The Estrogen Receptor

Julie Davey Dalsgaard Lund

Bachelor's Thesis

February 2005

Department of Chemistry
University of Aarhus
Denmark

Contents

Chapter 1

Introduction

This project is about the estrogen receptor and the different effects estrogen-like molecules can have on the receptor. The estrogen receptor is located in the nucleus or the cytoplasma of the cell. This means that the receptor can have a direct effect on the DNA transcription, when an estrogen enters the cell and binds to the estrogen receptor. In this case, the receptor is released from an inhibitory protein and undergo a conformational change and can now have an effect on specific DNA sequences; see [3]. The binding of other compounds than the natural estrogens can have similar effects. This means that other compounds can have an effect on the estrogen receptor and cause disruptions in the human body. For that reason, it is very important to be able to determine whether a given molecule has any estrogenic activity or not. Unfortunately, there are many different estrogens and since their structure do not need to resemble the structure of estradiol – the natural agonist for the estrogen receptor – for the molecule to have the same effects, the experimental process is very time-consuming. The primary goal of this project is to determine if it is possible to use computer-aided docking simulations to speed up the process. However, since I have examined pesticides in addition to known agonists and antagonists for both alpha and beta estrogen receptors, the secondary goal is determine if any of the examined pesticides have effects on either of the estrogen receptors. I have performed the docking simulations using the computer program Maestro[1]. The program docks the different compounds in the ligand binding domains of different estrogen receptors, which I have chosen using the Protein Data Bank[2] (PDB).

[1] http://www.schrodinger.com/Products/maestro.html
[2] http://www.pdb.org/

Chapter 2

Estrogen Receptor

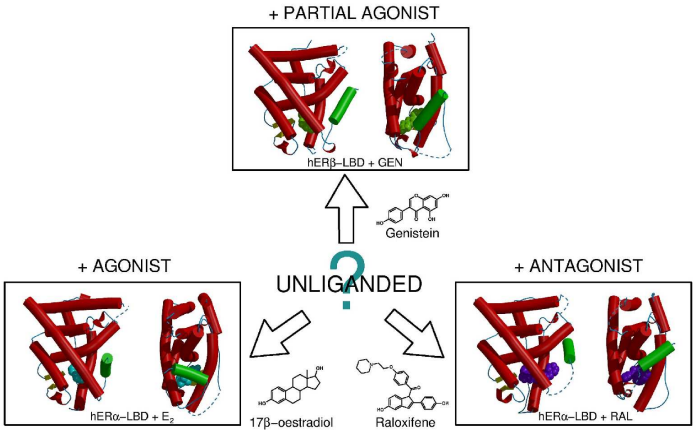
The estrogen receptor (ER) is a member of a large family of nuclear receptor transcription factors with specific domains associated with transactivation, DNA, and ligand binding; see [11]. The ER is a twelve helix protein, that is located in the nucleus of the cell, where it functions as a ligand-activated transcriptional regulator. Since the receptor is in the nucleus of the cell, it can only be affected by molecules that are small enough to pass through the cell membrane. Such molecules affect ERs to form dimers and affect the DNA through transcription of DNA within the nucleus; see [6].

The ER is located in several places in the female body, where it is the target of drugs against menopause and some cancer types as described in [11]. The receptor is located in the hypothalamo pituitary axis, breast tissue, liver, uterus, vagina, and in bone tissue. The location of the receptor affects the response of the receptor: ERs in different locations yields different physiological effects. Another thing that influences the effects of the ER is the kind of estrogens which bind to the receptor.

The binding pocket of the ER is a hydrophobic pocket. When a ligand binds to the ligand binding domain (LBD), there is a conformational change in the receptor which eventually leads to activation or deactivation of responsive genes. Every ligand interacts with a unique set of amino-acid residues in the hormone binding cavity and induces a specific orientation of helix 12. When an antagonist binds in the cavity, the bulky side chain of the ligand will prevent helix 12 from covering the bound ligand. On the other hand, when an agonist binds in the binding cavity, helix 12 covers the agonist and the binding cavity. Figure 2.1 on the following page shows the conformational state of the ER when binding an agonist and an antagonist.

There are two different ERs: ERα and ERβ; see [11]. ERα and ERβ have different responses and they are located in different tissue. It was thought for many years that there was only one ER, but it was noticed that estradiol had an effect on tissue where no ERαs were located. Today, it is known that there are more than one ER and this explains why estrogens can have an effect on tissue without ERαs. Some ligands bind to both receptors but have different effects on them. A ligand can be an agonist for ERα and an antagonist for ERβ. The structures of ERα and ERβ are only 47 percent identical. There is a difference in ligand binding ability and transactivational ability in the two ERs. The N-terminal transacti-

Figure 2.1 The conformations of ERα and ERβ LBD in the presence of an agonist (E_2), a partial agonist (GEN), and an antagonist (RAL). Helix 12 is shown as a green rod. The bound ligands are shown in space-filling form. Reprinted from [12].



vation (AF-1) domains have no similarities, but the ERs have quite similar DNA and ligand binding domains. The C-terminal LBD is multi-functional: It contains the ligand recognition site and regions for receptor dimerization and ligand-dependent (AF-2) transactivation; see [11].

The most examined receptor is ERα; see [8]. ERαs are found in the liver, uterus, vagina, bone tissue, hypothalamo pituitary axis, and in breast cancer tumors. The structure of ERα is different depending on whether an agonist or an antagonist is bound to it. When an agonist is bound to the receptor, helix 12 covers the binding site and the bound agonist. When an antagonist is bound to the receptor, helix 12 is flipped away from the binding site, because the antagonist makes it impossible for helix 12 to fold into place. For the receptor to have an effect, helix 12 has to be in the right position; see figure 2.1.

ERβ is found in different parts of the body than ERα. ERβs are located in the prostate, testis, ovary, and in some areas of the brain. Male and female bodies react differently when exposed to natural estrogens: Women are at risk of developing breast and endometrial cancer, whereas men can get gynecomastia. Furthermore, estrogens can disturbe the normal function of the male hypothalamus-hypophyseal-gonadal axis, which may lead to decreased libido, impotence, drop in androgen levels, and drop in sperm count; see [13]. If human embryos are disposed to estrogens, they may develop cryptorchidism, persistance of Mellerian in men, enlarged prostate, vaginal adrenosis, malformations of the female genital tract, and clear cell adenocarcinoma of the vagina.

ERβ also has helix 12, but when an agonist binds to the receptor, helix 12 does not overlay the LBD as it is the case for ERα; see figure 2.1 and [11]. The position of helix 12 in ERβ with a bound agonist is closer to the position of helix 12 in the antagonist position of ERα; that is pointing away from the LBD.

Chapter 3

Ligands

A ligand is any molecule that binds specifically to another. There are many known ligands for both ERα and ERβ, but there are also many unknown ones. The structure of the unknown ligands do not need to resemble the structure of estradiol for the molecule to have the same effects, and thus it is a time-consuming task to determine the effects a given compound has on the ER. It is very important to understand the effects of a compound, before it is marketed. The fact that there are many compounds on the market today which after thorough investigations are removed due to unfortunate side-effects, illustrates that it is not always easy to determine all the effects. To highlight key physiological and chemical aspects of the ligands I have used in my experiments, this chapter contains a brief description of these ligands.

3.1 Estradiol and Estriol

Estradiol is the natural agonist for the ER – its structure is illustrated by figure 3.1 on the next page. Estradiol is synthesized from cholesterol and secreted by the ovaries. It is not only present in women; it is also found in men, in which case it is secreted from their testes. Estradiol is the female sex hormone and it is essential in reproductive endocrinoligy and for growth. In case of an over-production – or an exposure to estrogen-like compounds – there is an increased risk of devoloping breast and endometrial cancer. When women reach menopause their ability to produce estradiol drops and some women develop osteoporosis because estradiol is essential for bone growth and maintainance. Estradiol lowers the concentration of low-density lipoprotein (LDL) and raises the concentration of high-density lipoprotein (HDL). Because HDL removes excess cholesterol from blood and tissue, women have a lower rate of coronary artery disease (CAD) than men. When a woman reach menopause, the risk of CAD is raised, if the woman is not in hormone replacement therapy; see [15] pages 622–624.

Estriol is a natural estrogen, which is produced during pregnancy. It is the major estrogen produced in the normal human fetus. There are indications that estriol may be less carcinogenic than estradiol. Research has shown that estriol does not induce endometrial growth to the extent of the other estrogens, even at doses where estriol is effective for the relief of postmenopausal symptoms; see [1]. The structure of estriol is shown in figure 3.2 on the following page.

Figure 3.1 The structure of estradiol

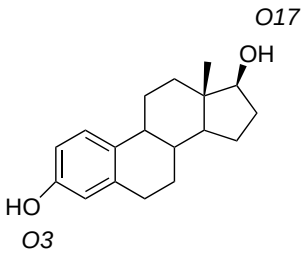
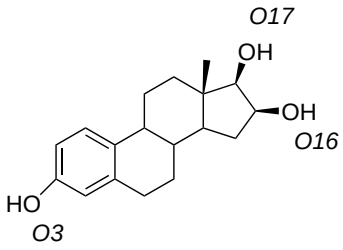


Figure 3.2 The structure of estriol

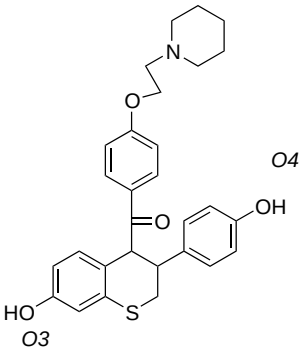


3.2 Raloxifen

Raloxifen is an antagonist for ERα and it was first developed to prevent breast cancer. Today, it is used to treat and prevent osteoporosis in menopause women and it is being tested for prevention and treatment of breast cancer. The structure of raloxifen is shown in figure 3.3 on the next page. Raloxifen is a very efficient antagonist in reproductive tissue, but it is a partial agonist[1] in bone tissue and lowers cholesterol levels in the blood. When raloxifen binds in the LBD of the ER, helix 12 is not able to close the LBD, because of the arm of the raloxifen molecule. When helix 12 does not come into place in the LBD, there is no activation of the receptor and no effect on DNA trancription; see [8] and [9].

[1] an agonist that yields a response which is less than maximal

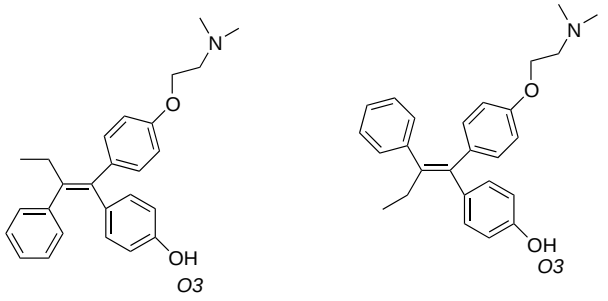
Figure 3.3 The structure of raloxifen



3.3 Tamoxifen E and Z

Tamoxifen is not the active compound in the body, but rather the prodrug for the active metabolit 4-hydroxytamoxifen (OHT). Tamoxifen is an agonist in some tissue and an antagonist in other tissue. In breast tissue, it is an antagonist, whereas in bone and uterine tissue, it is an agonist. Tamoxifen is the number one choice for the treatment of all stages of breast cancer; see [9]. Unfortunately, tamoxifen treatment has side-effects in the shape of hot flashes, vaginal bleeding, and skin aches; see [10] page 632. There are two different kinds of tamoxifen: Tamoxifen E and tamoxifen Z – structural differences between the two kinds are shown in figure 3.4. Tamoxifen E is an agonist and tamoxifen Z is an antagonist in ERα; see [8].

Figure 3.4 The structures of 4-hydroxytamoxifen E and Z

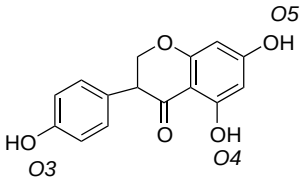


3.4 Genistein

Genistein is a partial agonist in ERβ and it is found in soya beans and soy products. Genistein is a phyto-estrogen – a plant-derived non-steroidal compound that has estrogen-like

biological activity. Phyto-estrogens are primarily produced as bacteracidal and fungicidal agents and may have a beneficial effect on and reduce the risk of breast and prostate cancer, heart disease and relieve symptoms associated with menopause. Genistein binds in ERβ in almost the same way as estradiol binds in ERα – see the structure of genistein in figure 3.5. Genistein is a partial agonist in ERβ; see [11].

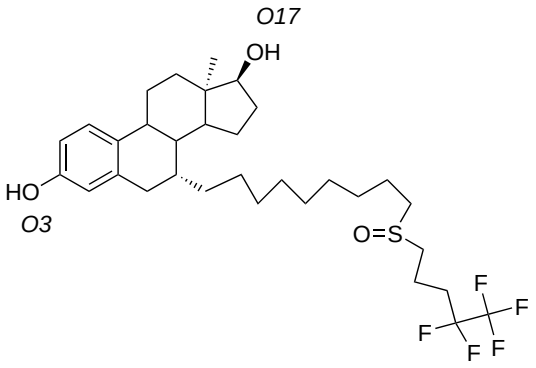
Figure 3.5 The structure of genistein



3.5 Fulvestrant

Fulvestrant (ICI-182780) is an antagonist for both ERα and ERβ – its structure is shown in figure 3.6. It is used to treat advanced ER-positive breast cancer in post-menopausal women. It is used primarily for cancer that has relapsed either during or following treatment with standard anti-estrogen treatments, such as tamoxifen. Unfortunately, it has a negative effect on bone density; see [9].

Figure 3.6 The structure of fulvestrant



3.6 Pesticides

I have examined the estrogen or anti-estrogen properties of 26 different pesticides. The selected set of pesticides and their structures are shown in figures A.1 on page 35 and A.2 on page 36. Structurally, none of the examined pesticides look like estrogens and they do not

have the characteristic steroid rings. This means that it is not possible to decide if they are estrogens or anti-estrogens – or if they do not bind to the ER at all – simply by looking at them. Pesticides do not need to look like estrogens or androgens to have an effect on ERs.

The problem with pesticides is that there are a lot on the market and it is only now that we are discovering that many of them are dangerous for the environment and for both animals and humans. Pesticides can have a whole range of severe effects and this is why article [13] argues that it is extremely important to find and remove the pesticides that have estrogenic effect from the market and to make sure that no new ones enter the market. The pesticides examined in this paper are used for many different purposes. The purposes include fungicides, insecticides, herbizides, bio-regulating, and plant growth regulating. The pesticides methiocarb, pirimicarb, propamocarb, chlorpyrifos, tolchlorfosmethyl, prochloraz, fenarioml, endosulfan, dieldrin, and chlorothalonil are all suspected of having an effect on the ER. They may cause cryptorchidism in sons of female gardners, reduce fertility in female green house workers and may cause an increased risk of breast cancer in females with increased concentrations of these compounds in their blood; see [3].

Chapter 4

Methods

This chapter is an overview of the computer program used for my docking simulations on ERs. I have used the docking program Glide[1] made by Schrödinger in 1998, which makes Glide one of the newest docking programs on the market. It is also one of the best docking programs on the market today. I have used version 2.0 of Glide, but because there are no articles on Glide 2.0 available, I have described how Glide 2.5 works. There have been made some improvements in Glide 2.5 compared to Glide 2.0, but the basics are the same.

4.1 Overview of Glide

Glide was made by Schrödinger, an American drug discovery technology development company, which was founded by Richard Friesner and William Goddard III in 1990. Schrödinger has developed several versions of Glide and I have chosen to use Glide 2.0 for my calculations in this bachelor thesis. Glide is one of many docking programs on the market, but one of the better ones. This is at least partially due to the fact that there is an ongoing development of Glide. At the time of writing, the latest version of Glide is version 3.0. The graphical user interface (GUI) for Glide is called Maestro – it lets the user interact with a three-dimensional model of the molecules. The results of docking simulations on ERs in Glide have not improved much from version 2.0 to 2.5, so I expect that my experiments could be repeated using version 2.5 without major differences in the outcome; see [7]. Glide is a grid-based ligand docking program. In Glide, you have a flexible ligand and a rigid protein. The average docking time is one minute per ligand with 0–10 rotable bonds on modest hardware. The binding mode predictions are good compared to other docking programs and the predictions of binding affinity are reasonably consistent with experimental data, but there is room for improvement. The docking method approaches a complete systematic search of the ligand conformation, orientation, and position. As described in [5], this docking method is more stable than stochastic methods.

[1] http://www.schrodinger.com/Products/glide.html

4.2 Scoring Functions

The scoring function in Glide is based on the ChemScore function of Eldrige et al as described in [4]:

$$\begin{aligned}\Delta G_{bind} =& C_0 + C_{lipo} \sum f(r_{lr}) + \\ & C_{hbond} \sum g(\Delta r) h(\Delta \alpha) + \\ & C_{metal} \sum f(r_{lm}) + C_{rotb} H_{rotb}\end{aligned}$$

The first term is the standard concentration. The second term is the lipophilic ligand-atom/receptor-atom pairs and the third term is the hydrogen bonds between the ligand and the receptor. The fourth term is the metal interactions and the fifth term is the rotations in the ligand. The factors f, g, and h are functions that gives full scores (1.00) for distances and angles within the normal limits of hydrogen bonds and hydrogen angles and a reduced score $(0.00 - 1.00)$ when the distances and angles are outside the limits.

Glide 2.5 uses a modified version of the ChemScore function called GlideScore; see [5] where this description originates. There are two forms of GlideScore in version 2.5 of Glide: The GlideScore 2.5 SP used by Standard-Precision (SP) Glide and the GlideScore 2.5 XP used by Extra-Precision (XP) Glide. SP is a "softer" function which is adept at identifying ligands that have a reasonable propensity to bind even in cases in which the Glide pose has significant imperfections. XP is a harder function that puts servere penalties on poses that violate established physical chemistry principles such as that charged and strongly polar groups be adequately exposed to the solvent. Since SP has loose parameters, allows more errors, and is faster to compute than XP, SP is often used for the first screening. The high penalties XP puts on wrongly docked ligands makes it ideal when more precisely docked ligands are needed and for removing false positives. The GlideScore function (SP) is as follows:

$$\begin{aligned}\Delta G_{bind} =& C_{lipo-lipo} \sum f(r_{lr}) + \\ & C_{hbond-neut-neut} \sum g(\Delta r) h(\Delta \alpha) + \\ & C_{hbond-neut-charged} \sum g(\Delta r) h(\Delta \alpha) + \\ & C_{hbond-charged-charged} \sum g(\Delta r) h(\Delta \alpha) + \\ & C_{max-metal-ion} \sum f(r_{lm}) + C_{rotb} H_{rotb} + \\ & C_{polar-phob} V_{polar-phob} + C_{coul} E_{coul} + \\ & C_{vdW} E_{vdW} + \text{solvation terms}\end{aligned}$$

The first term is the lipophilic interactions as in the ChemScore function. The hydrogen bonds are not as in the ChemScore function. Here the function is separated into differently weighted compounds, where the weight depends on whether the hydrogen bonds are between two neutral atoms, neutral and charged, or two charged atoms. The GlideScore also takes the metal bonds and the rotationally bonds into account. The seventh term of the GlideScore function handles the case where a polar – but not hydrogen bonding atom – is

found in a hydrophobic region. The function also takes Coulomb and van der Waals interactions between ligand and receptor into account. The last term in the function is based on the solvent that the docking is performed in; see [5].

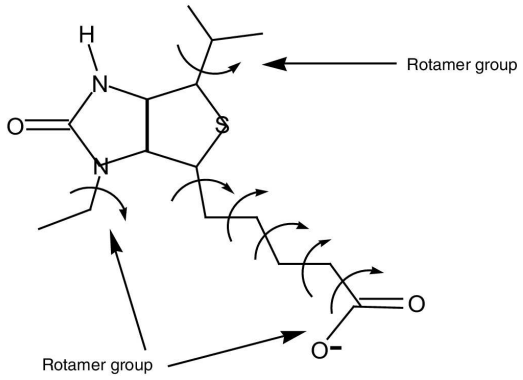
There is another scoring function in Glide: The E-model. The E-model is a combination of the GlideScore, molecule mechanics between the ligand and the protein, and the ligand strain. The E-model is used to choose the correct pose of the ligand in the LBD. The best scoring pose of each ligand – the pose with the lowest E-model score – are shown ranked by their GlideScore in the Glide Pose Viewer.

The common conditions for both E-model and GlideScore are that the lower the value, the better. The G-score must be lower than -8 before that you can be certain that the ligand binds and has an optimal effect on the receptor. Even though the G-score is higher than -8, the ligand can still have an effect on the receptor.

4.3 Ligand Conformations

In Glide, the ligand is divided into a core region and a number of rotatable groups. Each rotatable group is attached to the core of the ligand by at rotatable bond, but does not itself contain rotatable bonds. There are typically no more than 500 core conformations even with large and flexible ligands. The core plus all possible rotamer group conformations attached to the core are enumerated and docked as a single object in Glide. See figure 4.1.

Figure 4.1 Definition of core and rotamer groups. The four central torsions are part of the core. The methyl groups are not considered rotatable. Reprinted from [2].



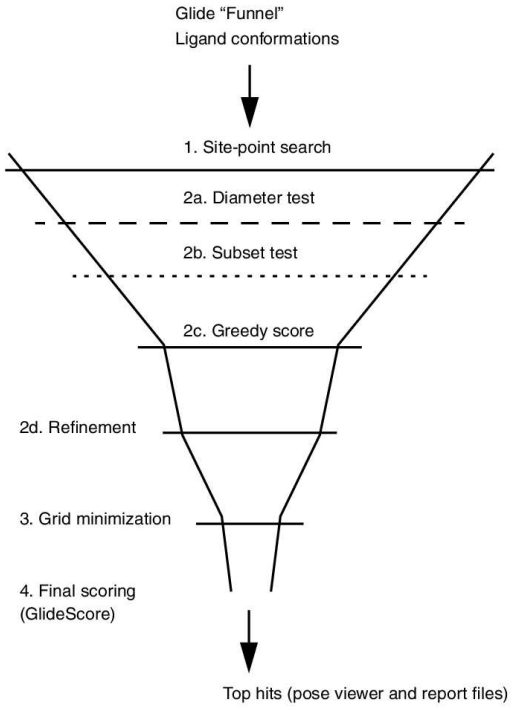
The steps involved in computing the GlideScore for a ligand/receptor pair is shown below. Figure 4.2 on the following page also illustrates the computation process.

1. The active site is searched for all possible positions for the core.
2. The atoms that lie within a specific distance of the ligand-diameter axis are examined.
3. Rotation of the ligand diameter is considered and interactions between the ligand and receptor are looked into.

4. The grid is minimized.

5. The final scoring is done and the GlideScore is recorded.

Figure 4.2 The Glide docking hierarchy illustrates how the search for ligand conformations, orientations, and positions is performed by step-wise minimization of the search space. Reprinted from [2].



Chapter 5

Results

In this chapter, I present and discuss the experimental results of my bachelor work. Some of the ligands I have examined are treated individually; others in logical ligand groups. The presentation focuses on how the ligands bind in the different ER sites and the the differences between known agonists and antagonists are discussed.

5.1 Selection of ER Structures

I have selected several different ER structures for my dockings. I have chosen both ERα and ERβ receptors, and I have used both agonist sites and antagonist sites for both types of receptors. To show the effect of having helix 12 far away from the LBD, I have also experimented with an empty ERα. The following table summarizes the configurations I have based my experiments on:

	Agonist	**Antagonist**	**Empty**
ERα	1GWR	3ERT, 1ERR	1A52
ERβ	1QKM	1HJ1	–

Some of the ligands described in chapter 3 are the co-crystalized ligand for one or more of the examined ERs. The following table shows the relations between these ligands and the ERs:

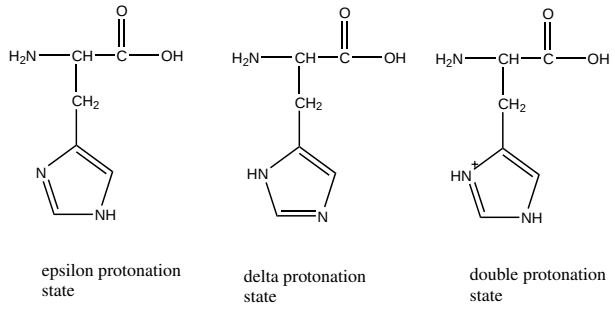
Structure	**1GWR**	**1A52**	**1ERR**	**3ERT**	**1QKM**	**1HJ1**
Ligand	Estradiol	Estradiol	Raloxifen	OHT-Z	Genistein	Fulvestrant

5.2 Protonation States of Histidin 524

The histidin residue in the LBD have one of the following protonation states: Epsilon, delta, or double. Figure 5.1 on the next page shows the structures of histidin in the different states.

Before I could perform any of my dockings, I had to decide on which protonation state to use. I docked each co-crystalized ligand in its receptor with three different protonation

Figure 5.1 Protonation states of histidin in the LBD



states. The following table shows the result of docking each ligand in the three different protonation states of histidin. The first column shows the different protonation state of histidin in the ERs. The second column shows the GlideScore and the third the E-Model. The fourth column is the number of hydrogen bonds between the ligand and the receptor and the last column shows the root mean square deviation (RMSD) values. The RMSD values are a measurement of how well the co-crystalized ligand docks in the protein. Lower values of RMSD indicates better results. For a docking to be good, the RMSD value should be lower than 2. Because the RMSD value is lower than 2 in all the calculations, it implies that all the ligands fit in their proteins.

Protonation states in 1GWR	**G-score**	**E-model**	**Hydrogen bonds**	**RMSD**
Epsilon	−8.55	−78.6	3	0.087869
Double	−8.63	−80.9	3	0.230453
Delta	−8.66	−78.9	3	0.279307
Protonation states in 1ERR				
Epsilon	−12.22	−116.4	3	1.448016
Double	−12.18	−117.2	3	1.456627
Delta	−11.98	−114.1	3	1.455565
Protonation states in 3ERT				
Epsilon	−10.04	−94.8	3	1.269553
Double	−10.07	−94.7	3	1.315565
Delta	−9.89	−93.9	2	1.295402
Protonation states in 1A52				
Epsilon	−8.37	−77.9	3	0.198767
Double	−8.08	−80.4	3	0.252196
Delta	−8.29	−77.7	3	0.106866
Protonation states in 1QKM				
Epsilon	−8.31	−82.3	3	0.237557
Double	−8.65	−83.4	3	0.237557
Delta	−8.44	−81.2	3	0.189662
Protonation states in 1HJ1				
Epsilon	−9.88	−89.7	2	1.533915
Double	−9.70	−87.7	2	1.661408
Delta	−9.41	−82.3	2	1.852222

The different proteins have different co-crystalized ligands. In 1GWR and 1A52, the co-crystalized ligand is estradiol, in 1ERR it is raloxifen, and in 3ERT it is OHT. In the two ERβs, 1QKM and 1HJ1, the co-crystalized ligands are genistein for 1QKM and fulvestrant for 1HJ1. The protonation state does not seem to have an impact on the number of hydrogen bonds in the different ERs; except for the delta protonation state of 3ERT.

For 1HJ1, the best protonation state is the epsilon protonation state of histidin. In almost all of the other dockings, the double protonation state has the lowest E-model and the differences in the G-scores are quite small. The differences in the RMSD values are also quite small and even though none of the dockings for the double protonation state shows the lowest RMSD value, it is still the best candidate because of the E-model. The reason for this is that it is the E-model scores that vary the most and that the double protonation state has very good E-model scores.

To ease comparisons between the different dockings, I have chosen to use the double protonation state for all of them. This is a reasonable choice, because the G-scores, E-models, RMSD values, and the number of hydrogen bonds are very similar for the different protonation states.

5.3 Intermolecular Interactions

Before I start going through my results I will discuss some governing conditions used for determining hydrogen bonds, the conditions for E-model and G-score and the important residues in the LBD. Hydrogen bonds are relative weak interactions between a hydrogen bond donor and a hydrogen bond acceptor. The hydrogen atom is partly shared between to electro-negative atoms such as nitrogen or oxygen. The hydrogen bond donor is the group that includes the atom to which the hydrogen has the strongest binding and the hydrogen itself. The hydrogen bond acceptor is the atom that has the weakest binding to the hydrogen atom.

Experiments have shown that the length of a hydrogen bond is $2.4 - 3.5$ Å between hetero-atoms and that an optimal hydrogen bond has an angle of 180° between O$\cdots$H$\cdots$N or between O$\cdots$H$\cdots$O. For that reason, I have chosen to disregard any hydrogen bond reported by Glide that does not have a length between $2.4 - 3.5$ Å and an angle of $180^\circ \pm 30^\circ$.

There are three very important residues in the LBD: Histidin 524 (His524), arginine 394 (Arg394), and glutamate 353 (Glu353). In the ERβ, the residues have different numbers, but this is just because ERβ lack some N-terminal residues compared to ERα.

5.4 Results of the Dockings

The following sections contain descriptions of the different dockings and discussions of the results. To verify the experimental results against the expected results, the descriptions are followed by conclusions on where each ligands docks best.

5.4.1 Known ER Ligands

The following table shows the docking results for the known ER ligands in the different sites. The left column contains the different ligands and the top shows the different ER structures. For each ligand, there are three rows with the docking results: The E-model, G-score, and

the number of hydrogen bonds as reported in the Glide Pose Viewer. Notice, that 1GWRa is identical to 1GWR except that the His524 in the LBD is mutated to an alanine 524 (Ala524).

Ligand	**Results**	**1ERR**	**3ERT**	**1GWR**	**1GWRa**	**1A52**	**1QKM**	**1HJ1**
Estradiol	G-score	−8.02	−7.28	−8.63	−7.47	−8.08	−8.13	−7.45
	E-model	−38.5	−20.2	−80.9	−64.4	−80.4	−47.8	−68.0
	Hbnd	3	3	3	3	3	2	3
Raloxifen	G-score	−12.18	−9.95	–	–	−10.71	–	−10.22
	E-model	−117.2	−79.2	–	–	−45.5	–	−90.2
	Hbnd	3	2	–	–	3	–	4
Estriol	G-score	−8.02	−7.28	−8.63	−7.47	−8.08	−8.13	−7.45
	E-model	−38.5	−20.2	−80.9	−64.4	−80.4	−47.8	−68.0
	Hbnd	5	2	3	3	3	3	3
OHT-Z	G-score	−9.77	−10.07	–	–	−10.39	–	−7.75
	E-model	−67.4	−94.7	–	–	−84.5	–	−68.8
	Hbnd	4	3	–	–	3	–	1
OHT-E	G-score	−8.64	−9.25	–	–	−9.36	–	−7.26
	E-model	−58.2	−54.0	–	–	−59.9	–	−62.8
	Hbnd	1	2	–	–	2	–	1
Genistein	G-score	−8.30	−7.86	−7.88	−6.91	−7.95	−8.31	−7.33
	E-model	−70.5	−53.3	−64.9	−61.0	−67.1	−82.9	−66.7
	Hbnd	6	4	4	3	4	3	5
Fulvestrant	G-score	−8.36	−10.18	–	–	−6.44	–	−9.70
	E-model	−18.8	−48.5	–	–	−5.6	–	−87.7
	Hbnd	1	3	–	–	–	–	2

The table shows that all ligands have been docked in all ER structures. Notice, that some of the hydrogen bonds reported in the table above do not satisfy the criteria set forth in section 5.3. Appendix B gives more details on the hydrogen bonds. All the co-crystalized ligands bind best in their natural ER, OHT-E binds best in the empty receptor (1A52), and estriol binds best in the agonist receptor (1GWR).

The agonists estradiol and estriol bind best in the agonist receptor in both the ERα and ERβ, whereas the partial agonist genistein binds best in the agonist ERβ. In ERα, it binds best in the antagonist site, 1ERR. According to literature, OHT-E is an agonist, but I have not been able to show this, because OHT-E does not dock in any of the agonist sites. Again, this is probably because I have not been able to perform a protein-flexible docking. The reason the agonists bind better in agonist sites is that the binding sites are smaller and the ligands can interact with helix 12 and the residues in the LBD. Therefore, the binding in antagonist sites are weaker, because of the lack of hydrophobic interactions with helix 12. The reason genistein acts differently is probably that in ERβ, helix 12 does not cover the LBD. For that reason, genistein prefers the antagonist site.

The antagonists OHT-Z, raloxifen, and fulvestrant do not bind in the agonist sites in either ERα or ERβ. This is probably because the protein is not flexible during the docking and therefore it is not possible for the antagonists to enter into the receptor; helix 12 is in the way, because of the arms of the antagonists. Furthermore, the agonist binding site is smaller than the antagonist binding site.

The results for 1GWR and 1GWRa with respect to the dockings of estradiol, estriol, and genistein shows that His524 has an effect: All the ligands have a hydrogen bond to His524; see appendix B. Furthermore, the G-scores and E-models are higher for the binding in 1GWRa compared to the binding in 1GWR.

5.4.1.1 Estradiol

Figure 5.2 shows how estradiol lies in both an agonist site (1GWR) and in an antagonist site (1ERR). Estradiol lies in the same way in both ERα and ERβ which implies that this must be the most favorable orientation. In figure 5.2, it is possible to see the difference in the position of His524 in an agonist site compared to an antagonist site.

Figure 5.2 Estradiol in 1GWR (pink) and 1ERR (green)

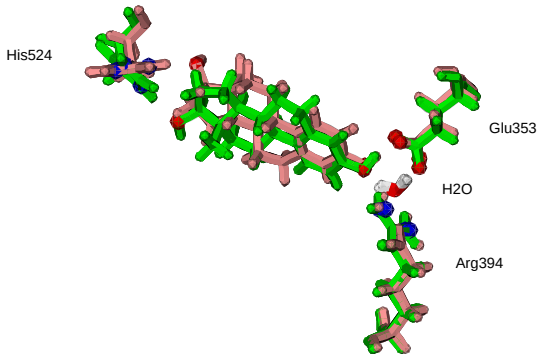


Figure 5.3 shows how estradiol lies in its natural receptor and the mutated receptor. The green receptor is the one where His524 is mutated to Ala524. The figure shows that estradiol lies in the same place for both receptors, and that the only difference is that the hydrogen bond to His524 is not formed in the mutated case.

Figure 5.3 Estradiol in 1GWR (pink) and 1GWRa (green)

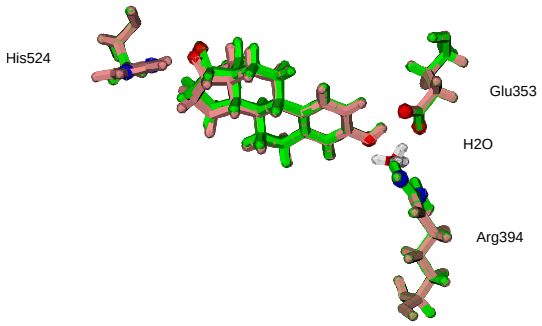
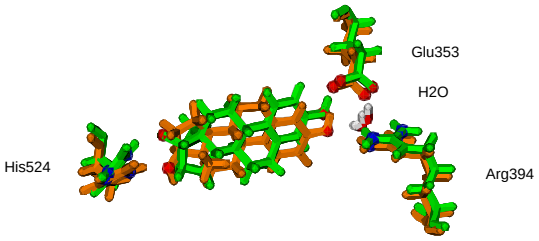


Figure 5.4 shows the binding of estradiol in ERβ. Again, estradiol lies in the same way in the agonist and antagonist sites. There is not a big difference in binding affinity between the agonist and antagonist site, and this may be because the two sites are quite alike as seen in the figure.

Figure 5.4 Estradiol in 1HJ1 (orange) and 1QKM (green)

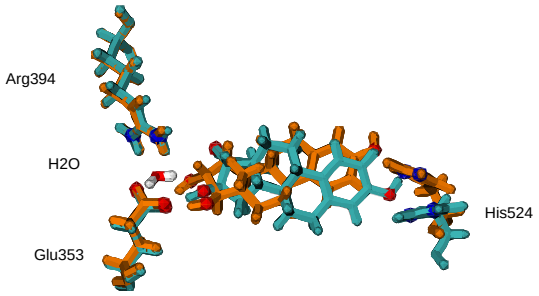


To conclude, estradiol binds best in its natural site (1GWR); as expected. Furthermore, His524 does have an effect on the binding affinity. It is not possible to conclude in which of the ERβ receptors estradiol binds best, because the agonist site has the best G-score and the antagonist site has the best E-model.

5.4.1.2 Estriol

Structurally, estriol looks very much like estradiol. The only difference is the extra hydroxy-group on the D-ring. Estriol have many hydrogen bonds in all the sites according to the tables in appendix B. About half of them could actually exist because they have an angle within $180° \pm 30°$. Estriol lies in different ways in ERα depending on if it is an agonist or antagonist site. Figure 5.5 shows the binding in the antagonist sites and figure 5.6 on the next page shows the binding in the agonist and empty sites.

Figure 5.5 Estriol in 1ERR (orange) and 3ERT (cyan)



In figure 5.5, the A ring of estriol faces His524, whereas in figure 5.6, the D ring is facing His524. In figure 5.7 on the following page, which shows the binding of estriol in ERβ, the D ring of estriol is facing the His524. Furthermore, it is possible to see the differences in the position of His524 in the agonist and antagonist site.

Figure 5.6 Estriol in 1GWR (pink), 1GWRa (green), and 1A52 (gray)

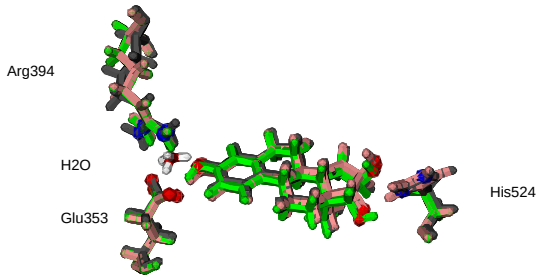
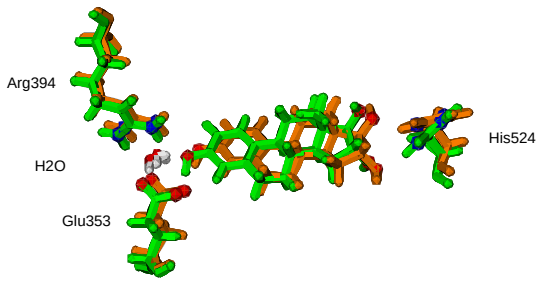


Figure 5.7 Estriol in 1QKM (green) and 1HJ1 (orange)



The conclusion is that estriol binds in all the ER structures, but for ERα, it binds best in 1GWR and for ERβ it binds best in 1HJ1.

5.4.1.3 Raloxifen

Figure 5.8 on the next page shows raloxifen in its natural ERα active site and in the empty ERα. Raloxifen in the empty ERα is a bit more stretched out and this may be because helix 12 is pointing away from the protein giving a larger LBD cavity. The stretching of the arm of raloxifen does have an effect on the results of the docking, because the E-model of raloxifen in 1A52 is only -45.5 compared to -117.2 for 1ERR. It is obvious that the position of the antagonist arm is of importance.

Figure 5.9 on the following page shows raloxifen in the two antagonist sites. Raloxifen lies in the same way in the two receptors and here one can see the difference in the position of the residues. His524 does not lie in the same place in the two receptors and this is probably because helix 12 is further away from the binding site in 1ERR than in 3ERT. Figure 5.10 on the next page shows the binding of raloxifen in ERβ.

In conclusion, raloxifen binds well in all antagonist sites, but it binds best in 1ERR, where it is the co-crystalized ligand. It does not bind in any of the agonist sites and this is because there is no room for raloxifen in the sites, since the docking is not protein-flexible.

Figure 5.8 Raloxifen in 1ERR (orange) and 1A52 (gray)

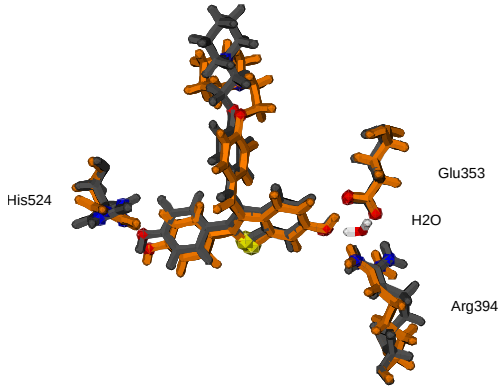


Figure 5.9 Raloxifen in 1ERR (orange) and 3ERT (cyan)

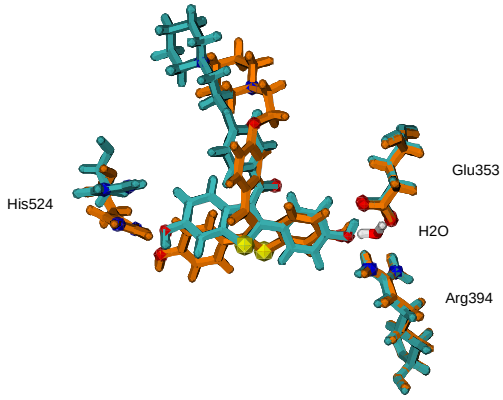
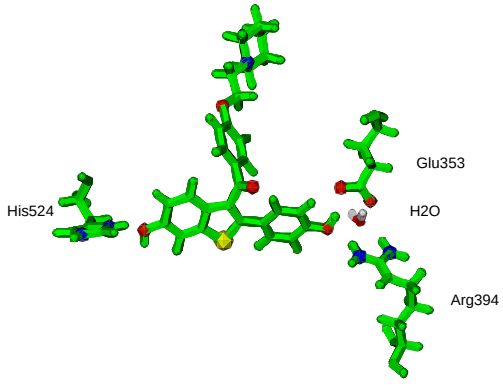


Figure 5.10 Raloxifen in 1HJ1 (green)



5.4.1.4 4-hydroxytamoxifen E and Z

The figures 5.11 to 5.14 (some of which can be found on the next page) show OHT-E and Z in the different antagonist sites. The two OHT conformations in the different sites have their arms and the phenol groups in the same places. There is a difference in the number of hydrogen bonds. The Z conformation has more bonds than the E conformation and this is because the phenol ring of the Z conformation lies closer to His524, Arg394, and Glu353 residues, which makes it easier to form hydrogen bonds. None of the conformations bind in the agonist sites; this is because they are too big to enter the active site cavity, when the protein is rigid in the docking. So even though the E conformation should be an agonist, I have not been able to show this.

Figure 5.11 OHT-Z (gray) and OHT-E (green) in 1A52

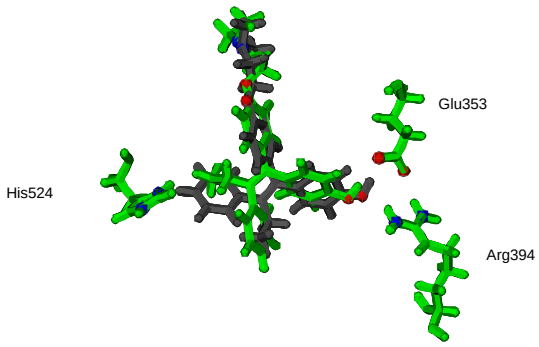
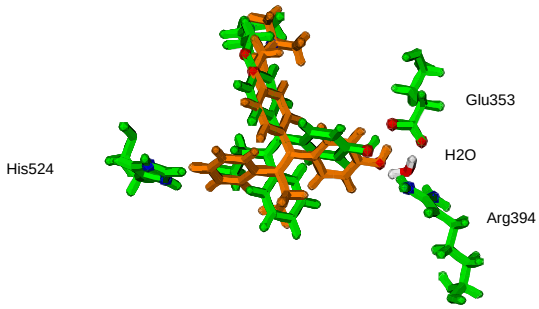


Figure 5.12 OHT-Z (orange) and OHT-E (green) in 1ERR



In conclusion, OHT-E binds best in the empty site (1A52) and OHT-Z binds best in 3ERT, where it is the co-crystalized ligand. Overall, OHT-Z binds better than OHT-E when considering the G-scores, E-models, and the number of hydrogen bonds.

Figure 5.13 OHT-Z (cyan) and OHT-E (green) in 3ERT

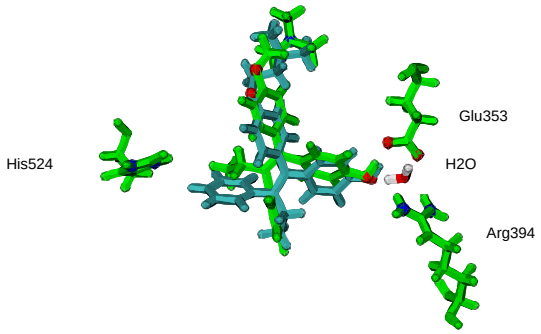
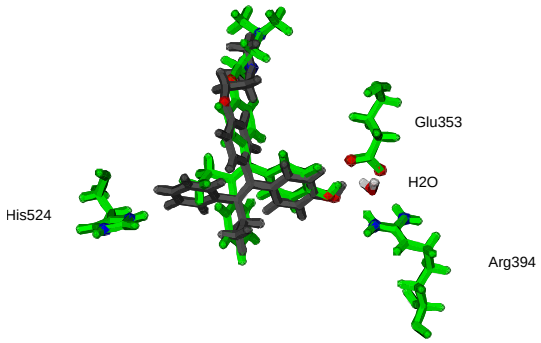


Figure 5.14 OHT-Z (gray) and OHT-E (green) in 1HJ1



5.4.1.5 Genistein

Figure 5.15 on the following page shows genistein in ERβ; both in 1QKM – the agonist site – and in 1HJ1 – the antagonist site. Genistein is the co-crystalized ligand in 1QKM and has the same orientation in the two receptors. It has more favorable hydrogen bonds in 1QKM than in 1HJ1. In the ERα, genistein has one orientation in 1A52, 1GWR, and 1GWRa (see figure 5.16) and another in 1ERR and 3ERT (see figure 5.17). This seems to indicate that the size of the active cavity is important for the orientation of genistein, when binding in the receptor. Again, the importance of His524 is shown, because genistein binds better in 1GWR, where it has a hydrogen bond to His524, than in 1GWRa.

To conclude, genistein in ERβ binds best in 1QKM, where it is the co-crystalized ligand, whereas in ERα it binds best in 1ERR. This may be because genistein is a partial agonist and it is a bit too big to fit in the rigid agonist sites.

Figure 5.15 Genistein in 1HJ1 (orange) and 1QKM (green)

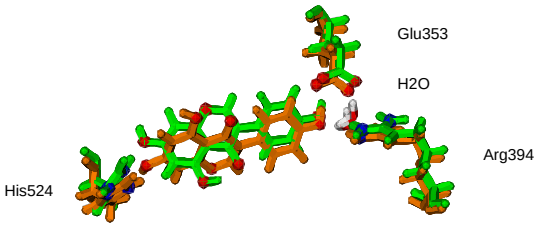


Figure 5.16 Genistein in 1GWR (pink), 1GWRa (green), and 1A52 (gray)

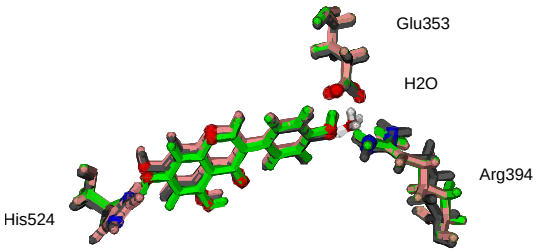
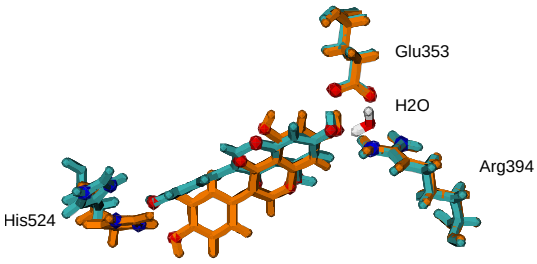


Figure 5.17 Genistein in 1ERR (orange) and 3ERT (cyan)



5.4.1.6 Fulvestrant

Fulvestrant is the co-crystalized ligand in 1HJ1 and in ERβ it only binds here. The reason why fulvestrant does not bind in 1QKM is probably that the LBD is not big enough when the protein is rigid. Fulvestrant has only one hydrogen bond in 1HJ1; the other bonds do not have favorable angles; see the table in appendix B.7. In ERα, fulvestrant binds best in 3ERT, but it also binds in the other sites. The E-models in ERα are quite low compared to the ERβ, and this may be because of the arm of fulvestrant and the lack of room in ERα. Figure 5.18 on the next page shows fulvestrant in ERβ and figures 5.19 on the following page and 5.20 on page 25 show fulvestrant in ERα. In all ERα sites, the arm is bent but the situation is worst in 1A52. In 1A52, fulvestrant has its arm in the active site and the ring system on the

surface of the receptor, so in this case it does not form any hydrogen bonds and the E-model is very low. The reason why it is possible for the ligand to bind in this way is probably that helix 12 is far away from the active site.

Figure 5.18 Fulvestrant in 1HJ1 (green)

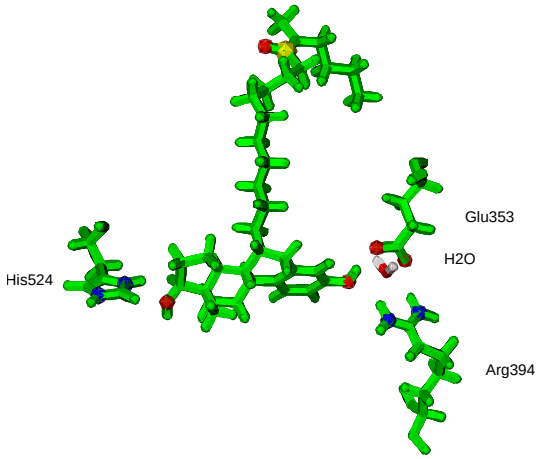
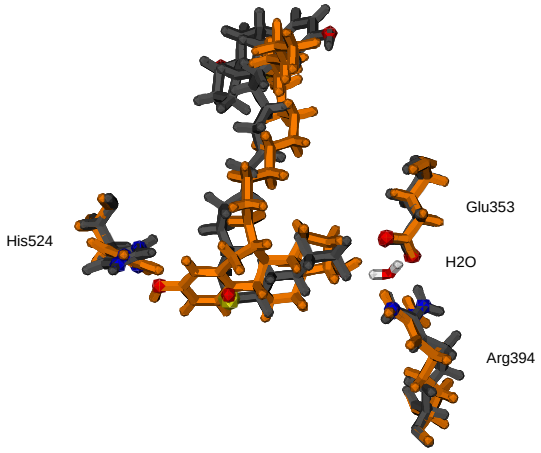
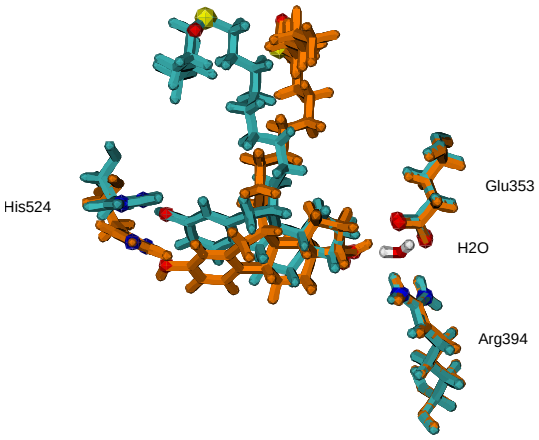


Figure 5.19 Fulvestrant in 1ERR (orange) and 1A52 (gray)



The conclusion for fulvestrant is that it only binds in one ERβ (1HJ1), but it binds in all the ERα antagonist sites. Fulvestrant does not bind in any of the agonist sites, because it is too big. It is not possible to fit the ligand in the smaller LBD, because helix 12 is in the way.

Figure 5.20 Fulvestrant in 1ERR (orange) and 3ERT (cyan)



5.4.2 Pesticides

This section contains the docking results for a set of pesticides, also tested experimentally by Vadim V. Sumbayev; see [14]. I have performed these dockings to investigate whether or not these pesticides have an effect of the ER. For each compound there is a G-score, an E-model, and the number of hydrogen bonds.

5.4.2.1 Primary Pesticides

The primary set of pesticides contain DDE and DDE-like compounds, along with other important pesticides. These are the pesticides which have been examined in a recent article; see [14].

None of the compounds bind as strongly as the co-crystalized ligands, but this does not mean that they do not have an effect on the ER. The differences in binding affinity for the different ERs are quite small, so in some cases it can be difficult to determine where the compound binds best.

DDE is the most examined pesticide and therefore I have focused my experiments on DDE and the compounds with similar structures. In ERα, DDE binds best in 3ERT and 1A52 according to the G-score and E-model. In ERβ, DDE binds best in 1QKM; again according to both the G-score and E-model. DDE does not have any hydrogen bonds in any of the sites, because it does not have any hydrogen bond donors or acceptors. The orientation of DDE is the same in all the sites.

The DDE-like compounds are TCDD, endosulfan, dieldrin, vinclozolin, iprodione, and paclobutrazol. My results show that the only ones that bind like DDE is vinclozolin and iprodione. TCDD, endosulfan, deildrin, and paclobutrazol have an agonistic effect on ERα. DDE, fenarimol, vinclozolin, and iprodione have an antagonistic effect on the ERα. For prochloraz, it is not possible to determine which effect it has, because it binds equally well in the agonist and antagonist sites.

Ligand	Result	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
DDE	G-score	−6.05	−5.97	−5.80	−5.90	−6.00	−6.58	−5.22
	E-model	−36.8	−45.1	−35.6	−33.6	−42.7	−46.8	−38.2
	Hbnd	0	0	0	0	0	0	0
TCDD	G-score	–	−3.77	−5.64	−5.87	−5.64	−6.32	−5.27
	E-model	–	−30.7	−18.6	−30.5	−25.9	−33.0	−40.0
	Hbnd	–	0	0	0	0	0	0
Fenarimol	G-score	−6.46	−4.75	−6.95	−6.94	−6.90	–	−6.07
	E-model	−48.0	−37.7	−19.8	−13.6	−40.6	–	−43.8
	Hbnd	0	0	0	0	0	–	1
Prochloraz	G-score	−7.19	–	−7.11	−6.81	−6.39	–	−6.46
	E-model	−52.7	–	−51.5	−45.1	−30.9	–	−52.1
	Hbnd	1	–	0	0	0	–	0
Endosulfan	G-score	−5.43	−5.40	−5.64	−5.54	−5.52	–	−5.18
	E-model	−33.8	−33.5	−41.3	−44.6	−42.8	–	−47.1
	Hbnd	0	0	0	0	0	–	1
Dieldrin	G-score	−5.59	−5.76	−5.58	−5.16	−5.52	–	−5.63
	E-model	−31.1	−30.9	−38.6	−29.5	−37.4	–	−34.0
	Hbnd	0	0	0	0	1	–	1
Vinclozolin	G-score	−6.35	−5.43	−5.93	−6.38	−5.96	−6.78	−5.45
	E-model	−44.1	−31.2	−34.4	−27.5	−36.6	−40.9	−37.2
	Hbnd	0	0	0	0	0	0	0
Iprodione	G-score	−6.51	−6.46	−7.00	−7.19	−7.24	−8.09	−5.80
	E-model	−50.5	−51.2	−33.3	−26.0	−38.9	−37.3	−47.1
	Hbnd	0	1	0	0	1	0	0
Paclobutrazol	G-score	−6.08	−5.81	−6.04	−5.84	−5.91	−6.80	−5.26
	E-model	−38.9	−42.4	−41.1	−43.8	−29.9	−34.0	−37.6
	Hbnd	0	0	0	0	1	0	0

The docking results are equally good for all the ERαs. DDE, TCDD, and paclobutrazol have an agonistic effect on ERβ. Fenarimol, prochloraz, endosulfan, dieldrin, vinclozolin and iprodione have an antagonistic effect on ERβ. The molecules are all quite small, but they do not resemble any of the known ligands for any of the ERs. This may be why they do not have a very low G-score and therefore not a very good binding. These compounds do not have many hydrogen bonds and this is because many of the ligands do not have hydrogen bond donors and acceptors near the three residues in the LBD.

5.4.2.2 Secondary Pesticides

The set of secondary pesticides have not been as thoroughly tested as the primary pesticides. Nevertheless, it is important to investigate the effects these pesticides may have on ERs.

Ligand	Resultat	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
Pirimicarb	G-score	−5.48	−3.65	−4.49	−4.92	−4.76	−5.20	−4.09
	E-model	−31.4	−27.9	−37.1	−37.9	−32.0	−15.9	−28.2
Fenpropathrin	G-score	−6.55	−5.35	–	–	−3.96	–	−4.72
	E-model	−38.9	−35.8	–	–	−26.4	–	−33.7
Phosetyl-aluminium	G-score	−3.33	−2.56	−4.69	−4.51	−4.81	−2.18	−2.95
	E-model	−45.1	−44.1	−43.0	−51.2	−61.3	−32.7	−52.5
Propamocarb	G-score	−4.56	−3.59	−3.94	−3.81	−4.32	−3.57	−3.39
	E-model	−37.7	−25.3	−29.8	−29.2	−35.3	−19.8	−34.0
Deltamethrin	G-score	−5.19	−5.45	−8.33	–	−7.56	–	−6.15
	E-model	−40.5	−53.1	−4.7	–	−27.3	–	−43.4
Diaminozide	G-score	−4.21	−3.91	−3.71	−3.47	−3.93	−4.08	−3.32
	E-model	−39.6	−38.3	−36.6	−37.4	−37.9	−40.1	−37.3
Methiocarb	G-score	−5.09	−4.45	−4.75	−4.41	−5.04	−5.27	−4.07
	E-model	−39.8	−39.7	−41.5	−40.1	−41.8	−43.3	−36.9
Dichlorvos	G-score	−3.82	−3.09	−3.29	−2.99	−3.31	−3.52	−3, 20
	E-model	−45.9	−30.4	−40.4	−26.7	−39.0	−33.6	−37.2
Chlormequat	G-score	−2.15	−1.66	−2.05	−1.78	−1.95	−2.18	−2.15
	E-model	−17.7	−26.4	−12.3	−17.3	−15.2	−19.4	−22.4
Metoicarbsulfoxide	G-score	−5.10	−4.18	−5.09	−4.81	−5.48	−5.05	−4.83
	E-model	−42.9	−39.2	−48.4	−40.6	−50.0	−43.1	−50.9
Tolclofosmethyl	G-score	−4.37	−4.58	−4.65	−4.52	−4.78	−5.15	−4.84
	E-model	−33.8	−38.2	−41.4	−40.3	−43.1	−46.3	−41.9
Ethephon	G-score	−3.36	−4.01	−3.71	−2.71	−2.50	−3.96	−2.87
	E-model	−51.8	−51.8	−44.2	−44.6	−50.2	−49.1	−47.2
Methomyl	G-score	−4.62	−4.14	−4.26	−4.47	−4.32	−4.16	−3.56
	E-model	−40.1	−38.7	−40.8	−37.9	−36.7	−42.2	−36.9
Chlorothalonil	G-score	−4.33	−4.41	−4.68	−4.39	−4.94	−5.19	−4.19
	E-model	−28.9	−18.2	−34.0	−24.0	−34.5	−35.2	−31.1
Tribenuron methyl	G-score	−6.65	−6.95	−8.35	−7.38	−7.72	–	−7.35
	E-model	−45.4	−49.2	−31.8	−29.6	−50.4	–	−56.9
Chlorpyrifos	G-score	−4.80	−2.76	−5.20	−4.89	−5.36	−5.58	−4.46
	E-model	−39.4	−28.2	−41.5	−34.7	−42.0	−30.1	−40.4
Dimethoate	G-score	−4.26	−3.99	−3.71	−3.48	−3.69	−4.24	−3.60
	E-model	−42.2	−38.5	−41.2	−33.2	−38.3	−33.0	−36.0

The table shows that the most common result of the dockings is that the ligand does not dock very well. Almost all the results show low G-scores and this means that the ligands do not have a direct effect on any of the ERs. The ligands may bind outside the active site, because the LBD can be too small or too big for the ligand. Tribenuron methyl and deltamethrin are the ligands that bind best. They both have G-scores less than -8. The E-model for the binding of deltamethrin in 1GWR is only -4.7 and this is because the deltamethrin is twisted and not in its optimal conformation. In 1A52, the G-score is -7.56 and the E-model is -27.3. The G-score is not quite as good as for the binding in 1GWR, but the E-model is much better. This is because the ligand is not as twisted as in 1GWR. Deltamethrin does not bind in 1GWRa and 1QKM sites. This must be because of lack of room. It is very interesting

that deltamethrin binds in 1GWR and not in 1GWRa, but I have no explanation for this behavior. The conclusion for deltamethrin is that it bind best in 1A52, because here both the G-score and the E-model are good. In 1GWR, the G-score is better but because the E-model is so low, I believe that deltamethrin binds best in 1A52.

Tribenuron methyl also has good G-scores and E-models; it appears that tribenurom methyl binds best in 1GWR or 1A52. The binding in 1GWR is the one with the best G-score and the binding in 1A52 is the one with the best E-model. Tribenuron methyl binds quite well in all the sites except in the 1QKM site. The E-models for the 1GWR and the 1GWRa sites are not as good as for the rest of the sites. This may be because these are the smallest sites. Compared to the position of the co-crystalized ligand, tribenuron methyl lies along the arm of raloxifen. In 1A52, the empty site, where the helix 12 is flipped away from the binding site, the ligand binds quite well and the structure of the ligand is very similar to the structure of the ligand before docking.

Some of the secondary pesticides are much smaller than the co-crystalized ligands. When the compounds are small, their G-scores are closer to zero and therefore the binding is weaker and the effect on the ER is smaller. It is possible that the effect on the ER is increased if two compounds bind in the LBD at the same time. The bigger compounds seem the have a greater effect on the ERs. Most of the compounds bind best in 1ERR or 1QKM, so I conclude that the compounds have an antagonistic effect on ERα and an agonistic effect on the ERβ, except for the two compounds with G-scores below -8 that have an agonistic effect on ERα.

Chapter 6

Conclusions

The purpose of this project was to determine if it was possible to use computer-aided docking simulations to speed up the process of determining if molecules have either estrogenic or anti-estrogenic effects. I wanted to show that by using computer-aided docking simulations, it is possible to get known ligands to dock in their natural ER. Furthermore, I wanted to examine if pesticides have an effect on the ER.

I expected to find that the known ligands would dock best in their natural ERs, because all the ligands and receptors are co-crystalized complexes. Furthermore, I expected that estriol would dock like estradiol, because the two ligands are very similar, and that OHT-E would bind like an agonist.

I found that the co-crystalized ligands all bound best in their natural sites, just as I had expected. Estriol binds best in the agonist site in 1GWR, which is also as expected because of the structural similarities between estriol and estradiol. OHT-E does not bind in any of the agonist sites, and this is probably because I have not performed a docking with both a flexible protein and a flexible ligand. When the protein is rigid, OHT-E is too big to fit in the small agonist sites. Therefore, it only binds in the antagonist and empty sites. I did not have any expectations to how the pesticides would bind, because their structures are so different from the known ligands. None of the pesticides bind very well, but it is not possible to conclude much, because their G-scores are so low. If the pesticides bind, it seems that most of them would give an antagonistic effect on ERα and ERβ.

There are many ways of making computer-aided docking simulations better. The most obvious one is to dock flexible ligands into flexible proteins. The problem with this approach is that it is very time consuming and therefore very expensive. As computers get faster and cheaper, it becomes easier to make dockings with flexible proteins more affordable.

All in all, computer-aided docking simulations is a promising novel way of determining the effect different compounds can have on receptors in the human body. I succeeded in reproducing the established results for the known ligands and found that pesticides may effect the estrogen receptor. The successful reproduction of the results for the known ligands gives hope to the future of this kind of research, because it creates new ways of testing compounds before sending them on the market.

Chapter 7

Experiments

This chapter describes the experimental process used to obtain the results presented in this thesis. Even though I have performed the calculations several times, this chapter only describes the final set of calculations in details.

My first attempt was made using the protein-ligand complexes with all water molecules present. This turned out to be a bad idea, because the water molecules interacted with non-ligand parts of the complexes and made the computing time longer.

The second time I performed the calculations, I did not consider the different protonation states of His524, and for that reason the calculations were all performed using the epsilon protonation state. As described in section 5.2, the choice of protonation state has an impact on the results. For that reason, I chose to redo the calculations based on the best protonation state, which I determined was the double protonation state. The final set of calculations upon which my results are based were thus performed using the double protonation state without superfluous water molecules

7.1 Selection of ER Structures

The protein data bank (PDB) is a collection of crystal structures for proteins with bound ligands and co-activators. I searched the PDB to find crystal structures of ERα and ERβ with bound agonists and antagonists. I found about 20 different ER structures with different ligands bound and from different tissue. The ER structures I chose had the ligands I wanted, were preferably from human tissue, and had the best possible resolution. Furthermore, I tried to find ER structures without too many co-factors bound. The following table shows the ER structures I chose along with the ligands and co-factors. Notice, that all the receptors are from human tissue except 1HJ1 which is from rat tissue.

PDB code	Ligand	Agonist/antagonist	Co-factors	Water	Resolution
1GWR	Estradiol	Agonist	Tif2 Nrbox3 Peptide	Yes	2.40 Å
1ERR	Raloxifen	Antagonist	–	Yes	2.60 Å
3ERT	OHT	Antagonist	–	Yes	1.90 Å
1A52	Estradiol	Agonist	–	No	2.80 Å
1QKM	Genistein	Partial agonist	–	Yes	1.80 Å
1HJ1	Fulvestrant	Antagonist	–	Yes	2.30 Å

7.2 Preparations

To prepare the proteins and ligands for docking, I downloaded each PDB file with the crystal structure of the protein, ligand, and co-factors. I examined each of the downloaded files to see if there where superfluous elements or elements that needed modification. I imported the structures into Maestro and looked at the protein/ligand complex. Then I separated the protein, ligand, and water molecules and removed the metal bonds between the ligand and protein in 1ERR.

The protein/ligand complexes were all dimers, so I truncated the complex to a more manageable size by deleting one half of the molecule. When importing complexes into Maestro, the program adds colors to the protein if there are any parts of it, it does not understand or parts that can have more than one conformation. I had a look at all of them and decided what to do in each case. After separating the protein, ligand, and water I looked at the different parts separately. In the ligand, I corrected the atom and bond types and set the formal charge of the ligand. I added hydrogens to the ligand and minimized it with the MMFFs force field and no solvent. After the minimization, I exported the ligand to a new file. Then I looked at the protein to see if anything had to be corrected there. I removed all of the hydrogen atoms and exported the protein. After this, I looked at the water molecules and discarded all but the one water molecule that makes a hydrogen bond with the ligand according to other studies; see [8]. Finally, I exported the water molecule.

After exporting the ligand, protein, and water molecule to separate files, I ran the `pprep` script on the protein file using the default settings. I redirected the output from the script to a log file (`pprep.log`) by using the command:

```
% $SCHRODINGER/utilities/pprep <basename>_prot.mae > pprep.log
```

The `pprep` script produces a modified protein file where all the residues except those that are close to the ligand or that form salt bridges are neutralized. I looked at the `pprep.log` file and imported the modified file `<basename>_protR.mae` into Maestro and here I looked at His524 in the active site and saw that it was in the epsilon protonation state. To decide which protonation state to base my calculations on, I made further calculations on all three types of protonation states. I added hydrogen atoms to the protein and to the water molecule and exported a combined file with the ligand, protein, and water. Then I ran the restrained minimization on the combined file with the command:

```
% $SCHRODINGER/utilities/impref <basename>_lp.mae > impref.log
```

Again, I used the default settings. After the restrained minimization I imported the structure with the name `<basename>_lp_ref.mae` and in Maestro, I separated the ligand, water, and protein before recombining the water and protein. I exported the ligand and the combined water and protein file. Now everything was ready for the grid calculations.

7.3 Calculations

The grid calculations were made using Glide from the co-crystalized ligand in the receptor with the default settings for everything, except the size of the inner grid box, which was different for the different ERs. The following tables summarizes the grid box settings for the ERs:

ER	**1ERR**	**3ERT**	**1GWR**	**1A52**	**1QKM**	**1HJ1**
Inner grid box	16	16	14	14	14	26

The grid calculations alone took between 12 and 45 minutes. Having calculated the grids, it was now possible to start the flexible dockings of all the ligands in each of the ERs. For each of the ERs, I used the co-crystalized ligand as the reference ligand. I had to change the number of rotatable bonds to 32, because some of my ligands have more rotatable bonds then the default value of 15. Furthermore, I changed the conjugated gradient minimization step setting to 5000 instead of the default setting. For each of the ERs, the docking took about 30 minutes; one minute per ligand.

After the docking was finished, I looked at the results with the Glide Pose Viewer, which shows the hydrogen bonds, the G-score and E-mode, and van der Waals interactions between the ligand and protein.

Bibliography

[1] Estriol. Avaliable from *www.aeron.com/estriol.htm*.

[2] FirstDiscovery 2.0 technical notes. Avaliable from *www.schrodinger.com/docs/impact2.0/-pdf/tech_notes/tech_notes.pdf*.

[3] ANDREASEN, P. Biocencors screening for endocrine disruptors by ligand-dependent conformational changes of nuclear receptors. *iNANO Centre, Aarhus University* (2003).

[4] ELDRIDGE, M., MURRAY, C. W., AUTON, T., PAOLINI, G. V., AND MEE, R. P. Empirical scoring functions: I. the development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des. 11* (1997), 425–445.

[5] FRIESNER, R. A., L.BANKS, J., MURPHY, R. B., HALGREN, T. A., KLICIC, J. J., MAINZ, D. T., REPASKY, M. P., KNOLL, E. H., SHELLY, M., PERRY, J. K., SHAW, D. E., FRANCIS, P., AND SHENKIN, P. S. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assesesment of docking accuracy. *Journal af Medicinal Chemistry 47* (2004), 1739–1749.

[6] GOODSELL, D. S. Molecule of the month: Estrogen receptor. Avaliable from *www.rcsb.org/pdb/molecules/pdb45_1.html*, 2000.

[7] HALGREN, T. A., MURPHY, R. B., FRIESNER, R. A., BEARD, H. S., FRYE, L. L., POLLARD, W. T., AND BANKS, J. L. Glide: A new approch for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal af Medicinal Chemistry 47* (2004), 1750–1759.

[8] JORDAN, V. C. Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 1. Receptor interactions. *Journal af Medicinal Chemistry 46*, 6 (2003), 883–908.

[9] JORDAN, V. C. Antiestrogens and selective estrogen receptor modulators as multifunctional medicines. 2. Clinical considerations and new agents. *Journal af Medicinal Chemistry 46*, 7 (2003), 1081–1111.

[10] KAMPMANN, J. P., BRØSEN, K., NIELSEN-KUDSK, F., AND NYBORG, N. C. B. *Basal og klinisk farmakologi*, 2nd ed. 2002.

[11] PIKE, A. C. W., BROZOZOWSKI, A. M., HUBBARD, R. E., BONN, T., THORSELL, A.-G., ENGSTRÖM, O., LJUNGGREN, J., GUSTRAFSSON, J.-Å., AND CARLQUIST, M. Structure

of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *The EMBO Journal 18*, 17 (1999), 4608–4618.

[12] PIKE, A. C. W., BRZOZOWSKI, A. M., AND HUBBARD, R. E. A structural biologist's view of the oestrogen receptor. *Journal of Steroid Biochemistry and Molecular Biology 74* (2000), 261–268.

[13] SONNENSCHEIN, C., AND SOTO, A. M. An updated review af environmental estrogen and androgen mimics and antagonists. *Journal of Steroid Biochemistry Molec. Biol. 65*, 1-6 (1998), 143–150.

[14] SUMBAYEV, V. V., BONEFELD-JØRGENSEN, E. C., WIND, T., AND ANDREASEN, P. A. A novel pesticide-induced conformational state of the oestrogen receptor ligand-binding domain, detected by conformational-specific peptide binding. *FEBS Letters*, 579 (2005), 541–548.

[15] WIDMAIER, E. P., RAFF, H., AND STRANG, K. T. *Human physiology: The mechanisms of body functions*, 9th ed. 2004.

Appendix A

Pesticides

Figure A.1 The structures of the examined pesticides

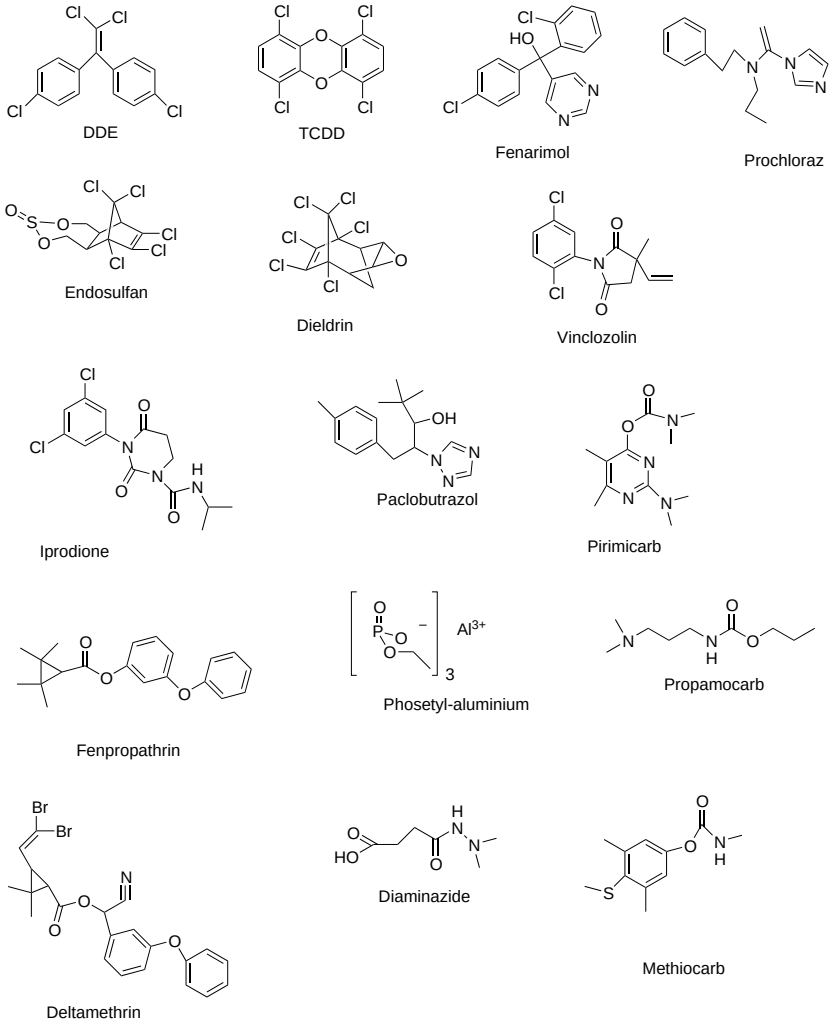
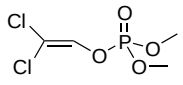
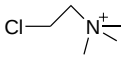


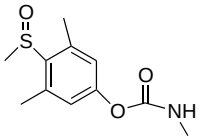
Figure A.2 The structures of the examined pesticides (cont.)



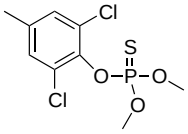
Dichlorvos



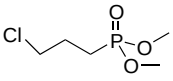
Chlormequat



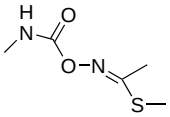
Metiocarb sulfide



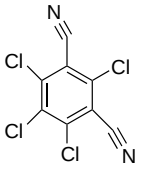
Tolclofos-methyl



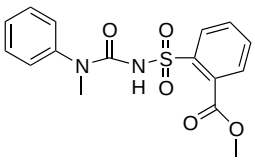
Ethephon



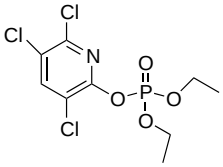
Methomyl



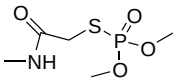
Chlorothalonil



Tribenuron methyl



Chlorpyrifos



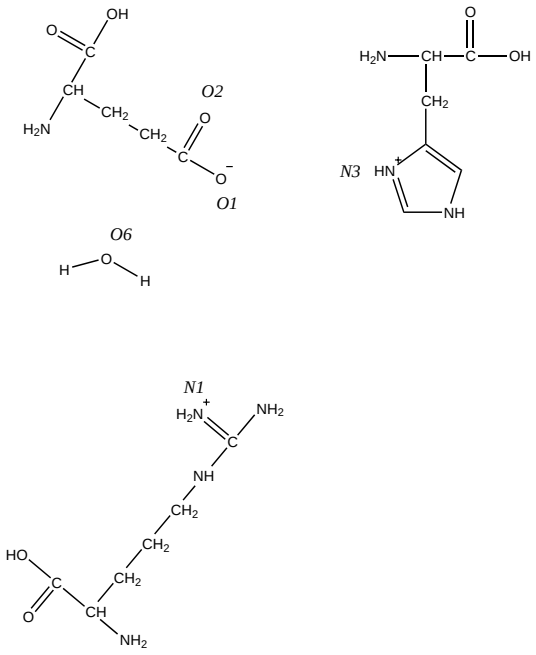
Dimethoate

Appendix B

Hydrogen Bonds

The tables show the lengths and angles of the hydrogen bonds for the different ligand-receptor complexes. I have chosen to show only the hydrogen bonds between the three important residues His524, Arg394, and Glu353 and the water molecule in the active site – see figure B.1. For some of the antagonists, there also exists hydrogen bonds between the arm and residues, but I do not consider those important in this study.

Figure B.1 The LBD with the three important residues and labels for the hydrogen bonds



The hydrogen bond distances are measured between hetero-atoms. I have only displayed the hydrogen bonds with distances smaller than 3.5 Å, because there is no point in measuring angles for non-existing bonds. For a hydrogen bond to be formed, the angle should be $180° \pm 30°$. However, not all the angles in the tables meet this requirement.

B.1 Estradiol

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3-H···O1	2.559 Å/ 169.3°	–	2.641 Å/ 118.0°	2.604 Å/ 120.8°	2.691 Å/ 175.6°	2.478 Å/ 162.3°	–
O3-H···O2	–	2.659 Å/ 179.4°	2.461 Å/ 174.3°	2.697 Å/ 171.1°	–	–	2.986 Å/ 165.2°
O3···H-N1	2.479 Å/ 129.1°	2.824 Å/ 127.1°	2.992 Å/ 145.0°	3.028 Å/ 147.9°	3.327 Å/ 142.9°	–	2.952 Å/ 118.2°
O3···H-O6	3.480 Å/ 109.7°	–	3.132 Å/ 104.3°	2.774 Å/ 99.7°	–	–	
O17···H-N3	2.956 Å/ 154.3°	3.425 Å/ 158.4°	2.809 Å/ 173.7°	–	2.627 Å/ 174.2°	3.048 Å/ 151.0°	2.795 Å/ 158.8°

B.2 Estriol

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3···H-N3	3.156 Å/ 162.2°	2.553 Å/ 111.6°	–	–	–	–	–
O17-H···O1	2.9727Å/ 155.2°	–	–	–	–	–	–
O17-H···O2	–	3.433 Å/ 99.8°	–	–	–	–	–
O16···H-N1	3.395 Å/ 154.1°	2.668 Å/ 169.6°	–	–	–	–	–
O16···H-O6	3.137 Å/ 122.9°	–	–	–	–	–	–
O16-H···O1	2.529 Å/ 168.9°	–	–	–	–	–	–
O3···H-N1	–	–	3.331 150.9°	2.998 Å/ 145.0°	3.006 Å/ 144.4°	2.885 Å/ 118.0°	3.205 Å/ 128.5°
O3···H-O6	–	–	3.078 Å/ 109.4°	2.847 Å/ 101.3°	–	–	3.490 Å/ 75.8°
O17···H-N3	–	–	2.820 Å/ 174.0°	–	3.122 Å/ 164.1°	3.365 Å/ 151.9°	2.845 Å/ 146.7°
O3-H···O1	–	–	2.758 Å/ 116.5°	3.285 Å/ 123.8°	2.792 Å/ 174.2°	2.450 Å/ 164.6°	3.481 Å/ 119.4°
O3-H···O2	–	–	2.758 Å/ 167.1°	2.722 Å/ 171.0°	–	3.320 Å/ 113.5°	2.784 Å/ 166.7°

B.3 Raloxifen

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3···H-N3	–	–	–	–	–	–	2.751 Å/ 146.7°
O4-H···O1	–	3.209 Å/ 121.7°	–	–	–	–	3.235 Å/ 137.9°
O4-H···O2	–	2.425 Å/ 172.1°	–	–	–	–	2.925 Å/ 159.2°
O4···H-N1	–	2.622 Å/ 108.8°	–	–	–	–	2.748 Å/ 133.2°
O4···H-O6	–	–	–	–	–	–	3.023 Å/ 112.6°
O4···H-N3	2.740 Å/ 147.5°	–	–	–	2.671 Å/ 81.5°	–	–
O3-H···O1	2.679Å/ 172.4°	–	–	–	2.738 Å/ 169.7°	–	–
O3···H-N1	2.841 Å/ 140.0°	–	–	–	3.162 Å/ 134.9°	–	–
O3···H-O6	3.342 Å/ 118.5°	–	–	–	–	–	–

B.4 4-hydroxytamoxifen E

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3-H···O1	2.618Å/ 171.4°	–	–	–	3.192 Å/ 159.2°	–	–
O3-H···O2	–	2.562 Å/ 168.7°	–	–	–	–	2.782 Å/ 163.3°
O3···H-N1	–	6.137 Å/ 124.8°	–	–	3.489 Å/ 121.4°	–	–

B.5 4-hydroxytamoxifen Z

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3-H···O1	2.575Å/ 166.3°	3.324 Å/ 131.4°	–	–	2.621 Å/ 171.5°	–	3.467 Å/ 133.6°
O3-H···O2	3.002 Å/ 126.5°	2.712 Å/ 171.5°	–	–	–	–	2.722 Å/ 165.3°
O3···H-N1	2.640 Å/ 149.1°	2.952 Å/ 145.6°	–	–	2.748 Å/ 130.8°	–	3.252 Å/ 127.9°
O3···H-O6	2.965 Å/ 112.7°	3.008 Å/ 117.3°	–	–	–	–	–

B.6 Genistein

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3···H-N3	2.520 Å/ 147.8°	3.457 Å/ 120.3°	–	–	–	–	–
O5-H···O1	2.495Å/ 174.2°	–	–	–	–	–	–
O5-H···O2	2.981Å/ 113.5°	2.614 Å/ 167.6°	–	–	–	–	–
O5···H-N1	2.713 Å/ 149.3°	2.672 Å/ 113.5°	–	–	–	–	–
O5···H-O6	2.993 Å/ 112.6°	–	–	–	–	–	–
O5···H-N3	–	–	2.525 Å/ 163.0°	–	2.483 Å/ 164.7°	2.829 Å/ 173.7°	2.733 Å/ 138.7°
O3-H···O1	–	–	3.129 Å/ 128.7°	3.125 Å/ 118.5°	2.637 Å/ 159.0°	2.620 Å/ 171.7°	3.163 Å/ 126.1°
O3-H···O2	–	–	2.723 Å/ 168.3°	2.544 Å/ 178.2°	3.052 Å/ 131.9°	3.452 Å/ 133.8°	2.659 Å/ 159.4°
O3···H-N1	–	–	2.711 Å/ 146.4°	2.615 Å/ 142.3°	2.544 Å/ 136.4°	3.137 Å/ 134.0°	2.832 Å/ 124.3°
O3···H-O6	–	–	2.920 Å/ 101.0°	2.817 Å/ 110.5°	–	3.277 Å/ 116.1°	3.332 Å/ 76.6°

B.7 Fulvestrant

Bond / ER	1ERR	3ERT	1GWR	1GWRa	1A52	1QKM	1HJ1
O3···H-N3	2.353 Å/ 157.0°	2.853 Å/ 127.3°	–	–	–	–	–
O17-H···O1	3.442Å/ 160.6°	–	–	–	–	–	–
O17···H-O6	3.486 Å/ 142.9°	–	–	–	–	–	–
O17···H-N3	–	–	–	–	–	–	2.947 Å/ 134.7°
O3-H···O1	–	–	–	–	–	–	3.481 Å/ 115.5°
O3-H···O2	–	–	–	–	–	–	2.697 Å/ 162.1°
O3···H-N1	–	–	–	–	–	–	3.306 Å/ 127.8°