pathways, and design specific drug target based on known three dimensional (3D)

structures. The large genomic data oriented projects brought the concept close to reality [4]. The detailed genomic sequence mapping, regulatory networks and metabolic pathways combined with single nucleotide polymorphism (SNP) data made easy to identify optimal drug targets. Access to high quality 3D structures with the drug target is a good starting point for rational drug design[4]. Several examples are available for rational drug target design with known 3D structures for HIV protease inhibitor amprenavir (Agenerase) and nelfinavir (Viracept) [5.6] and the influenza virus inhibitor zanaminivir (Relenza). Structure based drug design was applied to design the protein kinases inhibitors [7] such as Abl kinase [8], CDKs, EGFR kinase[9], Lck [10] and Src [11].

X-ray crystallography is the main method for structural determination of proteins. Structural domains of proteins are classified into classes of similar folds [12]. This made homology based protein model as the alternative in experimental structural determination. Homology modeling was used in several drug design projects. Enyedy [13] utilized the Bcl2 homology model to identify a novel inhibitor. Furet [14] applied homology modeling for rational design of Cyclin dependent kinase 1 inhibitors (CDK1).

The discovery of three novel ligand candidates [3] and homology modeling of Falcipain-2 provided information which leads to discover new drugs against malaria [15]. Recently, Schafferhans and Klebe [16] published a method for computational ligands docking with the protein binding sites. The homology modeling of BRCA1 genes with specific peptide binding structures, computational docking and recently developed free energy estimation protocols exhibited the interactions of drugs used in treatments and BRCA1. With the development of 3D modeling maps, pharmacokinetic studies were performed to locate the BRCA1 gene and associated undefined adverse interactions [17] Therefore, this attempt has been made to collect information and analyze the drug interaction in cancer therapy treatment.

MATERIALS AND METHODS

2.1. Selection of drugs

In this present study, anastrozole, extremestane, epirubicin and letrozole were used for homology modeling using the predicted BRCA1 gene. Anastrozole is a potent, selective non steroidal inhibitor. Epirubicin is the drug to eliminate cancer cells. Extremestane is an irreversible, steroidal activator for suicide inhibition. Letrozole is approved for early breast cancer with little side effects.

2.2. BRCA1 biomarker protein

The *BRCA1* gene is located on the long (q) arm of chromosome 17 at band 21, from base pair 38,449,843 to base pair 38,530,933 [18]. The genes BRCA1 and BRCA2 control the development of two protein compounds which suppress the tumors growth. These genes interfere with the protective action, diminishing the body's ability to defend the overgrowth of cancerous cells.

2.3. Computational methods

The methodology was described in four steps: Identifying a suitable template, making an optimal target template alignment, building the model and validating the model. Protein structure prediction and homology modeling has recently been reviewed [17].

2.4. Identifying suitable template and sequence alignment

Different online tools were used to select the protein of interest (BRCA1) and determine structures as the target 3D structure. Heuristic search methods such as BLAST [19] and FASTA [20] were used to find initial template (BRCA1 protein). HEX was an alternative protein locking program used for correlating the docking calculations. PIR database was searched by PIR-PSD using the profile generated from PIR database which exhibited similarity in unknown structure. The alignment of protein sequence was constructed by using CLUSTAL-X. The BRCA1 domain structure was submitted in project mode to SWISS model server.

2.5. Model building, evaluation and validation of model

The crystallographic structure of BRCA1 sequence was obtained from the protein database files. Structure was constructed using MODELLAR software which builds the model based on the satisfaction of spatial restraints [21]. This model was refined using SWISS PDB VIEWER [22]. Energy minimization of the representative model was evaluated by means of GROMOS 96 algorithm. The 3D structure validation mainly deals with the target protein, its energy value, overall quality factors and Ramachandran plot etc. The active site or domain region of the molecule was predicted by Castp server (www.sts-fw.bioengr.uic.edu/castp/calculation.php).

2.6. Docking with drugs

Docking plays an important role in drugs designing and docking process was carried out with the software Mol Dock. The specificity of the drug and the target protein depend on the binding sites. Using the comparative study, drug specific interaction were compared and predicted between the available drugs. The specific protein docking with specific receptors of cancer drugs were constructed with specific drug in 3D model.

RESULTS AND DISCUSSION

BRCA1 and BRCA2 are tumor suppresser genes involved in signaling and DNA repairs. Mutations in BRCA1 genes are responsible for 45% inherited breast cancer, 80% inherited breast and ovarian cancer [23]. In the present study, 3D homology modeling was carried out for the breast cancer type I susceptibility protein (BRCA1). The docking interactions were calculated for some of the drugs like anastrozole, epirubicin, exemestane and letrozole. The designed analogues would provide solutions for effective drug discovery. The primary 3D structure of the target BRCA1 human protein sequence was selected from the NCBI and loaded as the raw sequence. The loaded model was viewed in SWISS PDB viewer (Table. 1). The LOOK software package [18] and Segment Match Modeling (SegMod) were used to generate homology models through fragment based assembly [24]. The threading of target BRCA1 sequence with template sequence was performed (Fig. 1).

Fig 1 Screen shot showing the superimposed similarities between the alignments of vertical identities

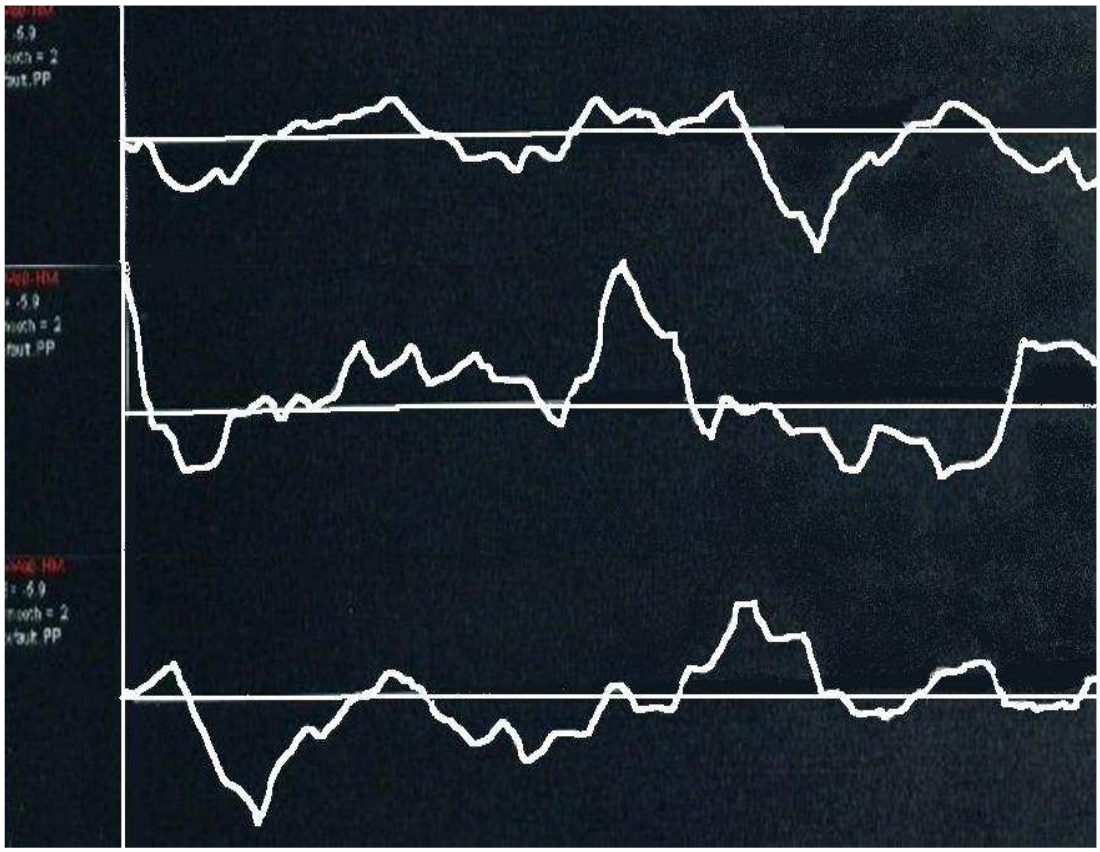
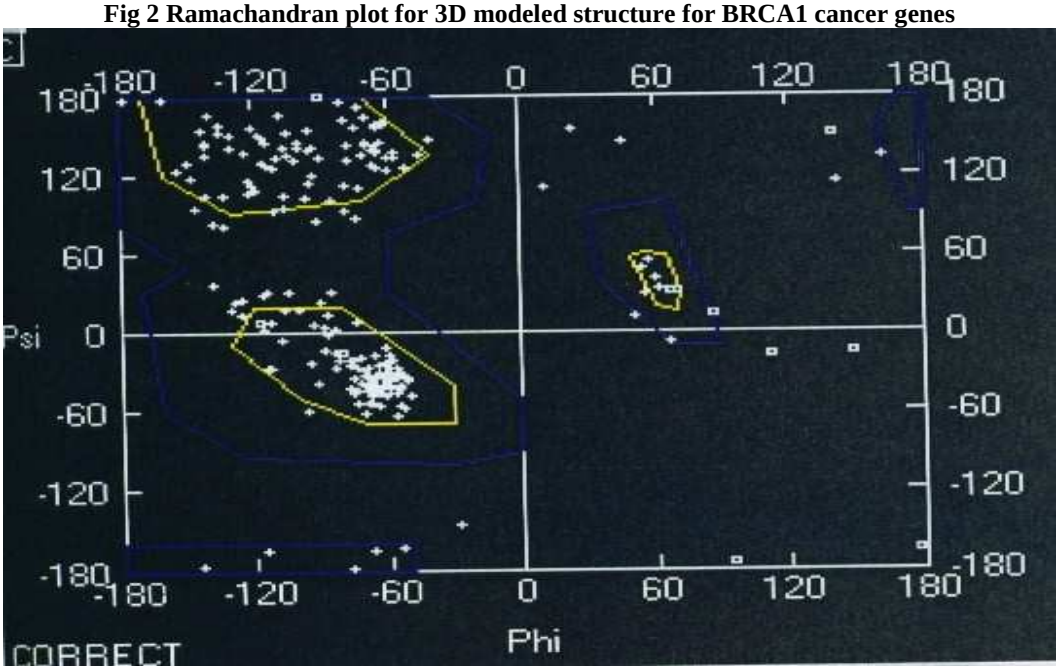


Fig 2 Ramachandran plot for 3D modeled structure for BRCA1 cancer genes



Homology modeling fetched 16 hits, out of 5 close homologous sequences. The sequence with low E values was selected from different homologous gene sequences. Homology modeling of the query protein was carried out using Rat BRCA1 Tandem crystal structure called BRCT region structure with 65% identity (Fig. 2). The C alpha backbones of the modeled query protein were super imposed with the template sequence and RMSD was calculated as 0.85 A°. The protein model built in this study was ribbon shaped which retained the general helical structure (Table. 2).

Schafferhans and Klebe [16] used gaussian functions to represent the physico chemical properties of the receptor and ligand. According to thermodynamic hypothesis, native protein conformations corresponded well to the global minima with the free energy 20. Energy of minimization value showed that the protein was well defined. Models with >50% sequence identity were in high quality, with ~1 Å root mean square (RMS) error for the main chain atoms (equal to medium-resolution NMR or low resolution X-ray structures). Models which have 30 – 50% sequence identity were in medium accuracy with RMS ~1.5 Å [25]. The hits from the database were evaluated further by molecular docking. The drug ligands and analogs were very specific for BRCA1 domain. It was noticed that the drug analog binds to the regions in adjacent with the BRCA1 domain sequence. A total of 44 ligand binding sites were determined of which the one with maximum area (1027.2) and volume (1325.3) was selected for neuraminidase protein of H1N1 sub type gene [28].

Ramachandran plot was used to visualize the glycine and all other residues which were placed in the allowed regions. The possible confirmation of the Phi ad Psi angles for the polypeptide was obtained. PROCHECK analysis revealed the residues within the limits of Ramachandran plot. Fig. 2 concluded that there were no disabled amino acids and it was considered as a good model. Loop building with SWISS PROTEIN DATABANK and energy minimization resulted in high reduced energy. SAVS structure analysis verified the BRCA1 protein identity as 100 and 90% (Fig.3). The scoring function indicated the likelihood which represented the favorable binding interaction [26]. The low energy indicated the stable and likely binding interaction [27]. The problematic residues were denoted in different colors like yellow, green and red respectively. Considering all these, this model was proved to be a valid model with overall quality 84.729 which supported the study of protein docking.

The drug ligands were collected for anastrozole, epirubicin, exemestane and letrozole and their analogues were also collected from KEGG database and their interactions were calculated from the distance of amino acids towards the modeled protein (Table. 2). The interactions of the drugs and its analogue confirmed the effective interaction between the chosen drugs. The RMS score value is -1.00. Docking analysis reveals that three water molecules directly affect the interaction of Zanamivir with neuraminidase [29].

According to Brinda [26], the distance between the target and drug ligands with less than 2.00 Armstrong showed high efficiency. Thus this study showed the interactions between the drugs which were found to be quite good (Fig. 4 & Fig. 5). The result obtained in this study might pave way to identify the potent drug target and suitable drugs for human. This modeled structural analysis could improve the present cancer therapy. The docking result showed that the drug chosen for the present study has high specificity and efficiency towards the target BRCA1 protein which causes tumor.

Fig 3 Ribbon model of BRCA1 gene depicting helices and strands in 3D structure

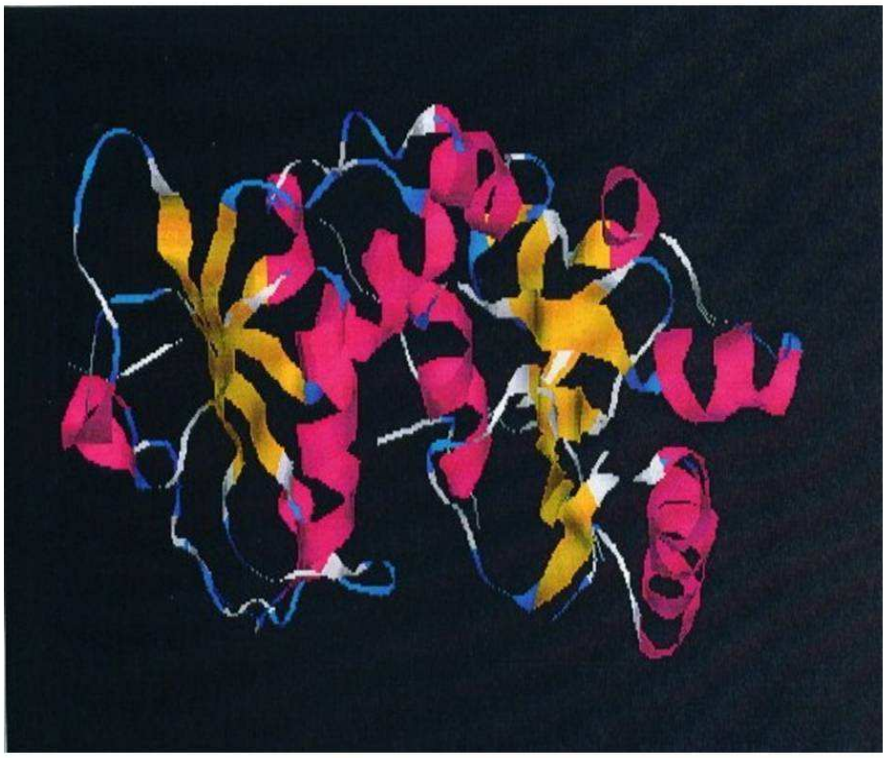
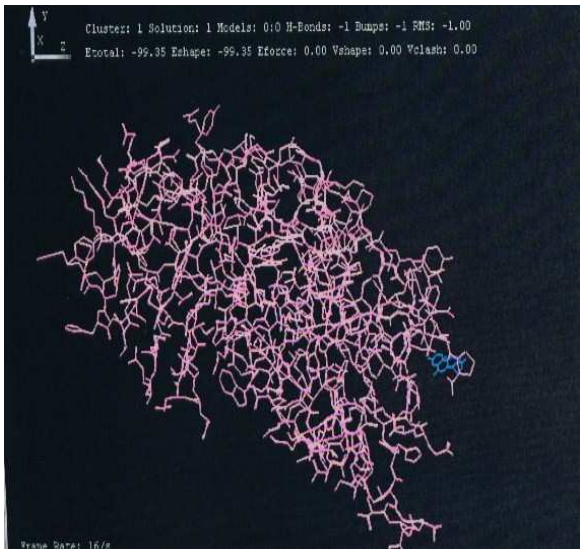
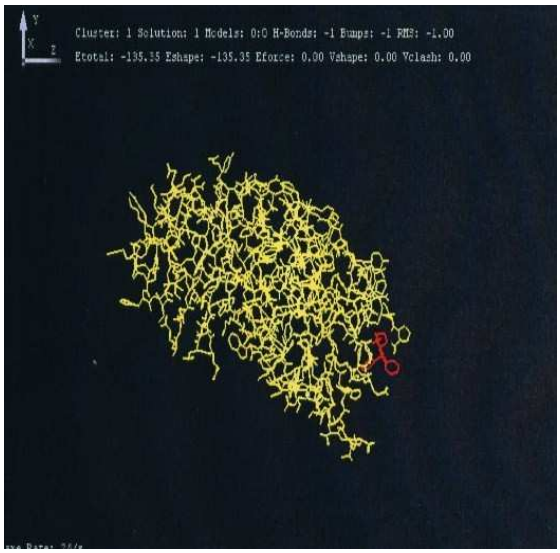
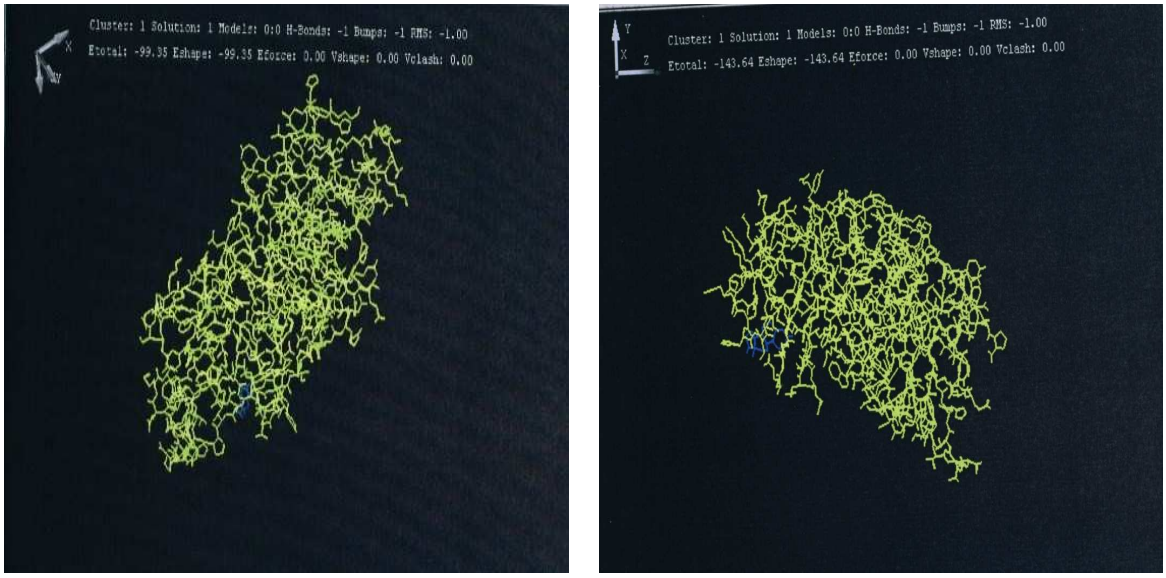


Fig 4 Screen shot showing the homology modeling of amino acid and drugs. A). Anastrozole b) Extremestane, c). Epirubicin d). Letrozole



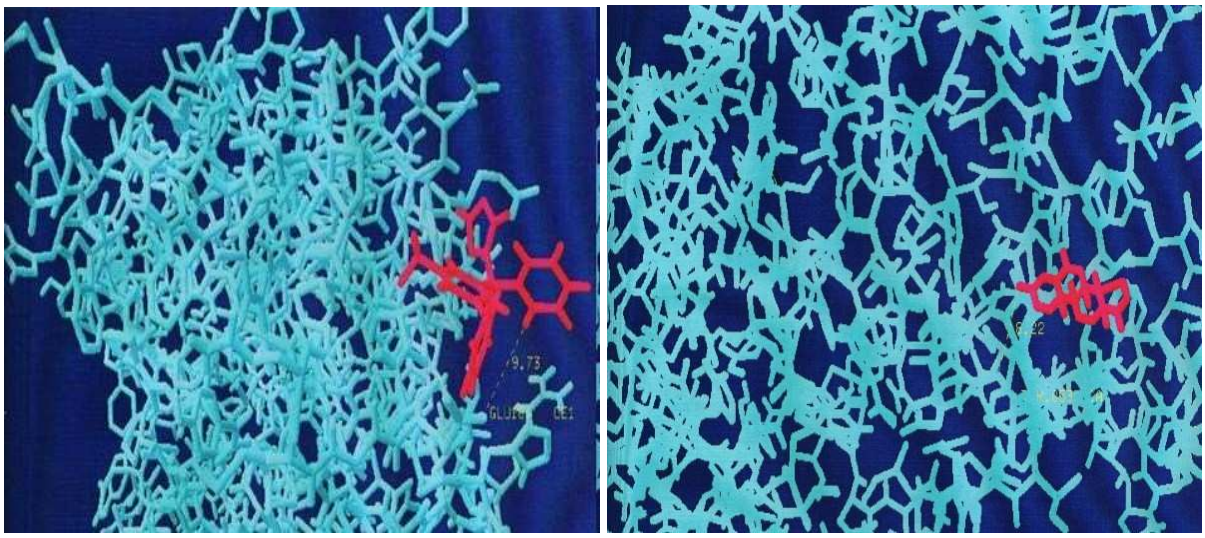




Clusters: 1 solution: 1 Models:0:0 H- Bonds:-1 Bumps:-1 RMS:-1

a) *Etotal: -99.35 Eshape:-99.35 Eforce:0.00 Vshape:0.00 Vclash:0.00*
b) *Etotal:-135.35 Eshape:-135.335 Eforce:0.00 Vshape:0.00 Vclash:0.00*
c) *Etotal:-99.35 Eshape:-99.35 Eforce:0.00 Vshape:0.00 Vclash:0.00*
d) *Etotal:- 143.64 Eshape: -143.64 Eforce:0.00 V Shape:0.00 Vclash:0.00*

Fig 5 Screen shot showing amino acid distance between protein and drug analogues with the binding sites. A). Anastrozole b) Extremestane, c). Epirubicin d). Letrozole



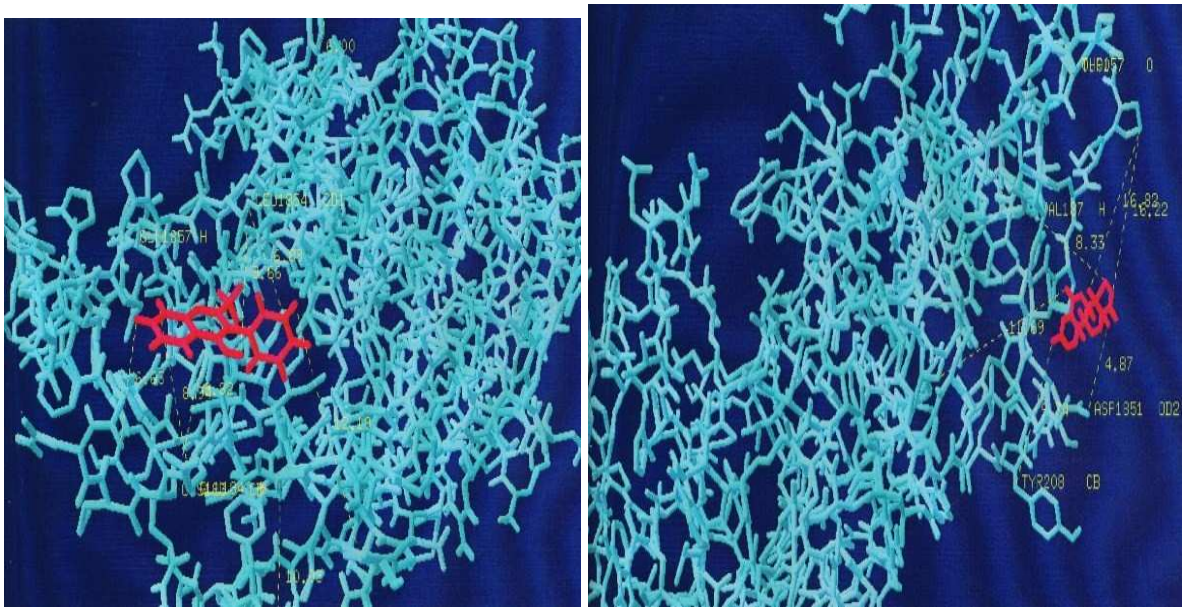


Table.1 In silico tools used for the homology modeling of BRCA1 gene

Program	Program used
NCBI	Storing and analyzing genetic and molecular data
SWISS model	Homology modeling server and protein modeling
BLAST	Search nucleic acid database and protein database (www.ncbi.gov/BLAST/)
HEX	Interactive protein docking and molecular super imposition program
KEGG	Specialized database of metabolic pathway
PIR	Protein identification resources (URL:http://pir.georgetown.program.edu).

Table.2 Analysis of amino acids and the distance between the ligands and modeled protein

Ligands	Amino acids	Distance in A°
Anastrozole	ARG 4,HIS 28,ILE 29,HIS 27,ME T 5	3.48 A°,1.23 A°,9.42 A°,6.02 A°,9.46 A°
Anastrozole analogue	PRO161,GLU184,CYS183,SIR1796LEU1854	4.26 A°,9.73 A°,5.95 A°,2.28 A°,4.52 A°
	THR207,GLU204,ASP206,LEU1854,TYR1853	8.99 A°,3.95 A°,5.16 A°,7.43 A°4.52
Epirubicin	GLU184, LEU1854, TYR208,CYS83,GLN1857	A°
Epirubicin analogue	LEU1854, GLU204, TYR208,VAL187, ASP1851	8.91 A°,8.66 A°,6.00 A°9.89 A°6.63 A°
Exemestane	HIS78,THR30,ARG25,LEU84,MET1728	4.95 A°,5.22 A°,9.14 A°,8.33 A°,4.87 A°
Exemestane analogue	LEU205,HIS160, LEU1854, ASP1851,GLU204 LYS66, MET138, GLU9,ALA107,GLN201	2.78 A°,7.55 A°,7.27 A°,5.12 A°,5.45 A° 1.97 A°,8.58 A°,4.85 A°5.33 A°,5.35 A°
Letrozole		6.70 A°,5.89 A°,7.15 A°6.05 A°8.44 A°
Letrozole analogue		

CONCLUSION

Suitable model of the 3D structure with potential target and the drug design step tried to find the optimal compound for moderating the normal function in the target sequence in selective and normal reversible way. Homology modeling has significant potential as a tool in rational drug design, in particular high throughput *in silico* screening or simulation approaches. The quality of the final structure mainly depends on the quality of the target-template alignment. Improvements in protocol alignment could improve the final model. Protein structures or ligands are not rigid systems, with high degree of flexibility. Improvements in these and other areas may finally turn homology based rational drug design into useful tool for the pharmaceutical industry.

REFERENCES

[1] Biggs, PJ., Warren, W., Ventt, S and Stratton, MR, **1993**, *Mutagenesi*, 8, 275-283.
[2] Greenstock, CL, **1993**, *Med Hypotheses*, 41, 513-523.
[3] Schapira, M., Raaka, BM., Samuels, HH and Abagyan, R, **2002**, *Proc. Natl. Acad. Sci*, 97,1008- 1015.
[4] Wieman, H., Kristin, T., Anderssen, E and Drablos, F, **2004**, *Mini-Reviews in Medicinal Chemistry*, 4, 793-804.
[5] Greer, J., Erickson, JW., Baldwin, JJ and Varney, MD, **1993**, *J. Med.Chem*, 37,8, 1035-1054.
[6] Wlodawer, A and Vondrasek, A, **1998**, *J. Annu. Rev. Biophys. Biomol. Struct*, 27, 249-284.
[7] Gray, NS., Wodicka, L., Thunnissen, AM., Norman, TC., Kwon, S., Espinoz, FH., Morgan, DO., Barnes, G., LeClerc, S., Meijer, L., Kim, SH., Lockhart, DJ and Schultz, PG, **1998**, *Science*, 281,533-538.
[8] Capdeville, R., Buchdunger, E., Zimmermann, J., Matter, A and Glivec, **2002**, *Nat.Rev. Drug Discov*, 1, 7, 493-502.
[9] Ghosh, S., Liu, XP., Zheng, Y and Uckun, FM, **2001**, *Curr. Cancer Drug Target*, 1,(2), 129-40.
[10] Zhu, X., Kim, JL., Newcomb, JR, Rose, PE, Stover, DR, Toledo, LM, Zhao H and Morgenstern, KA, **1999,** *Structure Fold. Des*, 7,651-662.
[11] Sawyer, T., Boyce, T., Dalgarno, B and Iuliucci, D, **2001**, *Opin. Investig. Drugs*, 10, 1327-1345.
[12] Govindarajan, S., Recabarren, R and Goldstein, RA, **1999,** *Protein*, 35, 329–332.
[13] Enyedy, IJ., Ling, Y., Nacro, K., Tomita, Y., Wu, X., Cao, Y., Guo, R., Li, B., Zhu, X., Huang, Y., Long, YQ., Roller, PP. Yang, D and Wang, S, **2001**, *J. Med. Chem*, 44, 513-517.
[14] Furet, P., Zimmermann, J., Capraro, HG., Meyer, T and Imbach, P, **2000**, *Aided Mol. Des*, 14, 5, 403-409.
[15] Sabnis, Y., Rosenthal, PJ., Desai, P and Avery, MA, **2002**, *Protein Sci*, 12,3,501- 509.
[16] Schafferhans, A., Klebe, G and Lig, P, **2001**, *J. Mol. Biol*, 307,407- 410
[17] Schonbrun, J., Wedemeyer, WJ and Baker, D, **2002**, Curr. Opin. Struct.Biol, 12, 3-42.
[18].Lengauer, C., Kenneth, W., Kinzler, KW and Volgelstein, B, **1998**, *Nature (Lond)*, 396, 643-649.
[19] Altschul, SF., Gish, W., Miller, W., Myers, EW and Lipman, DJ, **1990**, J. Mol. Biol, 15, 695-699
[20] Pearson, WR and Lipman, D, **1988**, *J. Proc. Natl. Acad. Sci*, 85,8, 2444-2448.

[21] Sali, S and Blundell, TL, **1993**, *J. Mol. Bio*, 234, 779-815.
[22] Guex, N and Peitsch, MC, **1997**, *Electrophoresis*, 18, 2714-2723.
[23] Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, DF., Evans, C., Deacon, J and Stratton, MR, **1999**. *J. Nalt. Cancer Ins*, 91, 943-949.
[24] Kolodny, R., Koehl, P., Guibas, L and Levitt, M, **2002**, *J. Mol. Biol*, 323,2,297-307
[25] Koehl, P and Levitt, M, **1999**, *Nat. Struct. Biol*, 6, 108-111.
[26] Brindha, V., Saravanan, A and Manimekalai A, **2009**, *Ind J. Science Technol*, 2, 2- 4.
[27] Kitchen, DB., Decornez, H., Furr, JR and Bajorath, J, **2007**, *Nature reviews Drug discovery*, 3,11, 935-949.
[28] Vijay, T., Shambhavi, S and Dwijendra Kumar, G, **2010**, *Archives of Applied Science Research*, 2, 6, 93-97.
[29] Tapar, KK., Rahul, D., Jawarkar, D., Vivek, V., Paithankar, P., Khataleb, N., Anil, M., Manikrao, Bhumesh, EW and Taibi, BH, **2011**, *J. Comput. Method. Mol. Design*, 2011, 1 (1): 1-6