Chapter 1

Design of a New Class of ER ligands
1. Introduction

In carcinogenesis, hormones are not only involved in initiation but also in the progression of the disease. However, due to an apoptotic effect, some hormones might have the opposite effect, for example reduce cancer risk or slow down tumor progression. On the other hand, a preserved functional hormonal pathway represents a target for anti-hormonal therapy in endocrine-related malignancies, such as breast, endometrial, ovary, prostate or testis cancers. Hormone-related carcinogenesis is of very complex aetiology due to the combination of genetic susceptibility, epigenetic changes and exposure to endogenous and exogenous hormones and hormone-like substances. Estrogen is the hormone involved in aetiology of carcinogenesis in responsive tissues, viz. the endometrium, breast, ovary and prostate, which are the models for studying hormonal carcinogenesis. The main effects of estrogen are mainly mediated through its receptors; thereby their expression patterns are one of the major determinants of incidence and progression of the disease.

Expression profiling in cancer is still an open issue, despite the intensive efforts in microarray-based studies. Estrogen and its receptor are mainly responsible for in majority of cancers of gynecological origin in female population. Since, Estrogen Receptor’s (ER’s) functional expression is confined to the initial stages of neoplastic transformation; designing of molecules that can interfere with estrogen receptor-function will be an effective strategy to prevent this type of cancer. β-estradiol (ES), the chief endogenous steroid hormone ligand of ER, primarily regulates a broad range of physiological processes such as growth, differentiation, and physiology of reproductive processes. Estrogen-bound ER classically acts as a transcription factor thereby regulating cell proliferation and influences pathological processes of the hormone-dependent breast, endometrial, prostate, ovarian and thyroid cancers. This makes ER one of the principal targets for cancer therapy. In order to get newer selective ER ligands with improve potency against the ER associated cancer; we have used molecular modeling techniques. It can remodel the experimental studies of protein-ligand complexes by modeling their structures and binding affinities in silico. Herein, we describe how we have designed some of the new class of ER ligands with improved potency.

2. Estrogen Receptor

Estrogen receptors were first identified by Elwood V. Jensen at the University of Chicago in 1958. The gene for a second estrogen receptor (ERβ) was identified in 1996 by Kuiper et al. in rat prostate and ovary using degenerate ER-alpha primers. Estrogen receptors (ERs) are a group of proteins found inside and on cells. They belong to the large family of nuclear receptors. These are the family of structurally related ligand-inducible transcription factors including steroid receptors (SRs), vitamin D receptors (VDR, LXR, and PPARs), thyroid/retinoid receptors (TR, RARs, and RXRs) and orphan receptors. All nuclear receptors share a similar structure and are activated by small lipophilic molecules like glucocorticoids, progesterone, estrogens (β-estradiol), retinoids and fatty acid derivatives. There are two classes of ER exists:
ERα and ERβ, which are members of the nuclear receptor family of intracellular receptors, and membrane estrogen receptors (mERs) (GPER (GPR30), ER-X, and Gα-mER), which are mostly G protein-coupled receptors. ERs regulate many different and sometime quite opposite processes in cells such as growth and differentiation. Estrogen is involved in morphogenesis of the mammary gland, prostate, and lungs. It is a tropic factor for neurons, and it is involved in maintaining the normal structure of mammary epithelium and fertility in females.

2.1. Estrogen action on Estrogen Receptor

There are several ways how estrogen and ER act on cells. The main and one of the first recognized is the interaction of estrogen with ER into estrogen-ER complex which may act directly on Estrogen response elements (ERE) in the promoters of estrogen responsive genes or may act by interacting with AP-1 or SP-1 complexes. This ligand-dependant activation of transcription via direct genomic action through ERE or tethering to other proteins is influenced by ligand, ER subtype and expression of different cell specific co-regulators. Another mechanism, so called rapid non-genomic mechanism is not well investigated, but it is associated either with membrane ER or with some other membrane or cytoplasmic protein. The ligand-dependant pathways of estrogen action are shown in Figure 1. The ligand-independent mechanism of ER activation involves growth factor signaling and kinases that phosphorylate ER and thereby activate it in the absence of a ligand.

![Figure 1. The ligand-dependant pathways of estrogen action (AP-1 and SP-1 indicate transcriptional factors). The ligand-independent mechanism of ER activation involves growth factor (GF) signalling and kinases that phosphorylate ER and thereby activate it in the absence of a ligand.](image)

In many cell types estrogen stimulates proliferation and inhibits apoptosis, but in others, estrogen induces apoptosis. Cell growth by estrogen is achieved by induction of transition from G1 to S-phase of the cell cycle, through upregulation of c-myc which controls cyclin D expression. ERα also helps in activation of mitosis. Conversely, in some breast and prostate cancer cells, after long-term estrogen deprivation, estrogen may induce estrogen mediated mitochondrial pathway of apoptosis. To date, apoptotic action of estrogen is mainly addressed to the upregulation of FasL gene expression either via direct genomic action.
on ERE in the promoter region of FasL gene \(^{36}\) or via interaction with AP-1 and SP-1 proteins. \(^{37,38}\) This dual role of a single hormone needs to be further resolved in the different cell types and physiological conditions. The apoptotic effects of estrogen may be applicable in the clinical management of postmenopausal breast cancer patients treated with tamoxifen, cancer patients treated with tamoxifen, the group in which estrogen may induce tumor regression as well in cases with estrogen negative breast cancer. \(^{39}\)

2.2. Role of estrogen in cancer pathology

There are two possible ways how estrogen initiates and promotes carcinogenesis

(i) via its proliferative effects which increase the numbers of cell divisions and accumulation of mutations in DNA

(ii) via estrogen metabolism, accumulation of intermediary products with genotoxic effect in cells

2.3 Types of cancers associated with Estrogen receptor

2.3.1. Breast cancer

It is well known that 50-80% of breast carcinoma patients have tumors with measurable ER\(\alpha\) levels, i.e. ER\(\alpha^+\) tumors. \(^{42}\) ER\(\alpha^+\) tumors are sensitive to endocrine therapy. \(^{43}\) More than half of ER\(\alpha^+\) breast carcinomas express progesterone receptor (PR) \(^{44}\) that mediates progesterone’s effects on in the development of the mammary gland and breast carcinoma, where estrogen signaling via ER\(\alpha\) is necessary to induce PR expression. \(^{45}\) The status of ER/PR (or steroid receptors), for primary breast cancer patients is accepted to provide potentially relevant information regarding the natural or clinical course of the disease. \(^{46}\) While steroid receptor status of the primary breast cancer is a prognostic indicator for patient's outcome, though it is considered a weak one, it has been proven to be a predictor of response to endocrine therapy since up to 80% of patients bearing ER+PR+ tumors respond to endocrine treatment. \(^{47}\)

The facts that steroid status is not a powerful prognostic marker, that some breast cancers patients with steroid receptors-negative status respond to endocrine treatment and that certain number of breast carcinomas will recur after such treatment in spite of steroid receptors- positivity emphasize the need for identifying markers complementary to SR status in order to improve the prognosis and prediction of breast cancer patients. Attention has been directed to estrogen-regulated proteins, including pS2, cathepsin D, and ER\(\beta\), assuming that these proteins may be indicators of a functional signal transduction pathway through which tumor cells respond to estrogen (or antiestrogen) stimulation. In addition, the new markers related to cell cycle regulation and those detectable in circulating DNA are under intensive investigation. The role of ER\(\beta\) and its predictive value in breast cancer is still not clear, but there is evidence that ER\(\beta\) transcription is down regulated during breast tumorigenesis. It has been shown that the expression of ER\(\alpha\) increases during the process of carcinogenesis, but the expression of ER\(\beta\) seems to decrease. \(^{48}\) ER\(\beta\) is under intensive
investigation and its role in breast cancer appears to be of additional predictive value. Selective estrogen receptor modulators (SERMs) are used in clinical management of ERα positive breast cancer. However, about 30% of breast cancers are initially ER negative and resistant to endocrine therapies.\textsuperscript{49} In addition, some initially ER positive breast cancers evolve to an estrogen-independent growth phenotype.\textsuperscript{49}

Moreover, it has become clear that ERs and PR in living cells represent a pool of different variant proteins; therefore, it is reasonable to propose that ER/PR status should now include ERα, ERβ, and PR receptors together with some of their isoforms and functionally active splice variants. The several isoforms of ERs described at the protein level and the presence of numerous mRNA splice variants suggest the possibility that expression profile of these variant protein tumors might be involved in tumor progression\textsuperscript{50} and tamoxifen resistance.\textsuperscript{51-53}

\textbf{2.3.2. Prostate cancer}

In addition to androgens, estrogens also play an important role in the development of the male reproductive system. Clinical studies show that increased estrogen serum level or increased estrogen/androgen ratio is associated with increased risk of prostate cancer. ERβ was initially cloned from prostate tissue.\textsuperscript{54} The facts that this receptor is highly expressed in prostate and that ERβ-, but no ERα-knockout mice displayed prostatic epithelial hyperplasia underlie the involvement of ERβ etiology of prostate hyper proliferative disease and cancer.\textsuperscript{55-58}

\textbf{2.3.3. Ovarian cancer}

The majority of ovarian cancers arise from the ovarian surface epithelium and only 5\% of from the granulosa cells. About 70\% of all ovarian cancers are ER-positive, but tamoxifen does not have therapeutic potential as in breast cancer. Regarding the ERs subtype, ERα is expressed in tumors of epithelial and stromal origin, whereas ERβ is expressed in granulosa cell tumors point.\textsuperscript{59,60} With analogy to other hormone responsive tissues, the role of estrogen can be expected in etiology of ovarian tumors (at least those that arise from epithelial cells). Moreover, in the etiology of ovarian carcinogenesis gonadotropin stimulation is included, increasing the endogenous estrogen levels produced by the granulosa cells during the reproductive age.\textsuperscript{61} To date, the role of estrogen and ERs in ovarian cancer is not elucidated; this is an issue that may be partially related to the expression of different ER isoforms in ovarian carcinomas.\textsuperscript{62,63} However, in recent years the results of studies in women on long-term hormone replacement therapy showed an increased ovarian cancer incidence (together with increased mortality and other diseases related to hormonal treatment).\textsuperscript{64-66} In general, progestins and ERβ seem to play a protective role against the development of ovarian cancer, like in breast and prostate carcinomas. Decreases in ERβ expression or increased ERα/ERβ ratio is reported in ovarian cancer.\textsuperscript{67,68}
2.3.4. Endometrial cancer

Although ERβ is expressed in most cell types in the uterus, the ERα is the main mediator of estrogen action in this organ. The vast majority of sporadic endometrial carcinomas are classified as type I carcinomas and are estrogen-related. These estrogen-related endometrial carcinomas are associated with long-term exposure to estrogen in the absence of sufficient levels of progesterone. Tamoxifen, routinely used as adjuvant treatment in breast cancer and in breast cancer prevention trials in perimenopausal women, increases the risk of endometrial cancer. This is a consequence of the agonist effect of tamoxifen on ERs in the uterus. The fact that ERβ is also expressed in the uterus suggests its potential role in the pathogenesis of this disease. However, its role in endometrial carcinogenesis is rather unclear, partially due to the controversial reports of estrogen induced-ERβ transcriptional activity on uterine cell lines. 70, 71

3. Estrogen ligands

Different ligands may differ in their affinity for Estrogen receptor alpha (ERα) and beta (ERβ) isoforms of the estrogen receptor: ES binds equally well to both receptors, estrone and raloxifene bind preferentially to the alpha receptor, estriol, and genistein to the beta receptor. 72

![Figure 2: Structures of different ER ligands acting via Estrogen receptor.](image)

Subtype selective estrogen receptor modulators preferentially bind to either α- or the β-subtype of the receptor. In addition, the different estrogen receptor combinations may respond differently to various ligands, which may translate into tissue selective agonistic and antagonistic effects. 72 The ratio of α- to β-subtype concentration has been proposed to play a role in certain diseases. 73 The concept of selective estrogen receptor modulators is based on the ability to promote ER interactions with different proteins such as...
Chapter 1

as transcriptional co-activator or co-repressors. Furthermore, the ratio of co-activator to co-repressor protein varies in different tissues. As a consequence; the same ligand may be an agonist in some tissue (where co-activators predominate) while antagonistic in other tissues (where co-repressors dominate). Tamoxifen, for example, is an antagonist in breast and is, therefore, used as a breast cancer treatment, but an ER agonist in bone (thereby preventing osteoporosis) and a partial agonist in the endometrium (increasing the risk of uterine cancer).

Once activated by estrogen, the ER is able to translocate into the nucleus and bind to DNA to regulate the activity of different genes (i.e. it is a DNA-binding transcription factor). However, it also has additional functions independent of DNA binding. The ER's helix 12 domain plays a crucial role in determining interactions with co-activators and co-repressors and, therefore, the respective agonist or antagonist effect of the ligand. To date, the main efforts should be directed towards the investigation of the newer selective estrogen receptor modulators in the terms of their selective targeting to either of the ER receptors.

4. Docking

4.1. Docking Overview:

In order to get newer selective ER ligands with improve potency against the ER action and associated cancer; we have used molecular modeling techniques. It can replace the experimental studies of protein-ligand complexes by modeling their structures and binding affinities in silico.

![Figure 3: Schematic diagram illustrating the docking of a small molecule ligand to a protein receptor to produce a complex.](image)

Molecular docking is a key computational chemistry technique that is routinely applied to drug discovery. This is applied to study molecular binding and how molecules bind to the target receptor. It predicts the preferred orientation of one molecule to a second when bound to each other to form a stable complex.

4.2. Key stages of Molecular docking:

4.2. 1. Target selection and preparation:

For a successful designing of any kind of ligand or drug molecules, selection of an exact target is very important. Drug is a molecule which binds to a specific region of cells, called the receptor. Drug can act as a poison if it is not binding to the desired receptor. Toxicity is nothing but our designed molecules binding to the undesired receptor. Therefore, the designing of molecules needs a good knowledge of receptor
structure. The target structure should be experimentally determined by (1) X-ray crystallography (2) Nuclear magnetic resonance. Docking has been performed successfully against homology models, although the reliability of the docking results depends heavily on the quality and bias of the homology modeling. A target may adopt different conformations in the unbound and bound states, and with different classes of ligands. To overcome this and other problems, molecular dynamics has found an increasing number of applications in conjunction with molecular docking. These range from preparing the apo-form of the target before docking to accounting for receptor flexibility, solvent effects, and induced fit, to calculating binding free energies and ranking docked ligands.

4.2. 2. Ligand selection and preparation:

The type of ligands chosen for docking will depend on the target. For lead discovery, crude filters such as net charge, molecular weight, polar surface area, solubility, commercial availability, and price-per-compound can be applied to reduce the number of molecules to be docked. For lead optimization, filters such as similarity thresholds, pharmacophores, synthetic accessibility, and absorption, distribution, metabolism, excretion, and toxicology (ADME-Tox) properties are additionally applied.

Most docking tools treat ligands flexibly, with the exception of ring conformations. In general, the more rotatable bonds in a ligand, the more difficult and time consuming the docking will tend to be. This is because the size of the search space increases exponentially with the number of torsions. More highly branched torsion trees lead to more difficult searches than do linear torsion trees. Rotation of conjugated bonds, such as in amides, carbamates, ureas etc. should be limited. One strategy to explore ring flexibility is to perform conformational analysis on any ring-containing ligands before docking.

4.2. 3. Result Analysis:

Information on the binding site of the receptor (Ligand binding domain) is the most important criterion of molecular docking. If there is no previous information regarding the location of the binding site, then the translational search space should encompass the entire surface of the receptor. This is known as “blind docking”. If the docking tool cannot encompass the whole target, then probable sites such as cavities large enough to contain the ligand(s) should be investigated separately. If there is previous information, such as ligands with known binding modes, active site residues, or mutagenesis data, then the search space can be reduced to focus on the region of interest, thus, simplifying the search problem. Molecular docking involves computationally exploring a search space that is defined by the molecular representation used by the method and ranking candidate solutions to determine the best binding mode. Thus, docking requires both a search method and a scoring function. Scoring functions can be empirical, knowledge based, or molecular mechanics based (Force-Field-based). i.e. the Auto Dock scoring function is based on the molecular mechanics force field AMBER, with two additional terms: one to model the desolvation free energy change on binding, which is based on atomic solvation parameters; and one empirical term to model the loss of conformational entropy on binding. The individual contributions to the total energy of binding,
Chapter 1

namely van der Waals, hydrogen bonding, electrostatic, desolvation, and a number of rotatable bonds in the ligand, were treated as independent variables. Molecular docking involves computationally exploring a search space that is defined by the molecular representation used by the method and ranking candidate solutions to determine the best binding mode. Thus, docking requires both a search method and a scoring function.

A Typical Docking Work Flow

This flow chart shows the key steps common to all docking protocols. The 3-D structures for the target macromolecule and the small molecule must first be chosen, and then each structure must be prepared in accordance with the requirements of the docking method being used. Following the docking, the results must be analyzed, selecting the binding modes with the best scores.

5. Present work:

5.1. Preparation of database:

Much research has been carried out in discovery and research of novel selective estrogen receptor modulators. Non-steroidal compounds were designed as inhibitors based on triphenylethylene structure resembling with tamoxifen for targeting estrogen receptor alpha (ERα). A computer-aided strategy was used to screen the molecular database containing 5000 molecules. Initially compounds of the different structure with known biological relevance were collected and then, designed a library of molecules resembling with 4-hydroxytamoxifen (4-OHT), which indeed is the active metabolite of tamoxifen. Tamoxifen was first approved drug for breast cancer treatment in the 1970’s (National cancer institute). Tamoxifen is a non steroidal anti-estrogen that antagonizes the action of estrogen and is effective in both the treatment and prevention of breast cancer. The designed data base contains 2D and 3D structures of all these designed molecules.

5.2. Preparation of Estrogen Receptor α (ER-α) as receptor for docking

Estrogen receptor-α (ER-α) complexed with 4-OHT is taken as a role model to target the drug action. The 3D structure of the human estrogen receptor-α (ER-α) (PDBID-3ERT) with 4-OHT was obtained from protein data bank (Figure 4).
Figure 4: Structure of human estrogen receptor alpha co-existing with 4-OHT. The phenolic \textit{OH} group of 4-OHT forms simultaneous H-bonds with amino acid residue Glu 353 and Arg 394. The basic amino ethyl side chain residue of OHT involves in a charge transfer or electrostatic interaction with amino acid Asp 351.

5.3. Lead molecules design

In this present work, we replaced the ethyl and phenyl ring of 4-OHT by oxindole unit to test the binding capacity with ER-\(\alpha\) (Figure 5). Oxindole is a privileged heterocyclic moiety, known to possess a wide range of biological activities including anti-inflammatory, anti-angiogenic, anti cancer, tyrosine kinase A inhibition and cyclooxygenase inhibition. Around 14 structurally related compounds were chosen for docking assay after in silico screening of initial 5000 molecules. The initial screening was performed by V-Rocs software. The designed lead molecules were energy minimized by VEGA-ZZ and NAMD software. After energy minimization of lead molecules, they were subjected to docking process with estrogen receptor. Molecular dynamics study was also used to obtain the globally minimized state of ligand and receptor.

Figure 5: Design strategy of bis-arylidine oxindole based molecules.
We have started our journey derived from synthetic diphenolethyne viz. the designed oxindole based diphenolethyne by incorporating an oxindole unit into the bisphenolic or bis-arylidene moiety and finally mimicking some molecules resembling tamoxifen/4-OHT (Series-I). We have observed that, by the introduction of more rigidity into the molecules, the binding affinity to ER increases. Hence, we used rigid core nucleus in all our designed molecules. We also predicted that the length of the basic side chain can be varied from 2 to 3 carbon spacer. One can notice that these molecules have the base skeleton of the active metabolite, 4-OHT, bearing the crucial hydroxyl and dimethylamino groups, which are known to maintain excellent ER-recognition.

![Figure 6: 2D structures of the newly designed ER ligands (Series I and Series II).](image)

In Series-II, we introduced an elongated oxindole unit with additional substituent on nitrogen. The advantage of using this moiety lies in the fact that, it reduces the number of possible orientation in the active site of ER and thereby increases the selectivity. Additionally, it increases the number of H-bonding interaction. We always intended for the designed compounds to have a higher selectivity for cancer cells with minimal toxicological liability.

Furthermore, we have succeeded in designing some molecules without any basic dimethylamino alkyl side chain (Series-III A & Series-IIIB) and in due course, we found that this molecule also retains excellent ER selectivity. It has been known previously that the basic side chain of these types of molecules is indeed
crucial to maintain the ER targeting abilities. The electrostatic or charge transfer interaction between the
amino acid ASP351 and the cationic dimethylamino groups is the key interaction which defines the ER
selectivity. From the molecular modeling results, we have found that the loss of energy in absence of the
basic side chain can be compensated by introducing additional H-bonds.

Figure 7: 2D structures of the newly designed ER ligands (Series-III A & Series-IIIB).

In fact we are the first group to design and establish that basic dimethylamino alkyl side chain is not an
essential criterion for a molecule to retain ER selectivity. Molecular modeling results are well supported by
promising biological results.

5.4. Molecular Docking

FRED, Autodock 4.0, Autodock-vina, Schrodinger maestros and Discovery studio are the
docking programme, used to investigate the best ligand binding affinity with ER-α.
FRED performs an exhaustive docking by enumerating possible poses of the ligand in LBD by rigidly rotating and translating each conformer with in binding site and filtering the pose ensemble by rejecting the poses that do not fit well in the ligand binding domain (LBD). FRED requires 3 specials file format: (1) Refined receptor structure without the co-crystallized ligands and the water molecules inside a sphere of defined radius centered on the ligand previously present on the crystal structure. (2) A box of atoms subset defining the region in which the ligand will be docked (3) Ligand clusters for Ligand conformations to simulate its flexibility. It takes as input the protein structure and a multi-conformer representation of the ligand to be docked. In the exhaustive search, each ligand conformation is systematically rotated and translated within the active site at a resolution of 1 Å. Every poses that passes a bump check is scored. The top scoring poses, across all conformers of the ligand, are refined by testing nearby rotations and translations at a resolution of 0.5 Å. FRED and HYBRID use an exhaustive search algorithm to dock molecules. Both programs treat ligand conformers as rigid during the docking process, although ligand flexibility is implicitly included by docking multiple conformers of each ligand. The protein structure is also treated as rigid during the docking process for both FRED and HYBRID; however, HYBRID is capable of using multiple conformers of the target protein and therefore can account for protein flexibility as well. Chemgauss4 scoring functions for the final refinement and scoring of molecules and the Chemgauss3 scoring function was used to score molecules during the initial exhaustive search.

HYBRID’s docking and scoring algorithm is identical to FRED’s, except that the scoring function FRED uses during the exhaustive search, Chemgauss3, is replaced by a ligand-based scoring function, the Chemical Gaussian Overlay (CGO). CGO scores are based on how well the docked molecule matches the shape and 3D arrangement of chemical features of the crystallographic ligand bound to the active site, rather than how well the docked molecule complements the active site.

Autodock programme combines a rapid energy evaluation through pre-calculated grids of affinity potentials with a variety of search algorithms. Kollman charges were added to the receptor along with the solvation parameters. Similarly, a random number of generations, energy parameters, step size parameters and output format parameters were accepted from docking parameter data to find the best confirmation among the designed lead molecules. Grid was generated within a cubic box dimension. The search was based on the Lamarickian genetic algorithm (LGA) and the results were analyzed based on ranked clusters of compound confirmations. For each ligand, a docking experiment was conducted with 50 simulations and conformations of lead molecules are visualized using PyMOL software.

AutoDock Vina treat docking as a stochastic global optimization of the scoring function, pre-calculating grid maps (Vina does that internally), and some other implementation tricks, such as pre-calculating the interaction between every atom type pair at every distance. It also uses the same type of structure format (PDBQT) for maximum compatibility with auxiliary software.
6. Results and Discussion

Through in silico screen, 14 compounds which were chosen for testing docking assay led to 4 potent ligands. The selective ligands discovered in this study were promising drug candidates to be used as molecular probes to explore the differences between ERα and ERβ. All the designed compounds satisfied the Lipinski's rule of five with zero violations and also the octanol/water partition coefficient (miLogp), a useful parameter for predicting drug transport properties like absorption, bioavailability, permeability and penetration, as well as topological molecular polar surface area (TPSA), number of atoms, their molecular weight (MW), number of hydrogen donors and number of hydrogen acceptors. A topological parameter is number of rotatable bonds and it describes the molecular flexibility of these compounds. This parameter was calculated for the ten lead molecules that satisfy the ‘rule-of-5’ and it is found that all the molecules have rotatable bonds in the range of 3-5.

In our designing work, we have mostly used the autodock vina and FRED software. Autodock vina is used by the most scientific groups across the globe. We have also used the new generation software FRED to obtain addition results. Nowadays, FRED (openeye scientific groups) software has been used by many groups due to its consistency and accuracy in predicting results. Openeye scientific groups are connected with many drug designing companies and laboratories. There are lots of other softwares available under openeye scientific software distribution for an accurate initial screening.

The structure of ligand binding domain (LBD) of human estrogen receptor-α complex with 4-OHT was taken as template to model the ligands (PDB code 3ERT and resolution of 1.90Å). First 4-hydroxytamoxifen was digitally removed from the active site of the receptor and then successively the designed bio ligands were docked at the active site of the receptor. All the heavy atoms of the amino acids of the cavity wall were immobilized and the side chains of amino acid residue Met 343, Met 421 were liberated for some compounds (7-14) since they are bulkier in this region. This was justified as this part of the cavity has been shown to be flexible. Energy minimization was then carried out by MMFF (Merck Molecular Force Field). The energetically optimal positions of the bio ligands were calculated by chemgauss scoring (by FRED software version 2.2.5) which uses Gaussian functions. FRED performs an exhaustive docking by enumerating possible poses of the ligand in LBD by rigidly rotating and translating each conformer with in binding site and filtering the pose ensemble by rejecting the poses that does not fit well in the LBD, then rank all the poses by Chemgauss scoring which uses Gaussian function. Scoring value consists of energy contribution from acceptor, donor, metal, aromatic and steric desolvation. Steric desolvation is an entropic term which measures the loss of torsional degrees of freedom upon binding. The desolvation energy is proportional to the volume around the atoms that are exposed to the solvent. Score value consist all sorts of energy contribution from acceptor, donor, metal, aromatic and steric desolvation. FRED treats each conformer as rigid during docking process, although the docking process is effectively flexible with respect to the ligand because multiple conformers of each ligand are docked into the site.
Score value (Table 1) of the two isomers 11 and 12, as depicted in Figure 8, are -109.17 and -120.13 respectively, indicating a strong association at the active site of the receptor. A precise modeling study revealed that two best conformations are possible for 12. In first conformation (a), the phenolic OH-group of compound-12 binds to the amino acid residues Arg394 and Glu353 with H-bonds. In second conformation (b) of 12, the phenolic OH binds to Leu387 and the NH of oxindole ring forms H-bond with Gly419. The basic side chains in both conformations make a stable charge neutralizing ionic interaction with an acidic amino acid residue Asp351. The 11 isomer (also has two best conformer) binds in similar fashion (figure 8) like 12 except the NH of oxindole ring forms a weak H-bond with Gly419 (figure 8) and thus slightly lowers the score value.

**Figure 8**: (a) structure of compound 12 in active site of ER-α; (b) structure of compound 11 in active site of ER-α; (c) superimposed structure of 12 with 4-hydroxytamoxifen at the active site; (d) structure of compound 14 in active site of ER-α; (e) structure of compound 13 in active site of ER-α; (f) superimposed structure of 14 with 4-hydroxytamoxifen at the active site.

Compound 13 and 14 in Figure 8 also show good affinity for the receptor. Score value of 14 is -111.46 and that of 13 is -117.13 which also indicate favourable association at the active site of the receptor. The phenolic OH-group of 13 and 14 binds to Arg394 and Glu353 respectively with H-bond and the N atom of basic side chain forms a stable charge neutralizing ionic interaction with Asp351. The N alkylating group PMB with an acceptor oxygen atom forms a stable electrostatic interaction with His524. Bring Table 1

Score values of 7, 8, 9 and 10 are -66.78, -95.19, -104.66, and -116.37 respectively. In 10 (figure 9), one of the phenolic OH binds to Met 421 with H-bond and other phenolic OH binds to Met 343 with another H-bond. In 9 (Figure 9), one of the phenolic OH, and the other OH group at 5-substituted position of oxindole
ring form H-bond simultaneously with Glu 419 where as only one of the two phenolic OH binds to Gly 420 for 7 (Figure 9) and no other notable H-bonding interaction found, thus observed a lower score value for 7.

Figure 9: (a) structure of compound 8 in active site of ER-α; (b) structure of compound 9 in active site of ER-α; (c) structure of compound 10 in active site of ER-α; (d) structure of compound 7 in active site of ER-α; (e) structure of compound 9 in active site of ER-α; (f) superimposed structure of 10 with 4-hydroxytamoxifen at the active site.

Table-1: Molecular modeling results obtained from FRED software of newly designed ER ligands. Score value is the total of all sort of interaction namely H-bond, steric desolvation, VdW interaction, electrostatic interaction, and chemgauss score is taken as default.

<table>
<thead>
<tr>
<th>Molecular code:</th>
<th>Total Score Value (chemgauss score)</th>
<th>H-bonding amino acids residue</th>
<th>Electrostatic interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-72.87</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-51.30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-58.36</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-66.96</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-71.21</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>
Another set of binding free energy calculation were done by Autodock 4.2 and Autodock vina to obtain the affinities of the ligands for the active site. The side chains of the amino acid residues Met 343, Met 421 were liberated and that of His524 was made free. All the heavy atoms of the other amino acids of the cavity wall were immobilized. A routine energy minimization was then carried out using the Merck Molecular Force Field (MMFF), to determine the best position for the bio ligand under determination. Autodock 4.2 uses an improved semi empirical free energy force field with an improved thermodynamic model of the binding process. Unlike Autodock 4.2, the scoring function in Autodock vina based upon an iterated local search global optimizer for optimization procedure, where the success of each step consists of a mutation and local optimization. Binding energies $\Delta E$ were calculated using the method $\Delta E = E_{\text{complex}} - E_{\text{protein}} - E_{\text{ligand}}$. For all compounds, binding to the ER is thermodynamically favoured, as evidenced by the negative binding energies and binding affinities for the ligand-receptor complex (Table-2). In compound 11 and 12 (figure 9), the phenolic OH forms H-bond with Arg394 and Glu353. The basic side chain interacts with Asp351 and the NH of indoline makes H bond with Gly420 (Auto dock shows the involvement of Glu419). Binding affinity values of 11 and 12 (figure 8) isomers are -11.1 kcal/mol and -10.5 kcal/mol.

<table>
<thead>
<tr>
<th></th>
<th>E (kcal/mol)</th>
<th>Side Chains of Amino Acids</th>
<th>Protein</th>
<th>Ligand</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>-92.61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-66.78</td>
<td>Met 421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-95.19</td>
<td>Met 421</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-104.66</td>
<td>Glu 419, Gly 420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-116.37</td>
<td>Met 421, Met 343</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-109.17</td>
<td>-9.15kcal/mol</td>
<td>Asp351</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-120.13</td>
<td>-8.79 kcal/mol</td>
<td>Asp351</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-111.46</td>
<td>-6.35 kcal/mol</td>
<td>His524, Asp351</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-117.13</td>
<td>Arg394, Glu353, Leu387, Gly419</td>
<td>His524, Asp351</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-hydroxy tamoxifen</td>
<td></td>
<td>Arg 394, Glu353</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bis-phenol</td>
<td>Arg394, Glu353</td>
<td>Asp351</td>
<td></td>
</tr>
</tbody>
</table>
(binding energies are - 9.15 kcal/mol and - 8.79 kcal/mol produced by auto dock 4.2 software). Compound 6 has similar interaction like 11 apart from the missing interaction with an acidic amino acid residue Asp351 due to lack of basic side chain. Binding affinity of 9 is -9.7 kcal/mol (binding energy is -8.61 kcal/mol). The binding affinities are slightly weakened for compound 13 due to the missing H-bonding interaction at indoline NH. The phenolic OH in compound 13 and 14 also binds with Arg394 and Glu353. The basic side chain has charge neutralizing ionic interaction with Asp351. The N alkylating group PMB with an acceptor oxygen atom forms a stable electrostatic interaction with Ile424.

Figure 10: (a) structure of compound 12 in active site of ER-α; (b) structure of compound 11 in active site of ER-α; (c) superimposed structure of 12 with 4-hydroxytamoxifen at the active site; (d) structure of compound 14 in active site of ER-α; (e) structure of compound 13 in active site of ER-α; (f) superimposed structure of 14 with 4-hydroxytamoxifen at the active site;
Chapter 1

**Figure 11:** (a) structure of compound 9 in active site of ER-α; (b) superimposed structure of 9 with 4-hydroxytamoxifen at the active site; (c) structure of compound 10 in active site of ER-α; (d) superimposed structure of 10 with 4-hydroxytamoxifen at the active site; (e) structure of compound 8 in active site of ER-α; (f) structure of compound 6 in active site of ER-α.

The binding affinity values of 7, 8, 9 and 10 are -8.2, -9.2, -8.8, and -9.2 kcal/mol respectively (while the observed binding energies are -8.0, -9.22, -9.24 and -9.11 kcal/mol respectively). In compound 10, one of the phenolic OH forms weak H-bonds with Arg394 and Glu353 whereas the other phenolic OH at 5-substituted indoline ring binds to His524 by another H-bond. One of the phenolic OH group in 9 forms H-bond with Met421 and the other OH group forms another H-bond with Glu419 (Auto dock shows involvement of His524). Same binding situation arises for 8; but here only one of the two phenolic OH groups binds with His524.

Initially, we have performed some molecular modeling studies on a series of Bis-phenolic compounds (synthetic diphenolethylene, 1,1-bis(4’-hydroxyphenyl)-2-phenylbut-1-ene). Preliminary results shows that bis-phenolic compounds were not antagonist in nature rather agonist in nature. Results also showed a poor binding in ligand binding domain (LBD) of ER and no evidence of any H-bonding or any other strong electrostatic interaction.

On the other hand, compound 1 and 6 of our designed ligand are bis-phenolic, but they can act as an antagonist on ER-α without having the prototypical antagonist basic side chain. We tried to superimpose the ligand with 4-OHT in its crystal structure (Figure 12) and we found that one of the phenolic OH of the 6 forms weak H-bonds with Arg 394A and Glu 353A with a distance of 3.0 Å and 2.7 Å respectively which is a characteristic interaction in 4-OHT ligand binding domain. The oxindole unit has one additional donor site (NH) and an acceptor site (=O) which is in favorable position to form H-bonds and electrostatic interaction with any amino acids in close proximity.

**Figure 12:** Superimposed structure of compound 6 with 4-OHT at the active site of human ER-α. The NH of oxindole forms a H-bonds with the oxygen of Gly420 at a distance of 3.0 Å and the oxygen of oxindole ring have electrostatic and phenyl ring has hydrophobic interaction with Leu387A, Leu 384, Met 343A, Leu 346A and Ala 350A, Ile 424 amino acids residues.
Chapter 1

Substituting with a halogen at 5 positions of oxindole unit (compound - 2, 3, 4, 5) does not increase binding energies values reveals the fact that there is no other strong interaction with halogen atoms with amino acids in close proximity. With a substitution at _NH’ of oxindole ring shows a substantial decrease in binding energies and in score values of 1. These observations also support the fact that, engagement of _NH’ H-bonds of oxindole moiety is important or key interaction for determining the binding modes of these bio ligands. When we try to super impose 1 with 4-hydroxytamoxifen it is shown that the characteristic H-bond with phenolic OH and Arg 394Aand Glu 353 is destroyed and thus with a consequence the binding energies decreases.

Table-2: Molecular modeling results obtained from Autodock and Vina software

<table>
<thead>
<tr>
<th>Molecular code:</th>
<th>Inhibition constant/Dissociation constant (Ki/Kd)</th>
<th>Binding Affinities</th>
<th>H-bonding amino acids residue</th>
<th>Electrostatic interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.57 uM</td>
<td>-9.2 kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1.28uM</td>
<td>-8.0kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>1.28uM</td>
<td>-8.0kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>605nM</td>
<td>-8.2kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>690nM</td>
<td>-8.6kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>632nM</td>
<td>-9.7 kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>Gly420</td>
</tr>
<tr>
<td>7</td>
<td>3.67uM</td>
<td>-8.2kcal/mol</td>
<td>Leu387</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>917nM</td>
<td>-9.2kcal/mol</td>
<td>His524</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>3.04uM</td>
<td>-8.8kcal/mol</td>
<td>Met421, Glu419</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>640nM</td>
<td>-9.2kcal/mol</td>
<td>His524, Arg394 , Glu353</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>212nM</td>
<td>-11.1kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>Asp 351, Gly420</td>
</tr>
<tr>
<td>12</td>
<td>362nM</td>
<td>-9.5kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>Asp 351, Gly420</td>
</tr>
<tr>
<td>13</td>
<td>4.49uM</td>
<td>-7.1kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>Asp 351, Ile 424</td>
</tr>
<tr>
<td>14</td>
<td>362nM</td>
<td>-10.5kcal/mol</td>
<td>Arg394 , Glu353</td>
<td>Asp 351, Ile 424</td>
</tr>
</tbody>
</table>
### Chapter 1

<table>
<thead>
<tr>
<th>4-hydroxy tamoxifen</th>
<th>5nM</th>
<th>-10.5 kJ/mol</th>
<th>Arg394, Glu353, Asp351</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bis-phenol</td>
<td>4.45 μM</td>
<td>-7.0 kJ/mol</td>
<td>Arg394, Glu353, Asp351</td>
</tr>
</tbody>
</table>

**Conclusion:**

Finally, the designed molecules were synthesized adopting different methodologies. The goal of ligand—protein docking is to predict the predominant binding mode(s) of a ligand with a protein of known three-dimensional structure. Successful docking methods search high-dimensional spaces effectively and use a scoring function that correctly ranks candidate dockings. Docking can be used to perform virtual screening on large libraries of compounds, rank the results, and propose structural hypotheses of how the ligands inhibit the target, which is invaluable in lead optimization. The setting up of the input structures for the docking is just as important as the docking itself; however, analyzing the results of stochastic search methods can sometimes become unclear.

**References:**


84. [http://autodock.scripps.edu/](http://autodock.scripps.edu/).


86. [https://www.schrodinger.com/maestro](https://www.schrodinger.com/maestro).


88. [https://www.pymol.org/](https://www.pymol.org/).