IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Pharmacophore Mapping , Structure similarity search and virtual Screeningof ligand for human Estrogen receptor alpha in the Treatment of Breast Cancer

Author- S Patel [1],Ashwani Mishra [1],Pooja Gupta [2] ,Gaurav Singh lodhi[3] Nidhi Mishra [4]

1.Department of Pharmacy ,Barkatullah University ,Bhopal

2. Lakshmi Narain College of Pharmacy Sanwer ,Indore.

3.Gaurav Lodhi RAJIV GANDHI COLLEGE OF PHARMACY,Bhopal

Abstract

Breast cancer is one of the leading type of cancer and according to future prediction it is likely to strike 1.57 million in 2025 from 1.46 million this year.In this research work an attempt was made to use the drug Discovery and design technique Scaffold hopping in order to find the best ,safe and effective molecule available in the existing market on the basis of insilico screening techniques .Total eleven molecules scrutinized were Desmethyl Tamoxifen Alpelisib ,Norendoxifen, Abemaciclib , Anastrozole , ,Epigallocatechingallate, 4 Hydroxytamoxifen Z)-2-[4-(1,2)-Diphenyl -1-Butenyl)-Phenoxy]-N, Ndimethylethanamine, Afimoxifene, Endoxifen and Tamoxifen.These molecules were subjected for ADME Study ,Druglikeness study , Pharmacophore mapping and Docking study using Human Estrogen Receptor.The best molecules which revealed all the best ,safe and promising attributes was Anastrozole.

1.INTRODUCTION

Breast cancer is a type of severe cancer that spread in the breast. Cancer starts due to cells begin to grow out of control manner . The Breast cancer cells usually form a tumour that can commonly be seen on an x-ray or felt as a lump.[1] Breast cancer occurs almost entirely in women, but men can get breast cancer, too.1As per the bi-annual report of Indian Council of Medical Research (ICMR) that Breast cancer is likely to strike 1.57 million in 2025 from 1.46 million this year[2]

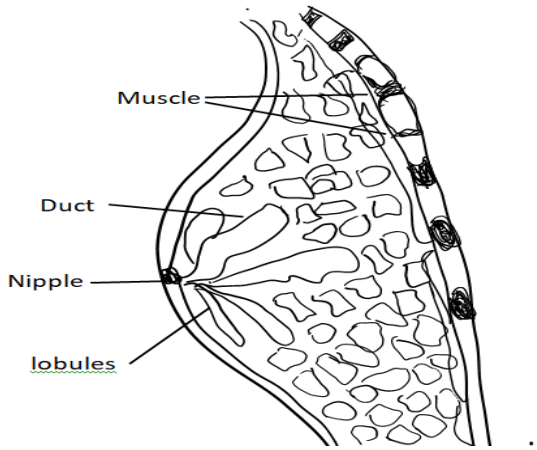


Figure-1 Breast Structure

Majority of breast lumps are benign not malignant. Non-cancerous breast tumorsmay grow abnormally, but generally they do not spread outside the breast. These types of Tumors do not cause risk of life though some types of benign breast lumps may cause increase a woman's risk of getting breast cancer.[3]

It is advised that any breast lump or change in breast needs to be checked by Physician toidentify if it is benign or malignant cancer and if it affectslife or future cancer risk.[4]

Estrogen is an important enzyme for survival and health in both genders.[5]Majority of the actions of estrogens are mediated by Estrogen Receptor (ER). A high level of Estrogen is linked with increased risk of Breast cancer which stimulating breast epithelial cell proliferation. TheestrogensReceptor exists in two forms ER beta and ER alpha . ERα and ERβ have different function in the progression of breast cancer. Maximum cases of the breast cancer is identified by abnormal expression of Estrogen Receptor α-positive affecting aa\round 70% of the patient of primary breast cancer.[6]Estrogen stimulated cell proliferation and increased tumor formation hence, inhibition of Estrogen Receptor alpha is most widely used approach for the prevention of Breast cancer[7]. The X-ray Crystallographic structure of the Human Estrogen Receptor with 1.6Å resolution was retrieved from Protein Data Bank (PDB), PDB ID: 2IOG was used as a potential anticancer drug target for the screening of the potent, safe and effective anticancer drug [8].Scaffold hopping is a key task of advanced medicinal chemistry for reasonable drug design.It aims to design molecules of novel scaffolds sharing same target biological activities toward known hit molecules. Scaffolding hopping depends on searching databases of known compounds that can't exploit vast chemical space[9].In this work as attempt was made to find the best molecule with the help of scaffold hopping in different database which is safe and promising through in-silico research work.

2.In-silico researchwork

2.1 Scaffold Hopping - In this workthe most potent anticancer drug used majorly for the treatment of the breast cancer has been subjected to the software for the pharmacophore mapping ,during pharmacophore mapping study number of hydrogen donor ,hydrogen acceptor , hydrophobic and hydrophobic pockets of the ligand was analysed .After the in silico analysis of the Pharmacophoric attributes of the molecules ,The selected molecule has been subjected for similarity search in freely available online database like Zinc 15

and CHEMBL .Total 380 molecules were selected from CHEMBL data base and 404 molecules were selected from Zinc Data base .

2.2 ADMEsar study-These selected molecules from scaffold hopping Alpelisib, DesmethylTamoxifen ,Afimoxifene,Norendoxifen, Abemaciclib, Anastrozole , Epigallocatechingallate, 4-hydroxytamoxifen ,(z)-2-[4-(1,2)-Diphenyl-1-Butenyl)-Phenoxy]-N,N-Dimethylethanamine,, Endoxifen and Tamoxifen.were then further comaratively analyzed for the absorption ,distribution ,metabolism and excreation attributes like Molecular weight , Num. heavy atoms ,Num. arom. heavy atoms, Fraction Csp3, Num. rotatable bonds ,Num. H-bond acceptors, Num. H-bond donors Molar Refractivity, TPSA ,Solubility , class .The pharmacokinetic parameters like GI absorption , BBBpermeate ,P-gp substrate ,CYP1A2 inhibitor, CYP2C19 inhibitor ,CYP2C9 inhibitor and skin permeation .

2.3 Druglikeliness-Druglikeliness attributes of selected molecules were based on the rules like Lipinski rule of 5 ,Ghose ,Veber ,Egan and Muegge.In this druglikeliness study the more emphasis was given to the violation exhibited by the selected molecules which is based on the lesser the violation of the molecule the better the druglikeliness property of the selected molecules .The last prediction was of Bioavailability which exhibits the therapeutic potential of the molecule.

2.4 Molecular Docking Study -EstrogenReceptor Protein was fetched from RCSB website in PDB format.The PDB ID 2IOG.The crystal structure of the protein was subjected for tunnel analysis and to visualize energetically allowed regions for backbone dihedral angles ψ against ϕ of amino acid residues to protein structure .The molecule exhibited best attributes Anastrozole was subjected for molecular docking with the EstrogenReceptor Protein using R package.

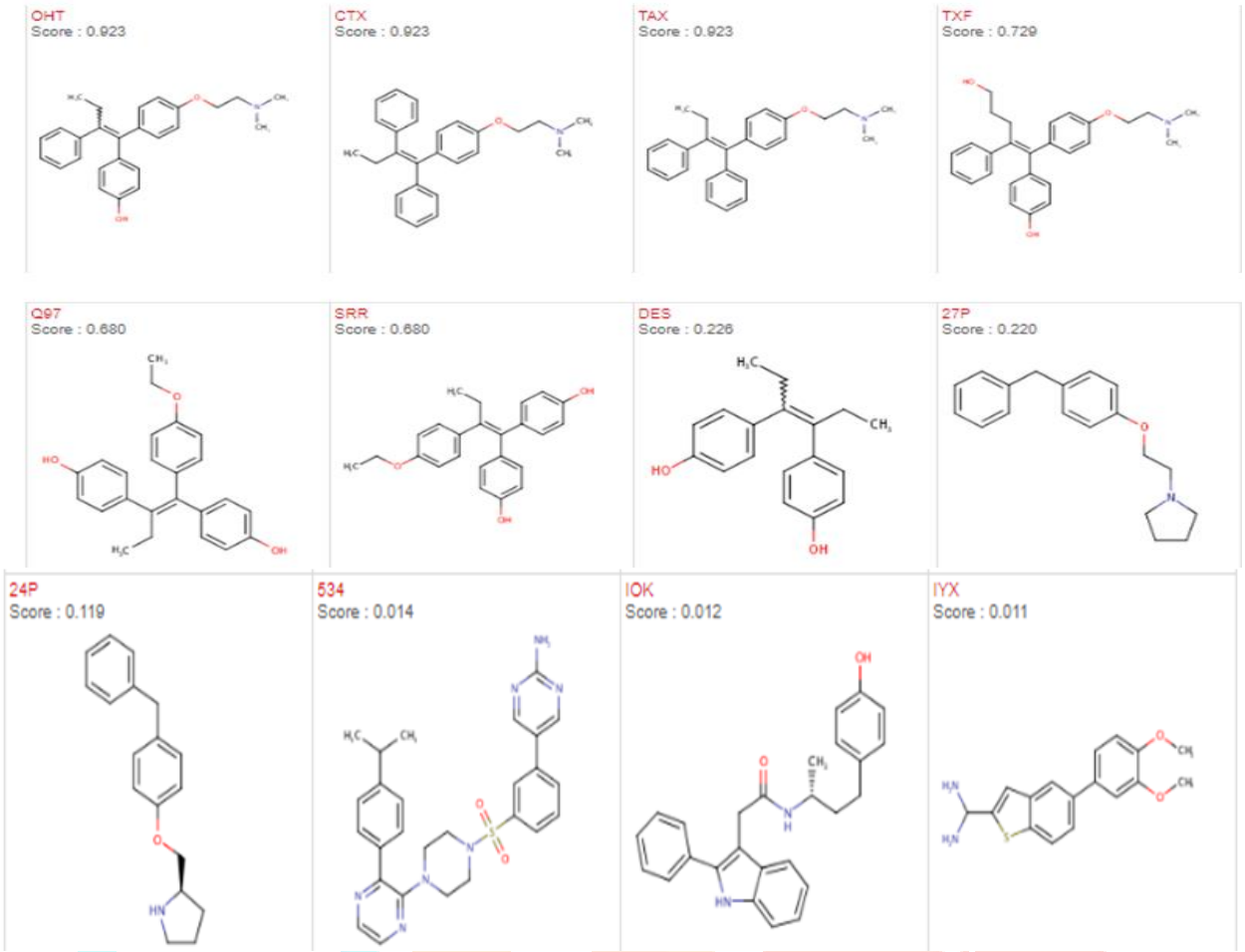


Figure -2 Result of the best match of Scaffold Hopping of the compounds having good scored out of 224 Structure

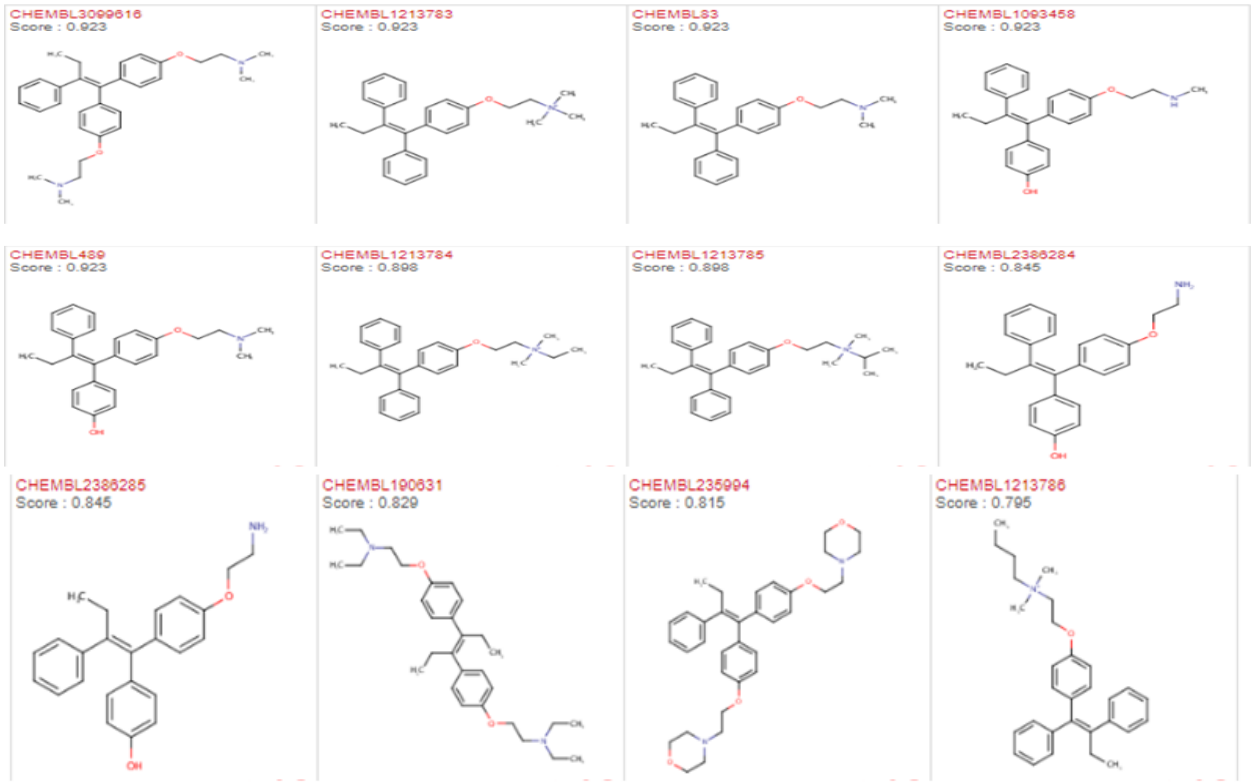


Figure -3 Result of best match of Scaffold Hopping of the compounds having good scored out of 380 Structure in CHEMBL data base

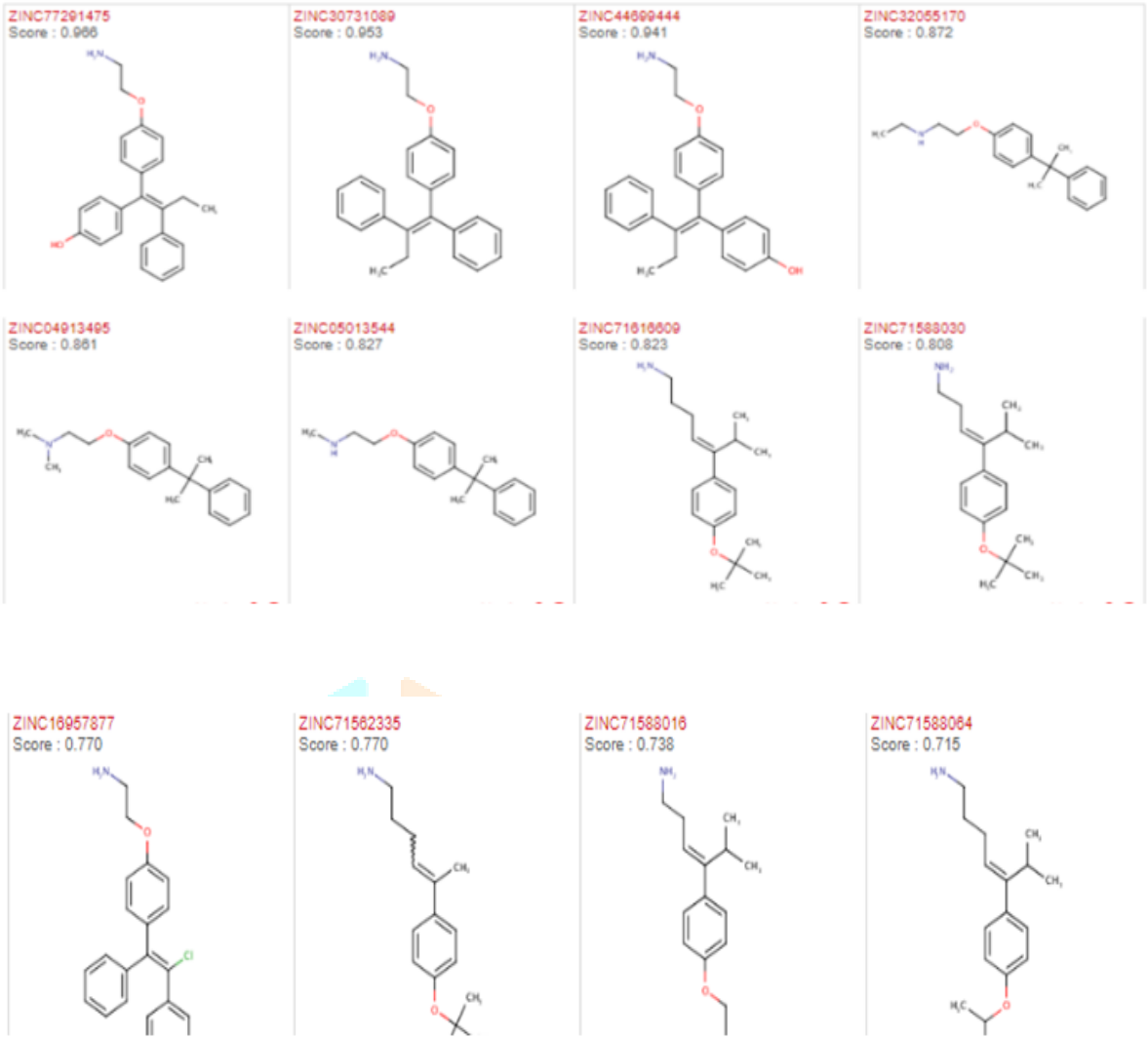
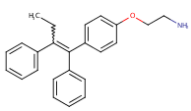
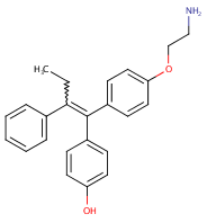
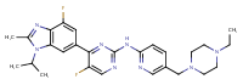
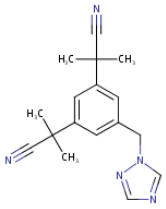
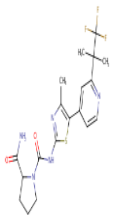
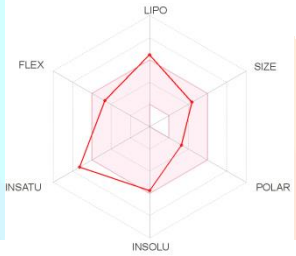
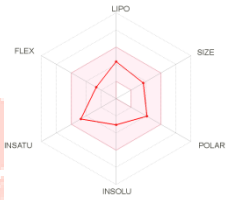
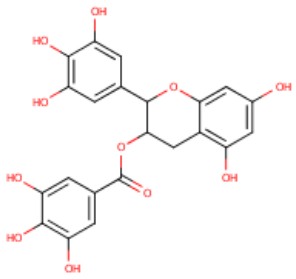
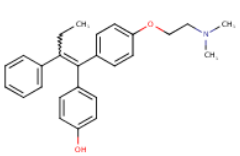
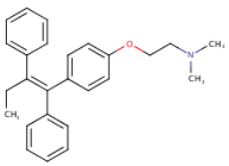
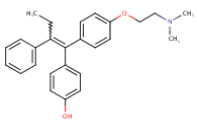
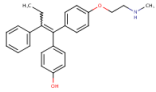
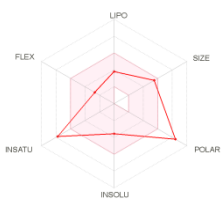
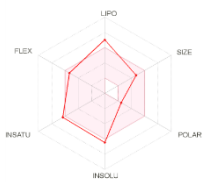
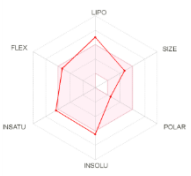
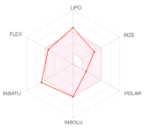
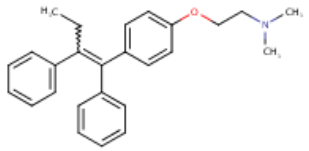


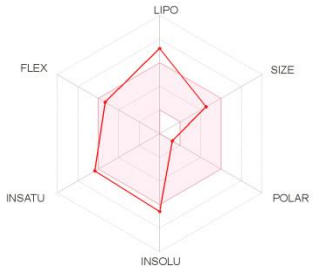
Figure -4 Result of best match of Scaffold Hopping of the compounds having good scored out of 404 Structure in Zinc data base.

Table -1 Plot and structure of Different molecules selected by Scaffold Hopping

DesmethylTamoxifen	**Norendoxifen**	**Abemaciclib**	**Anastrozole**	**Alpelisib**
				
				
Epigallocatechingallate	4-HYDROXYTAMOXIFEN	(Z)-2-[4-(1,2)-DIPHENYL-1-BUTENYL)-PHENOXY]-N,N-DIMETHYLETHANAMINE-	AFIMOXIFENE	ENDOXIFEN
				
				

Tamoxifen





3.2 Property determination of high score compounds

Table2 - Different ADME properties of the molecules .

Formula	**ENDOXIFEN**	**DesmethylTamoxifen**	**Norendoxifen**	**Abemaciclib**
Formula	C25H27NO2	C24H25NO	C24H25NO2	C27H32F2N8
Molecular weight	373.49 g/mol	343.46 g/mol	359.46 g/mol	506.59 g/mol
Num. heavy atoms	28	26	27	37
Num. arom. heavy atoms	18	18	18	21
Fraction Csp3	0.20	0.17	0.17	0.41
Num. rotatable bonds	8	7	7	7
Num. H-bond acceptors	3	2	3	8
Num. H-bond donors	2	1	2	1
Molar Refractivity	116.84	109.92	111.94	149.17
TPSA	41.49 Å²	35.25 Å²	55.48 Å²	75.00 Å²
Log *S* (ESOL)	-6.08	-5.90	-5.76	-5.36
Solubility	3.07e-04 mg/ml ; 8.22e-07 mol/l	4.32e-04 mg/ml ; 1.26e-06 mol/l	6.24e-04 mg/ml ; 1.74e-06 mol/l	2.22e-03 mg/ml ; 4.38e-06 mol/l
Class	Poorly soluble	Moderately soluble	Moderately soluble	Moderately soluble
	Pharmacokinetics			
GI absorption	High	High	High	High
BBB permeant	Yes	Yes	Yes	No
P-gp substrate	Yes	Yes	Yes	Yes
CYP1A2 inhibitor	Yes	Yes	Yes	No
Log K_p (skin permeation)	-4.09 cm/s	-4.02 cm/s	-4.37 cm/s	-6.66 cm/s
	Druglikeness			
Lipinski	Yes; 1 violation: MLOGP>4.15	Yes; 1 violation: MLOGP>4.15	Yes; 0 violation	Yes; 1 violation: MW>500

Ghose	Yes	Yes	Yes	No; 2 violations: MW>480, MR>130
Veber	Yes	Yes	Yes	Yes
Egan	Yes	Yes	Yes	Yes
Muegge	No; 1 violation: XLOGP3>5	No; 1 violation: XLOGP3>5	No; 1 violation: XLOGP3>5	Yes
Bioavailability Score	0.55	0.55	0.55	0.55

Table 3 - Different properties of the Drugs used to treat Breast cancer .

Formula	**Anastrozole**	**TAMOXIFEN**	**Alpelisib**	**Epigallocatechingallate**
	C17H19N5	C26H29NO	C19H22F3N5O2S	
Molecular weight	293.37 g/mol	371.51 g/mol	441.47 g/mol	C22H18O11
Num. heavy atoms	22	28	30	458.37 g/mol
Num. arom. heavy atoms	11	18	11	33
Fraction Csp3	0.41	0.23	0.47	18
Num. rotatable bonds	4	8	7	0.14
Num. H-bond acceptors	4	2	7	4
Num. H-bond donors	0	0	2	11
Molar Refractivity	83.81	119.72	111.31	8
TPSA	78.29 Å²	12.47 Å²	129.45 Å²	112.06
Log S (ESOL)	-3.04	-6.59		-3.56
Solubility	2.65e-01 mg/ml ; 9.04e-04 mol/l	9.56e-05 mg/ml ; 2.57e-07 mol/l	1.70e-02 mg/ml ; 3.85e-05 mol/l	1.27e-01 mg/ml ; 2.76e-04 mol/l
Class	Soluble	Poorly soluble	Moderately soluble	Soluble
GI absorption	High	Low	Low	Low
BBB permeant	Yes	No	No	No
P-gp substrate	No	Yes	Yes	No
CYP1A2 inhibitor	No	No	No	No
Log Kp (skin permeation)	-6.65 cm/s	-3.50 cm/s	-6.71 cm/s	-8.27 cm/s
Drug likeness				
Lipinski	Yes; 0 violation	Yes; 1 violation: MLOGP>4.15	Yes; 0 violation	No; 2 violations: NorO>10, NHorOH>5
Ghose	Yes	No; 1 violation: WLOGP>5.6	Yes	Yes
Veber	Yes	Yes	Yes	No; 1 violation: TPSA>140
Egan	Yes	No; 1 violation: WLOGP>5.88	Yes	No; 1 violation: TPSA>131.6
Muegge	Yes	No; 1 violation: XLOGP3>5	Yes	No; 3 violations: TPSA>150, H-acc>10, H-don>5
Bioavailability Score	0.55	0.55	0.55	0.17

Table 4- Different properties of the Drugs used to treat Breast cancer .

Parameters	**4-HYDROXYTAMOXIFEN**	**(Z)-2-[4-(1,2)-DIPHENYL-1-BUTENYL)-PHENOXY]-N,N-DIMETHYLETHANAMINE-**	**AFIMOXIFENE**
Formula	C26H29NO2	C26H29NO	C26H29NO2
Molecular weight	387.51 g/mol	371.51 g/mol	387.51 g/mol
Num. heavy atoms	29	28	29
Num. arom. heavy atoms	18	18	18
Fraction Csp3	0.23	0.23	0.23
Num. rotatable bonds	8	8	8
Num. H-bond acceptors	3	2	3
Num. H-bond donors	1	0	1
Molar Refractivity	121.75	119.72	121.75
TPSA	32.70 Å²	12.47 Å²	32.70 Å²
Water Solubility			
Log *S* (ESOL)	-6.45	-6.59	-6.45
Solubility	1.37e-04 mg/ml ; 3.54e-07 mol/l	9.56e-05 mg/ml ; 2.57e-07 mol/l	1.37e-04 mg/ml ; 3.54e-07 mol/l
Class	Poorly soluble	Poorly soluble	Poorly soluble
Pharmacokinetics			
GI absorption	High	Low	High
BBB permeant	Yes	No	Yes
P-gp substrate	Yes	Yes	Yes
CYP1A2 inhibitor	No	No	No
Log K_p (skin permeation)	-3.84 cm/s	-3.50 cm/s	-3.84 cm/s
Druglikeness			
Lipinski	Yes; 1 violation: MLOGP>4.15	- Yes; 1 violation: MLOGP>4.15	Yes; 1 violation: MLOGP>4.15
Ghose	No; 1 violation: WLOGP>5.6	- Yes	No; 1 violation: WLOGP>5.6
Veber	Yes	- Yes	Yes
Egan	Yes	Yes	Yes
Muegge	No; 1 violation: XLOGP3>5	Yes	No; 1 violation: XLOGP3>5
Bioavailability Score	0.55	-0.55	0.55

3.3 EstrogenReceptor Protein Analysis for Pockets and volume

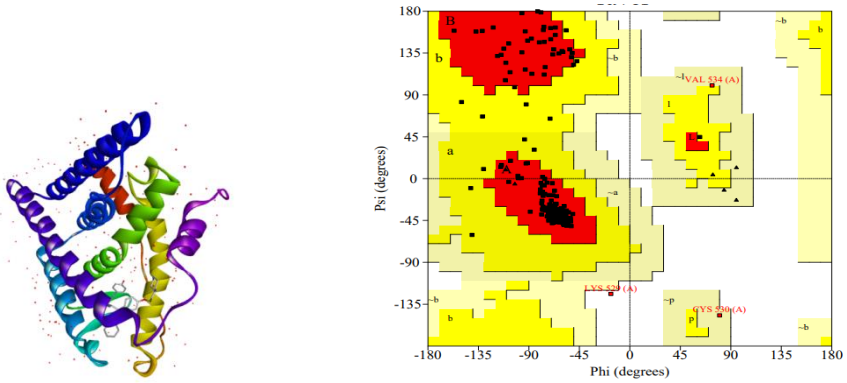


Figure -5 Protein ID 2IOG

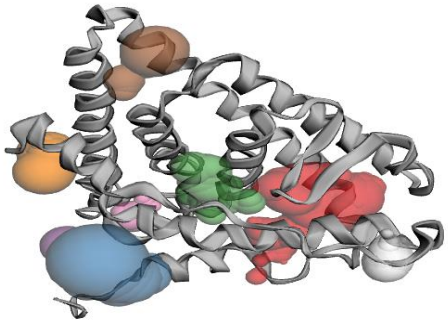


Figure 6 -Protein with 36 pockets.

Table-5 Shows Important pockets and color coding of Protein ID 2IOG

Pocket ID	Area SA	Color
1	262.7	Red
2	92.6	Blue
3	86.5	Green
4	75.3	Purple
5	41.2	Orange
6	40.1	Brown
7	11.8	Pink
8	11.4	White

3.4 Pharmacopore mapping of Anastrozole[10]

The Pharmacophore analysis of the Anastrozole Shows four hydrogen acceptor ,three hydrophobic group and one aromatic group.

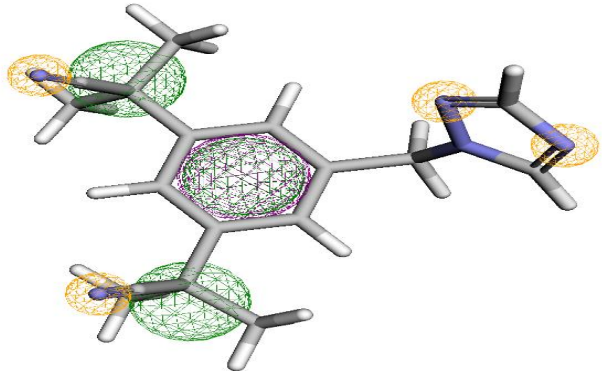


Figure 7 - Pharmacophoric Mapping of Anastrozole.

3.5 Binding Analysis of the Anastrozole with 2IOGEstrogen protein

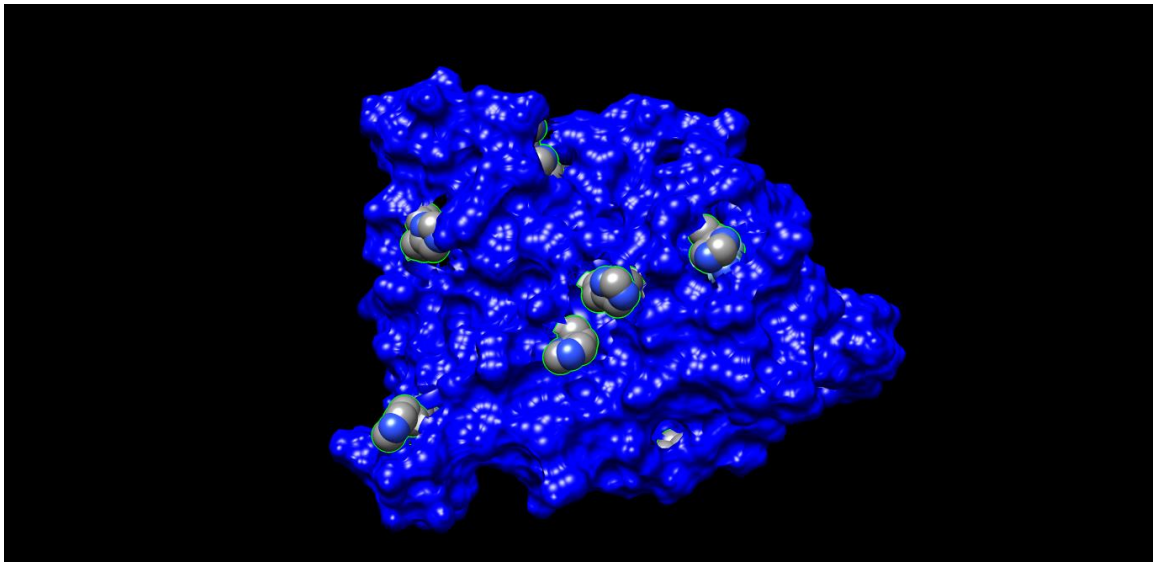


Figure 8 -Anastrozole ligand in the protein cavity

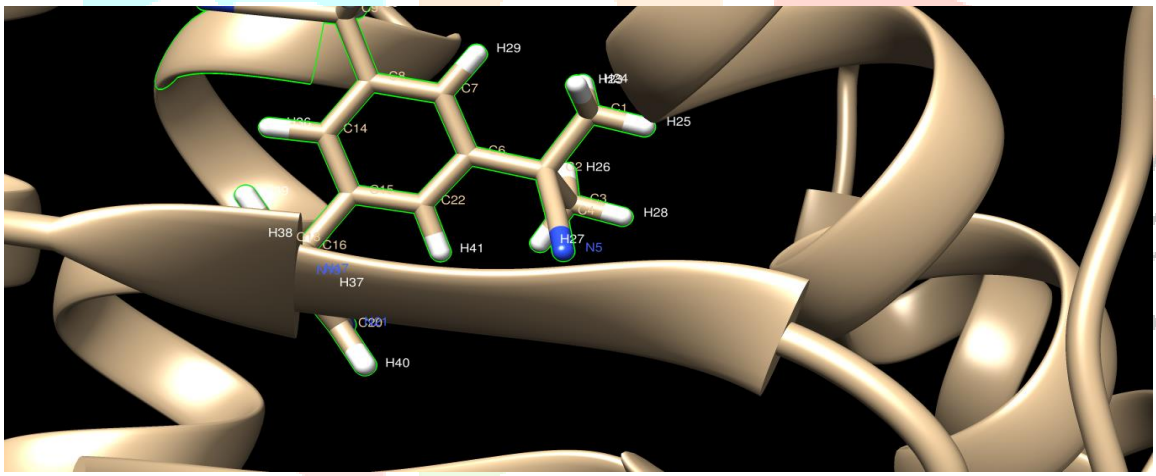


Figure 9 - Cluster Anastrozole ligand in to protein cavity and its binding .

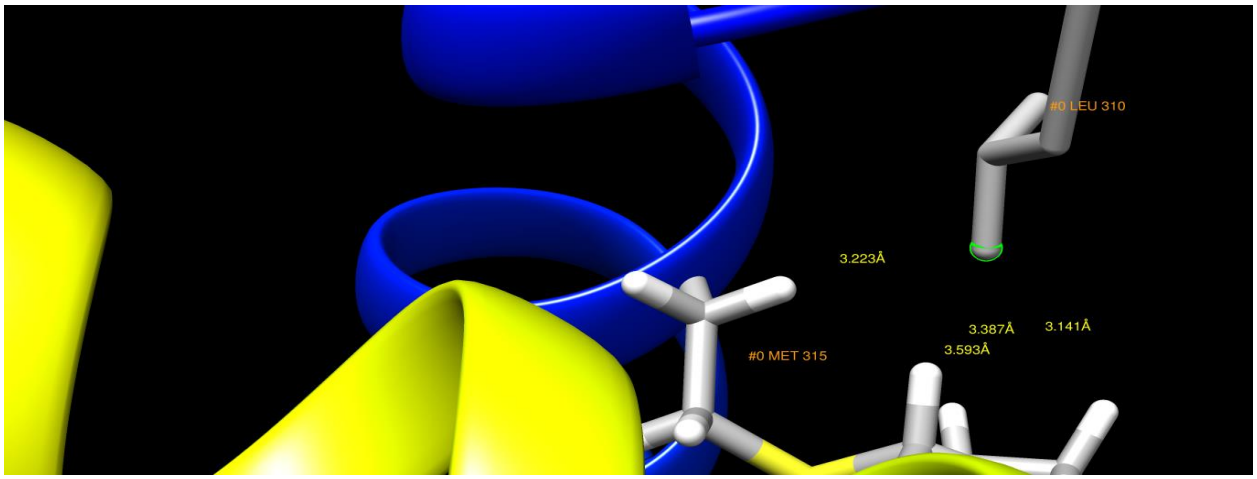


Figure 10 - Distance of theAnastrozole ligand with the protein residue exhibits moderate binding.

4 Result and discussion -

4.1.Scaffold hopping -From Scaffold hopping total 1008 structure were fond out of which on the basis of matching score only 32 structures were selected from different freely available data library and data sources .These structures were further subjected for dry lab screening by using radar plots.

Insilico Protein analysis was done and as per Ramchandran plot 93.3% Residues in most favoured regions [A,B,L] 208 amino acid , 5.4% Residues in additional allowed regions [a,b,l,p] 12 amino acid , 0.9% Residues in generously allowed regions [~a,~b,~l,~p] 2, in disallowed regions 1 amino acid and ,0.4%Number of non-glycine and non-proline residues 223 ,100.0%.This revealed that most of the amino acid are in the allowed region and the crystal structure of protein is good . Total 36 Pockets were found in the analysis of the protein tunnel .

4.2.Admetstudy- The adme study of Desmethyl Tamoxifen Alpelisib ,Norendoxifen, Abemaciclib , Anastrozole , ,Epigallocatechingallate, 4 Hydroxytamoxifen Z)-2-[4-(1,2)-Diphenyl -1-Butenyl)-Phenoxy]-N, Ndimethylethanamine, Afimoxifene, Endoxifen and Tamoxifen were performed and it was observed that Abemaciclib had high molecular weight , poorly soluble ,and the biggest point of its rejection from this study was its violation from Lipinski rule as well as Ghose .Only Anastrozole and Alpelisib exhibited less violation but in the Alpelisib it had more TPSA , Of the binding study only Anastrozole was chosen .Radar Plot Interpretation - The most promising compound were further scrutinized which wereDesmethylTamoxifenAlpelisib,Norendoxifen, Abemaciclib ,Anastrozole ,

,Epigallocatechingallate, 4 Hydroxytamoxifen Z)-2-[4-(1,2)-Diphenyl -1-Butenyl)-Phenoxy]-N,Ndimethylethanamine,Afimoxifene, EndoxifenandTamoxifen. The result revealed that drugs which are reported in the treatment of Breast cancer Epigallocatechingallate, 4-Hydroxytamoxifen,(Z)-2-[4-(1,2)-Diphenyl -1-Butenyl)-Phenoxy]-N,Ndimethylethanamine ,Afimoxifene, DesmethylTamoxifen and Norendoxifen all crossed the pink zone of the Radar curve which indicated that one of the parameter among properties (lipophilicity: XLOGP3 between −0.7 and +5.0, size: MW between 150 and 500 g/mol, polarity: TPSA between 20 and 130 $Å^2$, solubility: log S not higher than 6, saturation: fraction of carbons in the sp^3 hybridization not less than 0.25, and flexibility: no more than 9 rotatable bonds. were out of the optimum range .Only three structures exhibited better druggable properties which were Alpelisib,Abemaciclib and Anastrozole.

4.3 Drug Likeliness Endoxifen, DesmethylTamoxifen, Norendoxifen, Tamoxifen,Epigallocatechingallate ,4-Hydroxytamoxifen, Z)-2-[4-(1,2)-Diphenyl-1-Butenyl)-Phenoxy]-N,N-Dimethylethanamine, Afimoxifene and Abemaciclib revealed violation from the drug likeliness attributes based on 5 rules Lipinski, Ghose Veber, Egan and Muegge .The only two drugs who didn't show any violation were Anastrozole and Alpelisib,Which disclosed their good drugability among all the molecules selected by scaffold hopping .

4.4 Molecular docking exhibited that The lest energy -7.60Kcal/mol of the interaction was shown by theAnastrozole with Estrogen protein 2IOG As it has been reported that Most of the breast cancer is identified by abnormal expression of Estrogen Receptor α-positive affecting about 70% of the primary breast cancer patients. The activity of Anastrozole with this receptor predicted that among Desmethyl ,Tamoxifen ,Alpelisib ,Norendoxifen, Abemaciclib ,Anastrozole , ,Epigallocatechingallate, 4 Hydroxytamoxifen Z)-2-[4-(1,2)-Diphenyl -1-Butenyl)-Phenoxy]-N,Ndimethylethanamine, Afimoxifene, Endoxifen and Tamoxifen ,Anastrozole showed better effect in-silico for the binding with the estrogen receptor and it could be used potentially in the treatment of breast cancer in most promising way if the activity of these molecule will be performed on different models .

5 Conclusion - It has been proved in various research work that in-silico techniques have become a boon for the researcher since they not only can be done in laptops or computers but they also reduce the time required for the invention and research and would be able to provide the result with high accuracy and precision since it is hardly getting affected buy many causes of variation .It could be concluded from this research work that like the potential of Anastrozole has been determined in this research work , these sort of model can be used in different diseases to find the most potent molecule in the treatment of Patient. This predictive study could be studied by performing the clinical trial on the selected molecule for the final approval.

6.References –

1https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-2035247

2.https://www.livemint.com/news/india/icmr-data-shows-unequal-toll-of-cancer-on-women-11670349329355.html

3.Chan JCH, Chow JCH, Ho CHM, Tsui TYM, Cho WC. Clinical application of circulating tumor DNA in breast cancer.J Cancer Res ClinOncol. 2021;147(5):1431-1442.

4.Breast lump- Medline Plus, NIH, U.S. National Library of Medicine [Internet]. medlineplus.gov 2019 [Cited 18 July 2019]. Available from:https://medlineplus.gov/ency/article/003155.htm

5.Patel S, Homaei A, Raju AB, Meher BR. Estrogen: The necessary evil for human health, and ways to tame it. Biomed Pharmacother. 2018 Jun;102:403-411.

6. Patel S, Homaei A, Raju AB, Meher BR. Estrogen: The necessary evil for human health, and ways to tame it. Biomed Pharmacother. 2018 Jun;102:403-411.

7Duffy MJ.Estrogen receptors: role in breast cancer. Critical Reviews in Clinical Laboratory Sciences. 2006. 43(4): 325-347

8 .Kevin D. Dykstra, LiangqinGuo, Elizabeth T. Birzin, Wanda Chan, Yi Tien Yang, Edward C. Hayes, Carolyn A. DaSilva, Lee-YuhPai, Ralph T. Mosley, Bryan Kraker, Paula M. D. Fitzgerald, Frank DiNinno, Susan P. Rohrer, James M. Schaeffera and Milton L. Hammond. Estrogen receptor ligands. Part 16: 2-Aryl indoles as highly subtype selective ligands for ERα. Bioorganic & Medicinal Chemistry Letters 2007.2322–2328.

9 Zheng, S., Lei, Z., Ai, H. *et al.* Deep scaffold hopping with multimodal transformer neural networks. *J Cheminform* **13**, 87 (2021).

10 Meshram,R J., Baladhye ,V B., Gacche ,R N., Karale , B K., Gaikar,R B. Pharmacophore Mapping Approach for Drug Target Identification: A Chemical Synthesis and in Silico Study on Novel Thiadiazole Compounds .Journal of Clinical and Diagnostic Research. 2017 ;11(5): KF01-KF08.