Molecular Mechanism of interactions between Estrogen Receptor and Estrogen Receptor Selective Genotoxins

by

Annie S. Lee

S. B. Chemistry
Massachusetts Institute of Technology, 1997

SUBMITTED TO THE DIVISION OF BIOENGINEERING AND ENVIRONMENTAL HEALTH
IN PARTIAL FULFILLMENT OF THE REQUREMENTS
FOR THE DEGREE OF

MASTER OF TOXICOLOGY
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2000

Signature of Author:_____
Division of Bioengineering and Environmental Health
May 24, 2000

Certified by:_____
John M. Essigmann
Professor of Chemistry and Toxicology
Thesis Supervisor

Accepted by:_____
John M. Essigmann
Chairman, Department Committee on Graduate Students

Molecular Mechanism of Interactions between Estrogen Receptor and Estrogen Receptor Selective Genotoxins

by
Annie S. Lee

Submitted to the Division of Bioengineering and Environmental Health
on May 24, 2000 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in Toxicology

ABSTRACT

Although one million new breast cancer cases arise each year worldwide, therapies to treat the disease are limited. Conventional treatments including the chemotherapeutic agent, Tamoxifen, have had only limited success, often showing uncomfortable side effects. Our group has proposed a new scheme for a rational drug design. This scheme utilizes recent findings on the mechanism of cisplatin, the drug found to cure in excess of 93% of all testicular cancer cases. Cisplatin forms DNA adducts that are toxic. The toxicity of these adducts is enhanced by the recruitment of proteins that bind to the adducts and impede adduct repair. This thesis was an attempt to duplicate this "repair shielding" mechanism with another cytotoxin. Specifically, this toxin will be programmed to kill breast cancer cells. Breast cancer cells often overexpress the estrogen receptor protein. By synthesizing a drug that not only binds and damages the DNA but also binds the abundant proteins in the cells, thereby blocking the damaged site from DNA repair proteins, a selective treatment of cancer cells can be achieved.

In this study, the human estrogen receptor (hER) and the ligand binding domain of the hER genes were cloned into baculovirus expression vectors, establishing a system where a large quantity of the proteins can be expressed. The proteins expressed in insect cells were purified in one step, using the FLAG-epitope, yielding homogeneous proteins. The proteins were tested for binding to β-estradiol and were confirmed to be functional in ligand binding. They were also tested for their ability to bind the novel drugs synthesized to bind both the protein and the DNA. It was found that the ligand binding domain of the hER was capable of binding the drugs adducted to the DNA. In an effort to elucidate the mechanism of the protein-drug-DNA complex formation, an association experiment was carried out, which showed that the drug more readily bound to the protein than to DNA. However, a significant amount of the drug-protein complex still bound the DNA, if the ratio of the protein to the drug did not exceed 1.5.

Thesis Supervisor: John M. Essigmann
Title: Professor of Chemistry and Toxicology

Acknowledgements

First and foremost, I would like to thank my advisor, Professor John Essigmann. Working for John has been an integral part of my education at MIT, not only for my graduate studies, but also throughout my undergraduate years. His scientific and professional guidance has been instrumental in what I have achieved so far and what I am about to embark upon. Also, his support and understanding made his lab a pleasant, self-motivating and professional, place to work.

I came to believe that research is only a part of the learning experience in graduate school. A big part of the experience is interacting with other scientists in the field and learning to work with them. The members of the Essigmann lab have been excellent colleagues to learn from. I also thank them for making me feel like I was coming back home when I started in the BEH department. In particular, I would like to thank Hyun Ju Park for helping me design and perform many of the experiments in this thesis, for providing me with the protein model figures included in this dissertation, and for being an excellent mentor on this project. I am also grateful to Bob Croy for answering many questions I had all throughout the course of the project. I want to thank Jessica Kosa and John Essigmann for proof reading this thesis.

I would also like to thank my previous mentors, without whose teachings I would not have been able to accomplish this work. Kevin Yarema and Debbie Kreutzer Wintner have trained me, an undergraduate who knew nothing about laboratory work. They have left such a lasting impression on me that my lab habits often reflect theirs. I owe much thanks to Sang Seok Koh who taught me everything I know about baculovirus system and answered numerous questions via international calls and emails.

This thesis would not have been possible without the support of my friends and family. I would like to thank my friends for yelling at me to work when work needed to be done, for keeping me distracted when I needed to stay sane and for being the endless sources of procrastination they are during the course of this thesis work. Also, this dissertation could not have been written if it weren't for Marin's generous donation of her computer. Last but not least, I want to thank my family, who has been an unfailing source of support and comfort. Especially, I would not have been able to weather the last three years without the unconditional love and understanding of my parents.

Table of Contents

List of Figures

List of Tables

Chapter I. Introduction and Background

Research objectives

Each year, one million new cases of breast cancer are reported in the United States and approximately 400,000 people die from the disease. Most anti-cancer treatments available today have limited success. Many of the chemotherapeutic agents are not tumor-selective, causing many side effects during the treatment. Chemotherapeutic agents often kill dividing cells by damaging DNA. These drugs are not only lethal to cancer cells, but also to non-cancerous cells, causing many side effects. Moreover, in most cases, these treatments extend life, but do not completely eradicate cancer. One exception is cisplatin, which cures >93% of testicular cancers. Cisplatin displays its toxicity by binding to and damaging DNA (1). Proteins may bind to the drug adducts (2,3), blocking the lesion from being repaired by DNA repair enzymes. This information has become the foundation of an effort to design drugs that selectively kill tumor cells.

If a drug can form adducts on DNA and also can bind to a tumor tissue specific protein, it would result in an unrepairable lesion. Such drug would give tumor specificity during chemotherapy. The target tumor chosen for this was breast cancer cells, which often overexpress estrogen receptors.

Compounds containing estradiol-like moiety were synthesized and were tested in cell lines overexpressing estrogen receptor. However, as the research expanded, the need for homogeneous human estrogen receptor in the laboratory increased. The calf uterus cell extract as the source of the protein was inadequate, because of the low concentration

of the estrogen receptor among countless other cellular proteins in the extract. The main objective of this study was to develop and optimize the protein expression system that would produce a large quantity of the protein competent to bind ligands. The resulting purified protein was used to test the ability of the ER to bind the mechanism-based drugs synthesized in the laboratory, an important step in identifying potential new drugs.

This dissertation will be organized in the following manner:

Chapter I. Introduction and Background will focus on the background of the project including the features of the baculovirus expression.

Chapter II. Material and Methods will elaborate the details of the experiments.

Chapter III. Results will present the outcomes of the experiments.

Chapter IV. Discussion will summarize the conclusions that could be made from the experimental data.

Breast cancer and conventional therapy

Breast cancer is the number one cause of cancer death in women in the United States, with 45,000 women dying of the disease in 1998 and one million new cases of the disease expected to be reported around the world each year. Despite recent increases in awareness of the importance of early detection by regular mammography and self-examination has decreased the number of deaths, the disease continues to be a formidable public health problem.

The conventional course of action to treat breast cancer consists of surgical removal of the tumor, followed by radiation and chemotherapy. A long-term regimen of Tamoxifen, which acts to keep the tumor in remission, follows the combined treatment.

Although this multi-staged treatment has shown to decrease the breast cancer mortality, it is not without an expense. The initial chemotherapy gives rise to many widely known adverse side effects, such as bone marrow decimation, nausea and hair loss. It has been reported that Tamoxifen, although it is shown to decrease the tumor recurrence by 45%, increases the risk of endometrial cancer in women who take the drug. Thus, it is of essence that a new approach to the treatment of the cancer is explored. The laboratory has been working on designing drugs that would selectively kill tumor, therefore eliminating such side effects during chemotherapy.

Tumor selective genotoxins

The intellectual basis for the project of designing tumor selective anti-cancer chemotherapeutic agent stemmed from the study of the biochemical mechanisms of cisplatin (*cis*-diamminedichloroplatinum(II)). Regimens including cisplatin are widely used for the treatment of testicular cancer and have been shown to cure almost all of the patients treated with the drug (4), an impressive effect unseen in most other anti-cancer agents. Despite continuing efforts to elucidate the mechanisms of action of cisplatin, its biochemical workings at molecular level are still not fully understood. However, it has been shown that cisplatin kills cells, in part, by damaging the DNA of the cells (1). Subsequently, a group of proteins was identified that binds the DNA adduct formed by cisplatin (2,3). These proteins show a selective affinity for the DNA adducts (3), suggesting that the proteins play a role in cisplatin toxicity. One of the models suggested to explain this phenomenon was the "repair shielding" theory, which proposed that the binding of these proteins to the drug adduct physically blocks the damaged site from

DNA repair proteins thereby heightening the potency of the drug. It is this model that the design scheme of novel drugs selective toward cancer cells is based.

Many cancer cells are known to express proteins specific to the tissue type or cancer. By synthesizing compounds that recognize the tumor specific protein and also damage DNA, tumor selectivity can be achieved (Figure 1). These drugs, if administered into non-cancerous cells, will damage DNA, but the lesions will be readily repaired by the DNA repair machinery due to the absence of the proteins. The system chosen to test this approach was breast cancer cells. Many breast cancers overexpress estrogen receptor, providing a good target for the project. The compounds are composed of two domains, the DNA-reactive domain, or "warhead," and the protein recognition domain. The two domains are connected by a linker. The functional group selected for the warhead was a nitrogen mustard, which was known historically to form adducts and damage DNA. The protein recognition group was derived from, or functionally mimics, estradiol, the natural ligand of the tumor specific protein. Figure 2 shows the structures and the names of the compounds synthesized according to this scheme and used in experiments described in this dissertation.

The theory behind the scheme was confirmed in the study where 2PI-C6NC2 (Figure 2, B) was tested for its toxicity in ER(+) and ER(-) cells. The drug showed increased toxicity in ER-expressing cancer cells (5). It was calculated that approximately 20-25% of the DNA adducts would associate with the ER *in vivo*, thereby blocking the repair. It has since been demonstrated, that the DNA lesions associated with ER were repaired slower than the lesions formed in ER(-) cell line (Croy and Essigmann, unpublished data).

Estrogen receptor

The target protein for the tumor specific genotoxins developed in the laboratory is the estrogen receptor, a protein belonging to the nuclear hormone receptor superfamily. These proteins act as transcriptional regulators, which bind to specific DNA elements and activate transcription in response to the binding of a steroid hormone (6). The ER is a 67kDa protein found in a wide variety of species and is involved in the regulation of gene expression in tissues of female sex organs (7). It has been reported that more than 50% of human breast cancers overexpress the ER.

The human ER (hER) cDNA was cloned and completely sequenced by Green and the co-workers in 1986 (7). The hER cDNA is 1.8kb long and codes for a protein with 595 amino acid residues. Analogous to the typical structural scheme of the nuclear receptors, the ER is divided into 6 domains, designated A through F, classified according to the functions of each region (8). The N-terminal domains A and B contain the constitutive transcriptional activation function, AF-1. The C domain is the DNA binding domain (DBD), which shows a high degree of homology to the DNA binding domains of other nuclear hormone receptors (8). The D domain is designated as the "hinge" domain. The E domain is the ligand binding domain (LBD), which recognizes the hormone; the LBD, upon binding ligand, transforms the receptor to a transcriptionally active state (9). Finally, the F domain is the carboxy-terminal domain, which has variable homology among the receptor proteins.

The hormone ligand, as mentioned earlier, is estradiol (Figure 2A). Estradiol binds to ER tightly with the K_d of 0.35 nM (10). When estradiol is bound to the ligand binding domain of the ER, two molecules of the protein form a homodimer. This

complex, when bound to a specific region of the promoter, activates the transcription of gene connected to the promoter (11).

Figure 3A is a representation of the ER-LBD homodimer containing two molecules of estradiol in its binding domain. It was rendered from the crystal structure coordinates reported by Tanenbaum and co-workers (12). The structure of the protein was solved in complex with estradiol, giving insights into the interactions between the receptor and the ligand at the molecular level. Figure 3B illustrates the interactions in detail. The ligand is held in the binding cavity by a number of hydrogen bonds with the amino acid residues and water molecules. The D-ring hydroxyl group of the estradiol forms a hydrogen bond with His511. The phenyl hydroxyl group of the hormone makes a hydrogen bond with the carboxylate of Glu353 and a water-mediated hydrogen bond supported by Arg394. The same three residues were implicated as being important for ligand binding in an earlier study (13), confirming the structure of the binding cavity of the protein.

Figure 4 shows the ER-LBD binding E2-C6NC2 (Figure 2D), one of the drugs designed according to the strategy described earlier in this chapter. The drug molecule was docked into the protein rendered from the crystal coordinates reported by Tanenbaum et al. (12), and the complex was energy-minimized in water using the molecular modeling software, Insight II (Biosym). Since the protein recognition domain of the drug is estradiol, it is hypothesized that the interactions in the binding cavity of the protein would closely resemble the interactions between the protein and estradiol.

Baculovirus expression system

In the past, the estrogen receptor was obtained for work in this laboratory from a calf uterus cell extract. In addition to the unpleasantness of the preparation, the cell extract was not an adequate source of the protein, because it yielded only limited quantity of the protein. Also, because the extract contains other cellular proteins, the possibility of proteins other than estrogen receptor acting on the system could not be eliminated. Lastly, only bovine estrogen receptor, not human estrogen receptor, could be obtained from the preparation. Thus, there was a great need for a recombinant human protein expression system.

The baculovirus expression system has become widely used for recombinant eukaryotic protein production because of its ability to produce large quantities of biologically active eukaryotic proteins (14). By placing the gene coding for the protein of interest under the control of a strong baculoviral promoter, a high expression level of the protein can be achieved in insect cells. The proteins expressed in the cells can be purified using many protein purification tools commercially available, including the products developed for the purification of FLAG® epitope (Sigma) tagged proteins. In this section, a general overview and the features of the baculovirus expression system will be discussed. Also, a brief description of the isolation of the recombinant virus will be introduced.

The baculovirus vector uses the highly expressed and regulated *Autographa californica* nuclear polyhedrosis virus (AcMNPV) polyhedrin promoter (15) modified for the insertion of foreign genes. Insect cells provide an eukaryotic environment for posttranslational modification and folding which may be essential for producing

biologically active proteins (16). Posttranslational modifications of the gene products achieved by baculovirus expression systems closely resemble glycosylation, fatty acid acylation and phosphorylation normally present in their native forms (17). Baculoviruses infect insects and insect cells and they are not pathogenic to vertebrates or plants (18), making them safe and easy to handle. The stability of the recombinant viruses allows a long-term storage of the viruses.

The structure and the life cycle of the baculovirus makes it well suited for the use in protein expression. AcMNPV contains a double-stranded, circular, supercoiled DNA genome of approximately 134 kb (19). During AcMNPV infection, the viral genome is expressed in four different stages – immediate early, early, late and very late -- at which different genes are transcribed. The very late stage spans from approximately 18 hour post-infection until cell lysis. Two viral proteins are produced in large quantities during the very late phase, eventually amounting to ~50% of all cellular proteins. These are the polyhedrin protein, a viral structural protein of 29 kDa, and the p10 gene product, a non-structural protein of 10 kDa. Both proteins are non-essential for replication and formation of virus particles (20,21) and therefore have been targeted for use in expression vectors. In many of the widely used baculovirus expression vectors, the polyhedrin gene is replaced with the recombinant gene and the strong polyhedrin promoter will drive the transcription of the foreign gene, expressing the recombinant protein during the very late phase of the infection (22).

The host cells for the baculovirus, *Spodoptera frugiperda* (Sf) cells, were originally derived from pupal ovarian tissue of the fall armyworm (23). *Spodoptera frugiperda* IPLB-Sf-21AE (Sf21) cells and a cloned version of this cell line (Sf9) are

commonly used for the expression system. These cells can be grown either as a monolayer or in suspension culture at 25 to 28°C in a low protein media containing fetal bovine serum and antibiotics.

The first step to recombinant protein expression is the construction of the baculovirus transfer vector containing the gene of interest. The size of the AcMNPV genome is too large for routine molecular biological manipulations such as restriction digests and ligation. Instead, the target gene is inserted into a transfer vector (size <10 kb) and the completed plasmid is co-transfected into insect cells along with the AcMNPV viral DNA. Inside the cell, a complete expression vector is formed via homologous recombination between the transfer vector and the viral DNA.

Clontech's BacPAK6 system applied a number of modifications to the virus genome in order to improve the frequency of recombination. As illustrated in Figure 5, three *Bsu*36I sites were introduced into the viral DNA. It linearized the genome, which increased the frequency of recombination in insect cells (24). The *Bsu*36I digests also disrupt ORF 1629 gene, a gene essential for viral replication in host cells (25). Infectious viruses are formed only when a successful recombination takes place between the viral DNA and the transfer vectors such as pBacPAK8 and pBacPAK9 containing the ORF 1629 gene, giving rise to an intact ORF1629. On the other hand, recircularized viral DNA is not viable because it lacks ORF 1629. The *Bsu*36I digest also eliminates the polyhedrin promoter from the viral DNA, which is also restored, along with the target gene, by the recombination with the transfer vector. These features of Clontech's BacPAK6 system drastically increases the recombination frequency, reducing the need to screen many viral clones to obtain the desired recombinant virus.

Once the recombinant virus is produced by co-transfection, the virus is screened via plaque assay to acquire a pure stock of the virus. The production of clear, well-defined plaques during the assay demonstrates the viability and high titers of the recombinant virus. The production of the correct protein by the isolated virus is confirmed by Western blot analysis of the cell extracts from a small-scale infection. Then the virus expressing the target protein is amplified for further use. The details of these steps will be described in the next chapter, Materials and Methods.

The purification of the proteins expressed in the insect cells can be conducted in one step, using the FLAG epitope. The purification of a FLAG-tagged protein can be greatly simplified by using an anti- FLAG immunoaffinity column (16). FLAG sequence is an 8-amino acid peptide (DYKDDDDK), which can be easily detected by commercially available tools such as anti-FLAG monoclonal antibody and anti-FLAG immunoaffinity column resin. Production of FLAG-tagged protein is achieved by using transfer vectors, pSK277 or pSK278, which contain internal sequences coding for FLAG epitope (16). Because they are derivatives of pBacPAK8 and pBacPAK9, respectively, they contain all the features of the transfer vectors and can be co-transfected into insect cells with viral DNA to generate viruses expressing N-terminal FLAG-tagged proteins.

The baculovirus expression system, therefore, is an excellent system for the purpose of this project. Its capacity to yield a large quantity of the protein via quick and easy purification procedure will be able to fulfill the demands for the estrogen receptor.

Chapter II. Materials and Methods

Materials used

The pSV2NeoCMV vector containing hER cDNA was a gift from Dr. Ruth Sager. The baculovirus expression vectors pBacPAK9 (Clontech) and pSK278 were gifts from Dr. Sang Seok Koh of the Whitehead Institute. The *E.coli* strain DH5α [F-, ϕ80d*lac*ZΔM15, *rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17 (r_{K-}, m_{K+}), *sup*E44, *rel*A1, *deo*R, Δ(*lac*ZYA- *arg*F) U169] was used for the production of plasmids. All enzymes were purchased from New England Biolabs. pGEM plasmid preparation and [^{14}C]Hexene synthesis were carried out by Dr. Robert Croy.

Subcloning of hER and hER-LBD into baculovirus vector

The hER cDNA was amplified from pSV2NeoCMV-hER (7). Oligomers ALO1 (5'-CGCGGATCCATGACCATGACCCTCCACACCAAA-3') and ALO2 (5'-CCGCTCGAGTCAGACTGTGGCAGGGAAACCCTC-3') were used as primers for PCR reaction to obtain full-length hER cDNA. The fragment was double digested with BamHI and XhoI and inserted into baculovirus expression vectors, pBacPAK9 and pSK278, between the same sites to give ER and FLAG-ER, respectively. Oligomers ALO3 (5'-CGCGGATCCATGTCTGCTGGAGACATGAGAGCTGCC-3') and ALO2 were used as primers for PCR reactions to obtain the ligand binding domain (LBD) of hER. The PCR product was inserted into pBacPAK9 between BamHI and XhoI sites to give LBD. Similarly, PCR product obtained with oligomers ALO4 (5'-CGCGGATCCTCTGCTGGAGACATGAGAGCTGCC-3') and ALO2 was inserted into

BamHI and XhoI sites of pSK278 to give FLAG-LBD. The positive clones were confirmed by BamHI/XhoI digestion and the plasmids were isolated using Qiagen Plasmid Maxi Kit.

Culturing of Sf21 cells

Sf21 cells were cultured in ExCell420 (JRH Biosciences) insect cell media containing 10% Fetal Bovine Serum, 50 μg/mL Penicillin-Streptomycin (Life Technologies) and 2.5 μg/mL Fungizone (Life Technologies). The cells were grown at 27°C on 150x25 mm tissue culture plates or in suspension. For suspension cultures, 0.1% of Pluronic-F68 (Sigma) was added to the media.

Isolation of recombinant baculovirus

In order to produce a large quantity of the full-length and the ligand binding domain of hER protein in insect cells, recombinant baculoviruses expressing the proteins were constructed. The plasmids coding for ER (pAL3), LBD (pAL4), FLAG-ER (pAL5) and FLAG-LBD (pAL6) were co-transfected into Sf21 cells along with BacPAK6 viral DNA BsuI36 digest (Clontech) to produce recombinant viruses. The recombinant viruses were isolated by plaque assay and the expression of the appropriate proteins was confirmed by Western blot analysis of the whole cell extract from the Passage One (P1) virus. The virus was then amplified through subsequent passages in preparation of the large-scale expression and purification.

1. Co-transfection

For each construct containing estrogen receptor gene, 0.5 μg of plasmid DNA and 5 μl of BacPAK6 viral DNA were incubated at room temperature with Bacfectin provided by Clontech to form DNA-lipofectin complexes. The complexes were co-transfected into exponentially growing Sf21 cells on 6-well plates and the cells were incubated at 27°C. After 72 hours, the media containing the recombinant viruses was collected.

2. Plaque assay

Densely plated Sf21 cells were infected with co-transfection supernatant at various dilutions. The cell monolayer was overlaid with low-melting agarose (1% Sea Plaque agarose from FMC) into which cells can grow. Media was added after the agarose layer was set and the plates were incubated at 27°C. After 5 days, the cells were stained with Neutral Red (Life Technologies) and developed overnight. Plaques well isolated from the surrounding plaques were selected. Four plaques per construct were picked and placed in media overnight to release the virus from the agarose plug.

3. Passage one and whole cell extract preparation

Sf21 cells were infected with the virus released from the plaques. After 4 days, the media was collected and the cells were harvested. The media containing Passage one (P1) virus was stored at 4°C. The whole cell extract (WCE) was prepared by lysing the cells with lysis buffer (50 mM Tris-Cl pH7.6, 150 mM NaCl, 2 mM EGTA, 1% Nonidet P-40, 0.25% sodium deoxycholate, 1 mM PMSF, 2 mM benzamidine, 0.6 μM leupeptin,

2 μM pepstatin A, and 2 μg/mL chymostatin), pelleting out the cell debris and collecting the resulting supernatant.

4. Western blot analysis

The expression of the appropriate proteins was confirmed by immunoblotting. Whole cell extracts were separated on SDS-PAGE gels and the proteins were transferred to nitrocellulose membranes. The membranes were incubated with anti-hER raised against DNA binding domain (Affinity Bioreagents, Inc. MA1-310), "hinge" region (StressGen, SRA-1000) or ligand binding domain (StressGen, SRA-1010) in order to detect virus expressing proteins with the respective functional regions. WCE from baculoviruses derived from pAL5 (FLAG-ER) and pAL6 (FLAG-LBD) were blotted with anti-FLAG antibody (Sigma, F3165) as well, in order to confirm that the FLAG tag was intact in these proteins. The blots were incubated with anti-mouse IgG-HRP secondary antibodies. The signals were developed using SuperSignal ECL kit (Pierce) and were visualized by exposure to film.

5. Passage two preparation

The P1 viruses that expressed the correct proteins were pooled together and the viruses were amplified. $1.4x10^7$ cells plated on 150x25 mm plates were infected with 2 mL of P1 virus and incubated at 27°C. After 4 days, the media containing the viruses (Passage two--P2) were collected and were ready to be used for large-scale infection and purification of the proteins.

Expression and purification of FLAG-ER and FLAG-LBD

1. Large scale infection

Sf21 cells grown in suspension to a density of $1.8x10^6$ cells/mL were infected with P2 viruses. 500 mL of cells were harvested and the cell pellet was resuspended in 50 mL of P2 virus. Infection mixture was incubated at room temperature for 1 hour with occasional swirling. The volume was brought up to 500 mL with fresh media and the cell culture was grown at 27 °C for 60-66 hours. Cells were harvested and washed with phosphate-buffered saline. The media, the Passage three (P3) virus, was collected and stored at 4°C.

2. Cell extract preparation

The harvested cell pellets were resuspended in 7 mL of MTB100 buffer (50 mM HEPES-KOH pH 7.3, 100 mM KOAc, 25 mM MgOAc, 5 mM EGTA, 10% glycerol, 1 μM DTT, 0.01% Nonidet P-40, 1 mM PMSF, 2 mM benzamidine, 0.6 μM leupeptin, 2 μM pepstatin A, and 2 μg/mL chymostatin) and the cells were lysed by sonication. The lysate was centrifuged in the ultracentrifuge at 28,000 rpm, 4°C for 3 hours in the SW41 rotor. The resulting supernatant was the cell extract containing the recombinant proteins. An aliquot of the extract was saved for Western blot analysis.

3. Protein purification

The column with 1mL anti-FLAG resin (Sigma, A1205) was equilibrated at 4°C. The column was washed with100 mM glycine-HCl pH 3.5 and then with MTB300 buffer (same as MTB100, except 300 mM KOAc). The salt concentration of the cell extract

was adjusted to 300 mM with 5 M KOAc. The cell extracts were loaded onto the column 4 times to maximize the retention of the proteins by the resin. The column was washed with MTB300 buffer and then with MTB100 to bring the salt concentration back to 100 mM. The proteins were eluted with MTB100 buffer containing FLAG peptide (Sigma, F3290) in eight 1 mL-fractions. The presence of the proteins in each fraction was monitored by Bradford assay (26) and the fractions containing the proteins were pooled together. The excess salts, FLAG peptide and buffer were eliminated from the eluate using Centricons concentrators.

4. Protein quantification

The purified proteins were quantified by performing the Bradford Assay, using bovine serum albumin as the standard. The purity and the quantity of the purified proteins were determined by SDS-PAGE analysis. The protein bands were visualized by Coomassie stain.

Protein expression time course

50 mL of Sf21 cells per time point per protein were seeded on tissue culture plates at $0.7x10^6$ cells/mL. 2 mL of P2 virus were used to infect $3.5x10^7$ cells at room temperature. The cells were grown at 27°C and were harvested every 6 hours between 36 and 72 hours post-infection for FLAG-ER and every 12 hours between 36 and 72 hours post-infection for FLAG-LBD. The extracts were prepared as described above and Western blot analysis was performed with anti-LBD antibodies to detect the protein expression level at each time point.

Ligand binding assay

1. Estradiol competition

The purified FLAG-ER and FLAG-LBD proteins were diluted to 2.34 nM with BSA solution (10 mg/mL) and a stock solution of 100 nM [^{3}H]estradiol (1 mCi/μl, New England Nuclear) was prepared. Serial dilutions of unlabeled estradiol ranging from 20 μM to 0.02 μM were made in 50% DMF. The reactions containing 2.1 nM protein, 5 nM of [^{3}H]estradiol and varying concentration of unlabeled estradiol in a total reaction volume of 50 μl were incubated overnight at 4°C. Fifty μl of Dextran/charcoal mixture (0.5% Dextran, 0.5% charcoal in 10 mM Tris pH 7.6) was added to each tube and the tubes were incubated on ice for 30 minutes with occasional vortexing. The tubes were centrifuged at 4°C for 5 minutes and 75 μl of supernatant was counted in 5 mL of scintillation fluid.

2. Drug binding assay

In lieu of unlabeled estradiol, dilutions of Heptene (for FLAG-LBD only), Hexene and Amine Mustard (AM) were added to the competition reaction described above. From 1 mM stock solution of each drug, dilutions ranging from 20 μM to 0.1 μM were prepared for the reaction. The concentration of the test compound necessary to reduce receptor-bound radioactivity by 50%, the concentration at the inflection point, was determined from the plot of drug concentration vs. [^{3}H]estradiol CPM. The relative binding affinity (RBA) was determined by multiplying the molar ratio of the inflection point of estradiol and the test compound by 100 (5).

Gel mobility shift assay

1. Target preparation

Oligonucleotides C1 (5'-TTCTGGCCTCCTCT-3') and C2 (5'-CGCTCGAGAGGAGGCCAGAATTCGGA-3') were 5' phosphorylated with [γ-^{32}P] ATP (New England Nuclear) to give radiolabeled target DNA. Twenty pmol of each oligonucleotide were incubated at 37°C for 1 hour with 40 μCi [γ-^{32}P] ATP, T4 polynucleotide kinase and 10x kinase buffer. The C1 oligomer previously modified with one molecule of Heptene, called Hep1, was phosphorylated similarly. Free [^{32}P] ATP was eliminated by G25 columns at the completion of the reaction.

Labeled C1 was annealed with unlabeled C2, and unlabeled C1 was annealed with labeled C2 to give duplexes with one labeled strand and one unlabeled strand. The annealing reactions were carried out by heating to 80°C and cooling to room temperature overnight. The duplexes were called $C1_D$ and $C2_D$ respectively. Also, the duplex of Hep1 with unlabeled C2 was called $C1_D$-Hep1. $C1_D$-Hep1 was prepared by incubating at 4°C in 0.5 M NaCl and 25 mM Tris pH 8.0 overnight.

2. Drug reaction

Both $C1_D$ and $C2_D$ were allowed to react with drugs Heptene or Phenylindole (PI) to give DNA with drug lesions. Ten pmol of each duplex DNA were incubated with Heptene or PI at 37°C for 12-16 hours. The drug-adducted DNA was isolated by phenol-chloroform extraction followed by ethanol precipitation. The presence of the drug-adducted DNA duplex was confirmed by running a 16% denaturing gel and visualizing it on phosphorimager screen.

3. GMSA

Binding reaction of the protein to adducted DNA was performed in DNA binding buffer (15 mM Tris pH 7.9, 80 mM KCl, 4 mM DTT, 0.2 mM EDTA and 10% glycerol) with 10 μg of BSA and 1 μg of p(dIdC) carrier DNA as specified in Nardulli et al. (27). Twenty-five ng of FLAG-LBD and DNA duplex displaying approximately 15,000 CPM counts were added to the reaction. The reaction mixtures were incubated at 37°C for 20 min. The results were visualized by on a native, low salt, 5% polyacrylamide gel.

Protein-drug-DNA association mechanism

Reactions contained 15 μg of BSA, 1μg of pGEM (0.5 pmol), 0.15, 0.3, 0.45 or 0.6 μg of FLAG-LBD protein (4.5, 9, 13.5, 18 pmol, respectively) in 1M TE pH 8.0. [^{14}C]Hexene (specific activity of 53 mCi/mmol) was added to a final concentration of 200 μM. The reactions were incubated at 37°C for 1hour. Proteins in the reactions were eliminated by phenol/chloroform extractions and the DNA was recovered by ethanol precipitation in the presence of 6 μg of glycogen. The DNA pellet was resuspended in 10 μl of TE pH 8.0 and half of the recovered DNA was separated on a 0.8% agarose gel. After visualizing the bands with UV, the gel was dried and exposed on phosphorimager screen for 36 hours for imaging.

Chapter III. Results

Subcloning of hER and hER-LBD into baculovirus vector

The subcloning of full-length and ligand binding domain of the hER into baculovirus expression vectors was carried out as illustrated in Figure 6. The PCR reaction with pSVNeoCMV as the template and ALO1 and ALO2 as the primers gave a 1.8 kb fragment with a flanking BamHI restriction site at the 5' end and a XhoI restriction site at the 3' end. The fragment was double digested with BamHI/XhoI and ligated into BamHI/XhoI-digested pBacPAK9 and pSK278 baculovirus expression vectors. The 7.3 kb plasmids that resulted from the reaction were designated pAL3 and pAL5, respectively, and served as the construct for the ER and FLAG-ER baculoviruses, respectively.

The amino acid residues 282 to 595 code for the "hinge" region and the ligand binding region of the estrogen receptor protein (28). This region was subcloned into the baculovirus expression vectors to produce the ligand binding domain (LDB) of the estrogen receptor. The PCR reaction with pSVNeoCMV, ALO3 and ALO2 gave a 0.94 kb fragment also with BamHI and XhoI sites. Because the ligand binding domain is located near the C-terminus of hER, it does not contain an internal start codon (ATG) from which the protein can be translated. The codon was included in the primer, ALO3, so that the resulting PCR product would have a start codon immediately following the BamHI restriction site. The PCR product was inserted into pBacPAK9 as above and the 6.5 kb plasmid construct for LBD was named pAL4.

The PCR reaction with pSVNeoCMV, ALO4 and ALO2 yielded a 0.94 kb fragment that was used for the subcloning of FLAG-LBD. Because pSK278, a baculovirus expression vector with a built-in FLAG sequence, already contained a start codon (16), the PCR primer did not need to contain a separate start codon. ALO4 and ALO2 amplified the ligand binding domain and created BamHI and XhoI sites at the 5' and 3' ends, respectively. Restriction digest and ligation into pSK287 yielded a 6.5 kb plasmid expressing FLAG-LBD and it was designated pAL6.

Generation of baculovirus expressing ER and LBD

The recombinant viruses expressing ER, LBD, FLAG-ER and FLAG-LBD were produced by co-transfecting BacPAK6 with pAL3, pAL4, pAL5 and pAL6, respectively. Plaque pick assay enabled the visualization of the recombinant viruses' infection of Sf21 cells and isolation of the viruses. All viruses gave numerous well-defined plaques at all dilutions of the co-transfection supernatant. Four plaques for each virus were picked from wells infected with 10^{-3} or 10^{-4} dilutions of the cotransfection supernatant, because those dilutions gave most well-isolated plaques.

The viruses expressing the correct proteins were identified by Western blot analysis. The WCE of the cells expressing ER and FLAG-ER were blotted against anti-ER (raised against the DNA binding domain) antibody. FLAG-ER was blotted against anti-FLAG antibody to check for the presence of the FLAG tag. None of the 4 clones of ER displayed signals on the anti-ER blot (data not shown). The clones of FLAG-ER expressed the 67 kDa full-length protein (Figure 7A). Several smaller bands also appeared on the blot, recognized by both anti-ER and anti-LBD antibodies. These were

truncated forms of FLAG-ER containing the intact FLAG-tag, which is at the N-terminal end of the protein, and the DNA binding domain, which is in the front half of the protein. The P1 virus for FLAG-ER was pooled and was named ALBV1. The WCE from ER-expressing cells were blotted against anti-ER (anti-DNA binding domain) antibody. The Western analysis gave no signals for any of the four clones (data not shown). The production of the recombinant virus for ER was attempted again, starting from subcloning of the plasmid construct. The yet-to-be-confirmed recombinant virus was named ALBV3.

The WCE of the cells expressing LBD and FLAG-LBD were blotted against anti-ER (raised against LBD or the "hinge" region) antibody. FLAG-LBD extracts were also blotted against anti-FLAG antibody. The clones of FLAG-LBD expressed the 33 kDa protein with the intact FLAG tag. Similarly, the clones of LBD yielded the 33 kDa protein (Figure 7B). The Western blot analysis showed that the FLAG-LBD protein was slightly larger than LBD due to the 8 amino acid FLAG tag. FLAG-LBD and LBD viruses were named ALBV2 and ALBV4, respectively.

Expression and purification of FLAG-ER and FLAG-LBD

The full-length, FLAG-tagged ER protein was purified using the anti-FLAG affinity resin. One major band and two minor bands co-eluted near 67 kDa (Figure 8). The major band was approximately 80% of the total proteins eluted. The truncated proteins which contained the intact FLAG-tag may have co-purified with the full-length FLAG-ER, but would have been removed in the concentration step where Centricon-50 eliminated everything whose molecular weight was smaller than 50 kDa. Using the one-

step purification procedure, approximately 30 μg of the FLAG-ER was purified from 500 mL of Sf21 cells.

In order to obtain the homogeneous FLAG-ER, the baculovirus (ALBV1) was re-isolated by repeating the plaque assay, but the contaminant bands could not be eliminated. There have been reports suggesting that hER can be posttanslationally phosphorylated in Sf9 and MCF-7 cells (29,30). It is possible that one of the contaminating bands was the phosphorylated form of ER. Also, the extraneous bands may be the broken-down forms of FLAG-ER, if the protein had become unstable during the purification process.

FLAG-LBD was also purified by the one-step procedure. The 33 kDa protein was approximately 70% homogeneous with one contaminant protein roughly 67 kDa in size (Figure 8). A 1 L Sf21 culture yielded 30-45 μg of the FLAG-LBD.

Protein expression time course

The time course protein expression was conducted in order to optimize the protein expression procedure. The Western blot analysis using anti-LBD antibody indicated that the highest FLAG-ER expression was 54-hour post-infection (Figure 9A). The protein expression level had dropped off by 62-hour, the infection time designated in the original protocol.

The expression level of FLAG-LBD peaked at 60-hour post-infection and decreased afterwards (Figure 9B). This coincided with the original protocol, which specified between 60 and 66 hour infection at 27°C. For the subsequent protein

expression attempts, the infection times were adjusted according to the findings from this experiment.

Competitive binding of FLAG-LBD to estradiol and drugs

The ability of the FLAG-LBD to bind to estradiol, the natural ligand of the ER, was examined. The amount of radiolabeled estradiol bound to FLAG-LBD as a function of unlabeled estradiol concentration yielded a sigmoidal curve indicating that the FLAG-LBD is functional in binding its natural ligand (Figure 10A). The recombinant protein displayed the inflection point of approximately 10 nM, which was comparable to the inflection point reported in the previous study (5).

Upon confirming that the recombinant protein actively bound the ligand, its ability to bind to different drugs was examined. The FLAG-LBD showed binding to both Heptene and Hexene (Figure 10B and 10C), although the binding was not as strong as that of estradiol. The plot of radiolabeled estradiol amount in relation to the drug concentration showed a sigmoidal curve for both drugs, but the curve was shifted to the right (inflection point = 34.1 nM) compared to the curve for estradiol indicating that the binding of the drugs was weaker than that of estradiol. The relative binding affinity (RBA) of Hexene to estradiol was approximately 39. The RBA of Heptene could not be easily compared, because different batches of purified protein and different lots of tritiated estradiol were used for the two assays. However, similar assays done using calf uterus whole cell extract containing estrogen receptor proteins showed similar RBAs for the two compounds (Croy and Essigmann, unpublished data).

The assay using AM showed no association between the FLAG-LBD and the compound (Figure 10B). This was expected, since the compound lacks the hormone moiety, which would interact with the ligand binding domain.

Competitive binding of FLAG-ER to estradiol and drugs

The competitive binding of FLAG-ER to estradiol and the drugs showed very similar results to those of FLAG-LBD. The amount of radiolabeled estradiol bound to the protein showed a sigmoidal curve, indicating that the protein actively bound the ligand (Figure 11A). Similarly, in the reaction where Hexene was competing against the labeled estradiol for binding to the protein showed binding, although the affinity was not as high as that of estradiol (Figure 11B). As it was for FLAG-LBD, the reaction with AM showed no competition (Figure 11B).

Gel shift mobility assay

A gel mobility shift assay (GMSA) was performed to assess the ability of the FLAG-LBD to interact with DNA modified with different drugs. In the presence of FLAG-LBD and drug-damaged DNA, the protein-DNA complexes via drug were formed and band shifts were observed (Figure 12 lanes 3, 4, 5, 8 and 9). The lanes with DNA alone or unmodified DNA with the protein, however, did not show the slow-moving band (Figure 12 lanes 1, 2, 6 and 7). In both $C1_D$ and $C2_D$ complexes, duplexes modified with Heptene had stronger affinity towards FLAG-LBD than did those modified with PI (Figure 12 lanes 3 and 4, 8 and 9). Also, a much stronger association between the protein

and the DNA was observed when there were multiple drug adducts on the DNA, compared to the mono-adducted DNA (Figure 12 lanes 4 and 5).

This assay confirmed that the recombinant protein was functional in binding not only free drugs, but also in forming complexes with DNA modified by the drugs.

Protein-drug-DNA association mechanism

This experiment was performed in order to investigate the kinetics of the protein-drug-DNA formation. As shown in Figure 13, two different paths of association are available when the radiolabeled drug ([^{14}C]Hexene) is added to a reaction containing DNA (pGEM) and the protein (FLAG-LBD)—the drug may react with the DNA first (B) or it may bind to the protein first (D). The formation of the protein-drug-DNA complex (C) from B has been demonstrated with GMSA, but it has not been shown if the same complex can be formed from D.

In this reaction, the proteins were removed from the reaction mixture by phenol/chloroform extraction and the remaining DNA was resolved on an agarose gel and was visualized with UV. The phosphorimager analysis of the same gel showed the radiolabeled drugs adducted to the DNA.

The UV visualization of the plasmids reacted with the drug in the presence of different amounts of FLAG-LBD showed similar intensity bands (Figure 14A). However, the phosphorimager analysis of the same gel showed a decrease in the ^{14}C signals as the amount of the protein added to the reactions increased. This indicates that as the protein concentration in the reaction increases, the association between the protein and the drug probably occurs more frequently than that between the drug and the DNA.

This path does not completely take over, shown by the existence of the bands in the phosphorimage up to the point where the ratio between the protein and the drug adducts was 9:1 (Figure 14B lanes 2 to 5). However, when the ratio of the protein to the drug adduct was approximately 1.5:1 (using the estimation of 6 drug adducts per pGEM plasmid (Croy and Essigmann, unpublished data)), a significant amount the drug reacted with the plasmid (Figure 14B, lane 2). The faint bands above the main bands in lanes 1 and 2 of both the UV visualization and the phosphorimage were probably the nicked form of the plasmid, formed by the reaction with the drug.

Chapter IV. Discussion

This study describes the expression and the purification of the human estrogen receptor (hER) and the ligand binding domain of the hER using the baculovirus expression system. In order to investigate the mechanism and the optimization of the estrogen receptor-selective genotoxins, the production of a large quantity of pure proteins was required. The use of the baculovirus system allowed the expression of the recombinant proteins in insect cells, thereby eliminating the need to work with the cell extracts prepared from animal organs containing estrogen receptors.

The plasmid constructs containing hER cDNA in baculovirus expression vectors, the recombinant viruses derived from the constructs and the proteins expressed by the viruses are listed in Table 1. It was demonstrated that a one-step purification yielded proteins competent to bind the natural ligand, estradiol, as well as drugs containing estradiol moiety. 30 to 45μg of FLAG-LBD was purified from 1L of Sf21 cells and approximately 30μg of FLAG-ER was purified from 500mL of Sf21 cells.

Purity of the Proteins

Much of the purity and the yield of the recombinant proteins seemed to depend on the status of the Sf21 cell culture at the time of infections. The same amount of cells at the same point in their growth curves infected with the same ALBV2 virus yielded different amount and purity of FLAG-LBD protein. Therefore, it was crucial that the Sf21 cells were healthy -- growing rapidly and have normal morphology -- at the time of the infection and that the cell culture was replaced with cells from the frozen stock at a

regular interval. It has been suggested that high passage number of cells decreased virus production (18).

A few observations were made regarding the purification of the FLAG-ER and FLAG-LBD. FLAG-ER co-eluted with two contaminating bands similar to FLAG-ER in size. If one of the bands was the phosphorylated form of the protein, the treatment with phosphotases would collapse that band, increasing the purity of the protein. On the other hand, if the contaminating bands were broken-down form of the protein because the protein was not stable enough during the purification, it has been speculated that eluting the protein in the presence of a trace amount of estradiol may enhance the stability. This, however, would not be an ideal method of the protein purification for the purpose of this study, since the examining the interaction between the protein and the drugs is an essential aspect of the study. The protein purified with estradiol, and therefore is bound to the ligand, is less likely to interact with the drugs, given the protein's higher affinity towards estradiol than that towards the drugs being studied. Because of the possibility that an incompetent recombinant virus was the cause of the impure protein, co-transfection of pAL5 and isolation of new virus are being repeated.

FLAG-LBD also co-eluted with a protein 66kDa in size. The Western blot analysis of the cell extract just prior to the purification showed no corresponding signal at 66kDa, indicating that the contaminating band does not contain a portion recognized by anti-ER antibodies and is not related to ER protein. It is possible that the protein is a cellular protein that contains amino acid sequence recognizable by anti-FLAG affinity resin.

Baculovirus expressing ER

The failure to produce baculovirus that expressed untagged ER may be due to the plasmid DNA from which the virus was derived. The fact that the co-transfection supernatant gave many clear plaques and yet all four plaques isolated failed to express the correct protein suggests that the plasmid clone selected for the co-transfection was the source of error. It is possible that the selected clone contained a mutation, which was not detected by the restriction enzyme analysis following the miniprep of the clones, caused the expression of a different protein or no expression at all. The possibility of ER being toxic or unstable to be expressed in the cells were excluded, since there have been a number of purification of ER via baculovirus system reported in the literature (29,31).

Activity of the recombinant proteins

Some observations about the activity and the overall quality of the purified recombinant proteins could be made from the binding curves of the proteins. First, the binding curves for the FLAG-ER and FLAG-LBD with estradiol (Figures 10A and 11A) both show sigmoidal curves, indicating that both proteins bind to their natural ligand. However, there was approximately 4-fold difference in the maximum amount of the tritiated estradiol bound to FLAG-ER and to FLAG-LBD, although the same number of receptors, calculated from the protein concentrations, was used for both sets. This suggests that the purification of FLAG-LBD was more successful than that of FLAG-ER and that the amount of the active FLAG-ER in the final purification product was less than previously thought. It is likely that most of the protein of the major band in Figure 8 was contamination or inactive FLAG-ER. Therefore, it will be beneficial to attempt the

purification of FLAG-ER by remaking the recombinant baculovirus, starting from co-transfection. Further characterization with the more pure FLAG-ER would be very informative.

Secondly, although the purity of the two proteins differed greatly, the quality of the active proteins seems to be very similar. The two binding curves for estradiol in Figure 10A and 11A, when scaled to account for differing y-value range, overlap almost completely (Figure 15A). Despite the differences in the absolute binding amount of estradiol, the two proteins' bindings towards the ligand at different ligand concentration were similar. The curves overlapped almost completely and they have almost the same inflection point at ~10nM (FLAG-ER: 11.7nM, FLAG-LBD: 10.9nM), indicating that the amount of unlabeled estradiol need to occupy half the active proteins is very similar. The same phenomenon was found in the binding curves against Hexene. Despite the differences in the absolute CPM values -- in other words, the amount of the labeled estradiol bound to the proteins -- the curves, once again, almost overlapped (Figure 15B) and they showed similar inflection points at ~45nM (FLAG-ER: 41.2nM, FLAG-LBD: 34.1nM). These findings suggest that although one purification process was more successful than the other was, both preparations contained active proteins and that the two proteins display same binding affinity toward the ligands.

Lastly, the percent activity of the FLAG-LBD purified was calculated using the scintillation counts of the tritiated estradiol when the protein was reacted with the labeled estradiol only. The calculation was carried out using the obtained CPM value, the protein and the labeled estradiol amounts and taking into account the specific activity of the labeled estradiol (72Ci/mmol) and the counting efficiency of the scintillation counter

(~40%). It was estimated that approximately 89% of the proteins added to the reaction were bound to estradiol, therefore presumed active. However, the accuracy of this estimation was limited to the accuracy of the protein concentration concluded from the gel quantification as well as the purity and the activity of the tritiated estradiol.

Formation of the protein-drug-DNA complex

It had been demonstrated with the gel mobility shift assay discussed earlier chapter that when drug-damaged DNA is present, FLAG-LBD binds to the complex via the drug. However for an *in vivo* application, DNA and the estrogen receptor protein are already present in the cell and the drug is added. This experiment was performed to simulate the situation in the cell by adding the drug into a reaction where the DNA and the protein are already present. As mentioned in the Results chapter of this report, the drug seems to associate with the protein more often, especially when more protein is added to the reaction. It was not possible to deduce from this experiment whether the protein-drug complex can still bind to DNA, since all three states—B, C, and E in Figure 13-- would appear the same on the phosphorimager analysis. Thus, major bands in the phosophorimage can be either the results of B, E or the combination of the two. In either scenario, it indicates that the three-component complex can be formed if the ratio of the protein to the drug adducts is approximately 1.5:1.

Conclusion and Significance of Results

This study explored the one-step purification of recombinant proteins, FLAG-ER and FLAG-LBD. The proteins were purified to 70% homogeneity (in the case of FLAG-

LBD) and in large quantities. The proteins displayed binding to estradiol and to the free drugs containing an estradiol moiety and showed high activity. The FLAG-LBD protein also bound to drugs adducted to DNA, demonstrating that the protein-drug-DNA complex, whose formation is vital to the efficacy of estrogen receptor selective genotoxins, can be assembled. Also, it was shown that such complexes still form when the drugs are free to bind either the DNA or the protein first.

The production of the purified ER proteins also provides material for many other aspects of the studies pertaining to the estrogen receptor selective genotoxins. For example, purified proteins are a vital component of the repair assay being developed in the lab, which investigates the efficiency of the repair mechanisms towards DNA treated with the tumor selective drugs in the presence of the estrogen receptor.

Future directions

The results of this studies gave rise to many questions that should be further investigated. First, as mentioned earlier in this section, re-production of FLAG-ER virus in order to obtain proteins of higher purity and activity is necessary. Although according to the scheme of tumor-selective genotoxins, the region of the protein in direct contact with the drug and the DNA is the ligand binding domain and the formation of the protein-drug-DNA complex is demonstrated using the FLAG-LBD, the validity of the system cannot be assumed. The full-length estrogen receptor is approximately twice the size of the ligand binding domain purified during this study and the added bulk of the protein may have a slightly different interaction with the other components. Also, the DNA binding domain, the region missing from the ligand binding domain, may affect the

interactions differently. Therefore, it is essential that the full-length protein of good quality is obtained and further characterized.

Also, the kinetics of the protein-drug-DNA complex association should be studied more in depth. The formation of the complex is vital to the efficacy of the estrogen receptor selective genotoxins and it is important to understand the nature of the components. To follow up on the initial findings presented in this work, experiments elucidating the association kinetics of the three-component complex when the protein is already bound to the drug needs to be performed. Such information will be essential for the optimization of the drug dosing.

Another study that would give valuable insights into the molecular mechanism of the drugs is mutational studies of the protein. Different mutant forms of the estrogen receptor protein can be made and used to investigate the interaction between the protein and the drug. In order to achieve this, the residues of the proteins, which are suspected to be important in binding the ligands should be mutated by site-directed mutagenesis. Altering these residues could affect the protein's binding affinity towards the natural ligand and examining the mutations' effects on the drug-binding capability will improve our understanding of the interaction between the protein and the drugs on the molecular level.

There have been a number of studies with mutant estrogen receptor reported in the literature (30,32,33). Of these, two mutant proteins seem to be promising candidates. Glu353Gln mutant showed a 9-fold decrease in binding affinity towards estradiol (33) and His524Ala mutant displayed a 12-fold decrease in biding affinity towards the ligand (32). They were two of the three amino acids identified as residues important in binding

estradiol according to crystallography studies (12,13). Both form hydrogen bonds to the hormone, holding it diagonally across the cavity of the protein.

The construction of the plasmid DNA containing the mutated hER gene followed by the expression and the purification of the mutant proteins in the baculovirus expression system will provide these proteins to be used in the investigation.

References

1. Bruhn, S.L., Toney, J.H., and Lippard, S.J. (1990) Biological processing of DNA modified by platinum compounds. *Prog. Inorg. Chem.* **38**, 477-516.
2. Bruhn, S.L., Pil, P.M., Essigmann, J.M., Housman, D.E., and Lippard, S.J. (1992) Isolation and characterization of human cDNA clones encoding a high mobility group box protein that recognizes structural distortions to DNA caused by binding of the anticancer agent cisplatin. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 2307-2311.
3. Toney, J.H., Donahue, B.A., Kellett, P.J., Bruhn, S.L., Essigmann, J.M., and Lippard, S.J. (1989) Isolation of cDNAs encoding a human protein that binds selectively to DNA modified by the anticancer drug cis-diamminedichloroplatinum(II). *Proc. Natl. Acad. Sci. U. S. A.* **86**, 8328-8332.
4. Feuer, E.J., Brown, L.M., and Kaplan, R.S. B.A. Miller, L.A.G. Ries, B.F. Hankey, C.L. Kosary, A. Harras, S.S. Devesa, and B.K. Edwards (eds.) (1993) SEER Cancer Statistics Review: 1973-1990, pp.XXIV.1-XXIV.13, National Cancer Institute: Bethesda, MD.
5. Rink, S.M., Yarema, K.J., Paige, L.A., Tadayoni-Rebek, B.M., Solomon, M.S., Essigmann, J.M., and Croy, R.G. (1996) Synthesis and biological activity of DNA damaging agents that form decoy binding sites for the estrogen receptor. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 15063-15068.
6. Yamamoto, K.R. (1985) Steroid receptor regulated transcription of specific genes and gene networks. *Annu. Rev. Genet.* **19**, 209-252.

7. Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J., Argos, P., and Chambon, P. (1986) Human oestrogen receptor cDNA: sequence, expression and homology to v-*erb*-A. *Nature* **320**,134-139.

8. Krust, A., Green, S., Argos, P., Kumar, V., Walter, P., Bornert, J., and Chambon, P. (1986) The chicken oestrogen receptor sequence: homology with v-*erb*A and the human oestrogen and glucocorticoid receptors. *EMBO J.* **5**, 891-897.

9. Mangelsdorf, D.J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P., and Evans, R.M. (1995) The nuclear receptor superfamily: the second decade. *Cell* **83**, 835-839.

10. Tora, L., Mullick, A., Metzger, D., Ponglikitmongkol, M., Park, I., and Chambon, P. (1989) The cloned human oestrogen receptor contains a mutation which alters its hormone binding properties. *EMBO J.* **8**, 1981-1986.

11. Katzenellenbogen, J.A. and Katzenellenbogen, B.S. (1996) Nuclear hormone receptors: ligand-activated regulators of transcription and diverse cell responses. *Chem. Biol.* **3**, 529-536.

12. Tanenbaum, D.M., Wang, Y., Williams, S.P., and Sigler, P.B. (1998) Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 5998-6003.

13. Brzozowski, A.M., Pike, A.C.W., Dauter, Z., Hubbard, R.E., Bonn, T., Engstrom, O., Ohman, L., Greene, G.L., Gustafsson, J., and Carlquist, M. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* **389**, 753-758.

14. O'Reilly, D.R., Miller, L.K., and Luckow, V.A. (1992) Baculovirus Expression Vectors: A laboratory Manual. W.H. Freeman, New York.

15. Luckow, V.A. and Summers, M.D. (1988) Trends in the development of baculovirus expression vectors. *Bio/Technology* **6**, 47-55.

16. Koh, S.S., Hengartner, C.J., and Young, R.A. (1997) Baculoviral transfer vectors for expression of FLAG® fusion proteins in insect cells. *BioTechniques* **23**, 622-627.

17. Hasnain, S.E., Jain, A., Habib, S., Ghosh, S., Chatterji, U., Ramachandran, A., Das, P., Venkaiah, B., Pandey, S., Liang, B., Ranjan, A., Natarajan, K., and Azim, C.A. (1997) Involvement of host factors in transcription from baculovirus very late promoters—a review. *Gene* **190**, 113-118.

18. Merrington, C.L., Bailey, M.J., and Possee, R.D. (1997) Manipulation of baculovirus vectors. *Mol. Biotech.* **8**, 283-297.

19. Ayres, M.D., Howard, S.C., Kuzio, J., Lopez-Ferber, M., and Possee, R.D. (1994) The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. *Virology* **202**, 586-605.

20. Smith, G.E., Vlak, J.M., and Summers, M.D. (1983) Physical analysis of *Autographa californica* nuclear polyhedrosis virus transcripts for polyhedrin and 10,000-molecular weight protein. *J. Virol.* **45**, 215-225.

21. Williams, G.V., Rohel, D.Z., Kuzio, J., and Faulkner, P. (1989) A cytopathological investigation of *Autographa californica* nuclear polyhedrosis virus p10 gene function using insertion/deletion mutants. *J. Gen. Virol.* **70**, 187-202.

22. BacPAK™ Baculovirus expression system Product protocol. Clontech Laboratories, Inc.: Palo Alto, CA.

23. Vaughn, J.L., Goodwin, R.H., Tompkins, G.J., and McCawley, P. (1977) The establishment of two cell lines from the insect *Spodoptera* frugipera (Lepidoptera: Noctuidae). *In Vitro* **13**, 213-217.

24. Kitts, P.A., Ayres, M.D., and Possee, R.D. (1990) Linearisation of baculovirus DNA enhances the recovery of recombinant virus expression vectors. *Nucleic Acids Res.* **18**, 5667-5672.

25. Possee, R.D., Sun, T.P., Howard, S.C., Ayres, M.D., Hill-Perkins, M., and Gearing K.L. (1991) Nucleotide sequence of the *Autographa californica* nuclear polyhedrosis virus 9.4kbp *Eco*RI-I and R (polyhedrin gene) region. *Virology* **185**, 229-241.

26. Bradford, M.M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye binding. *Anal. Biochem.* **72**, 248-254.

27. Nardulli, A.M., Lew, D., Erijman, L., and Shapiro, D.J. (1991) Purified estrogen receptor DNA binding domain expressed in *Escherichia coli* activates transcription of an estrogen-responsive promoter in cultured cells. *J. Biol. Chem.* **266**, 24070-24076.

28. Seielstad, D.A., Carlson, K.E., Katzenellenbogen, J.A., Kushner, P.J., and Greene, G.L. (1995) Molecular characteriziation by mass spectrometry of the human estrogen receptor ligand-binding domain expressed in *Escherichia coli*. *Mol. Endocrinol.* **6**, 647-658.

29. Obourn, J.D., Koszewski, N.J., and Notides, A.C. (1993) Hormone- and DNA-binding mechanisms of the recombinant human estrogen receptor. *Biochemistry* **32**, 6229-6236.

30. Reese, J.C. and Katzenellenbogen, B.S. (1991) Mutagenesis of cysteines in the hormone binding domain of the human estrogen receptor. *J. Biol. Chem.* **266**, 10880-10887.

31. Beekman, J.M., Cooney, A.J., Elliston, J.F., Tsai, S.Y., and Tsai, M.J. (1994) A rapid one-step method to purify baculovirus-expressed human estrogen receptor to be used in the analysis of the oxytocin promoter. *Gene* **146**, 285-289.

32. Ekena, K., Weis, K.E., Katzenellenbogen, J.A., and Katzenellenbogen, B.S. (1996) Identification of amino acids in the hormone binding domain of the human estrogen receptor important in estrogen binding. *J. Biol. Chem.* **271**, 20053-20059.

33. Ekena, K., Katzenellenbogen, J.A., and Katzenellenbogen, B.S. (1998) Determinants of ligand specificity of estrogen receptor-α estrogen *versus* androgen discrimination. *J. Biol. Chem.* **293**, 693-699.

FIGURES

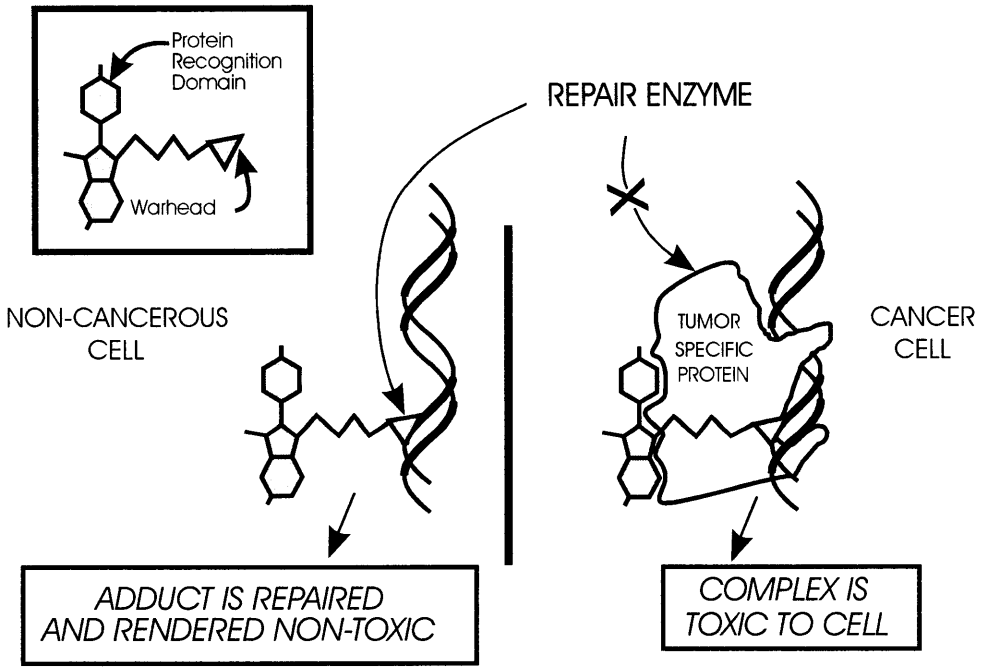


Figure 1. Design of programmable anti-cancer drugs

This is the underlying scheme of the design of the tumor-specific drug. The box inset shows the basic design of the drug which consists of two domains. The "warhead" domain reacts with the DNA forming an adduct. The protein recognition domain is derived from the ligand of the protein abundant in the cancer cells. In a cancer cell, the lesion formed by the drug would also bind the tumor specific protein, blocking the damaged site from the DNA repair enzymes. This complex would be highly toxic to the cell. However, in a normal cell where such protein is not expressed, the drug adduct would be readily repaired, showing no toxicity.

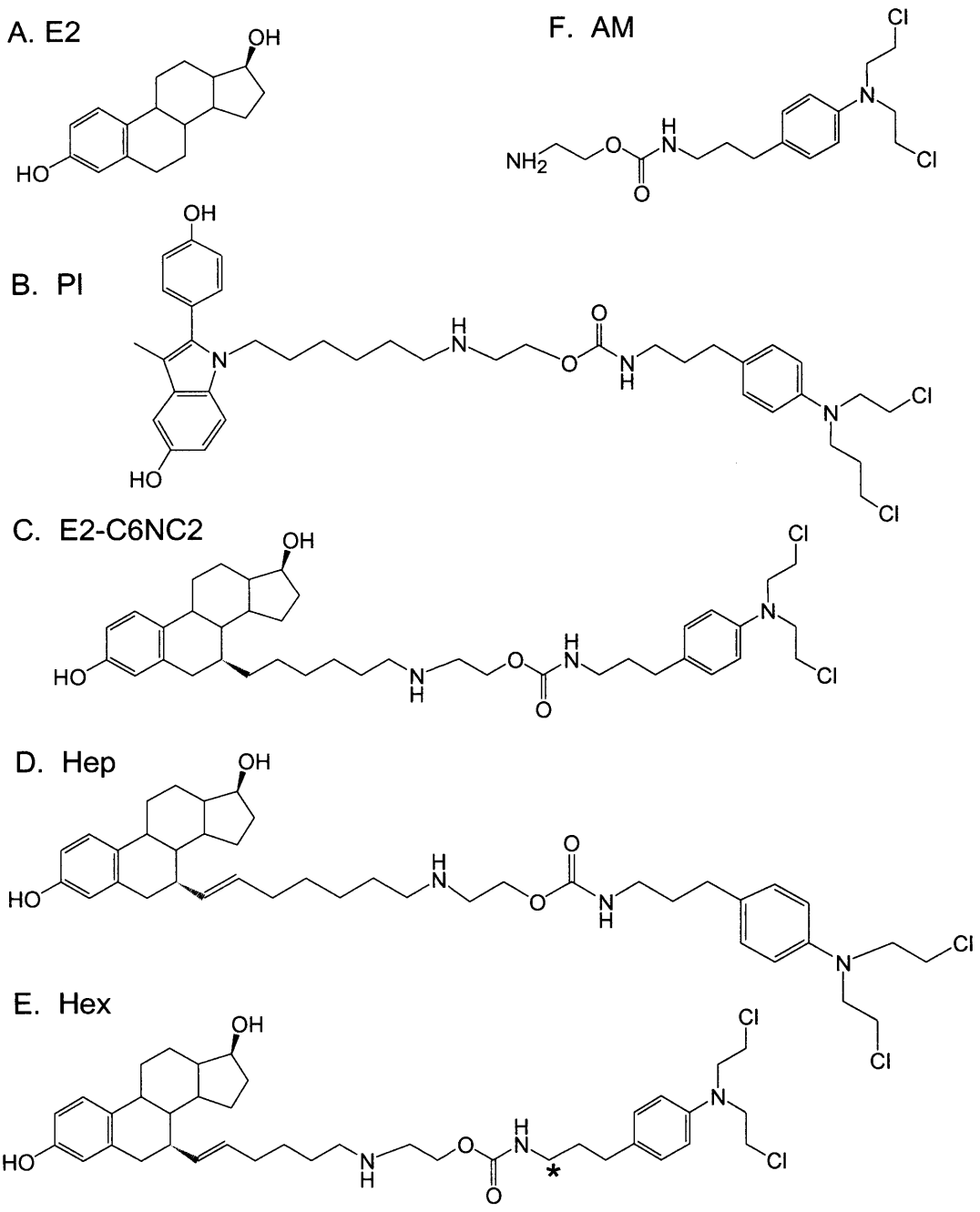


Figure 2. Chemical structures of estradiol and the drugs used in this study
All drugs are composed of the DNA binding domain, a nitrogen mustard, and the protein recognition domain. The protein binding domain is either an estradiol or otherhormone- like moiety. A. Estradiol, the natural ligand of the estrogen receptor protein. B. PI-C6NC2, the protein recognition domain is a phenylindole. C. E2-C6NC2, same as B except the phenylindole is replaced with estradiol. D. E2-Heptene, the alkene derivative of the 7-carbon linker. E. E2-Hexene, the alkene derivative of C. * indicates the position of the 14C label in [14C]Hexene. F. Amine mustard (AM), the control compound lacking the protein recognition domain and the carbon linker leading up to the amine.

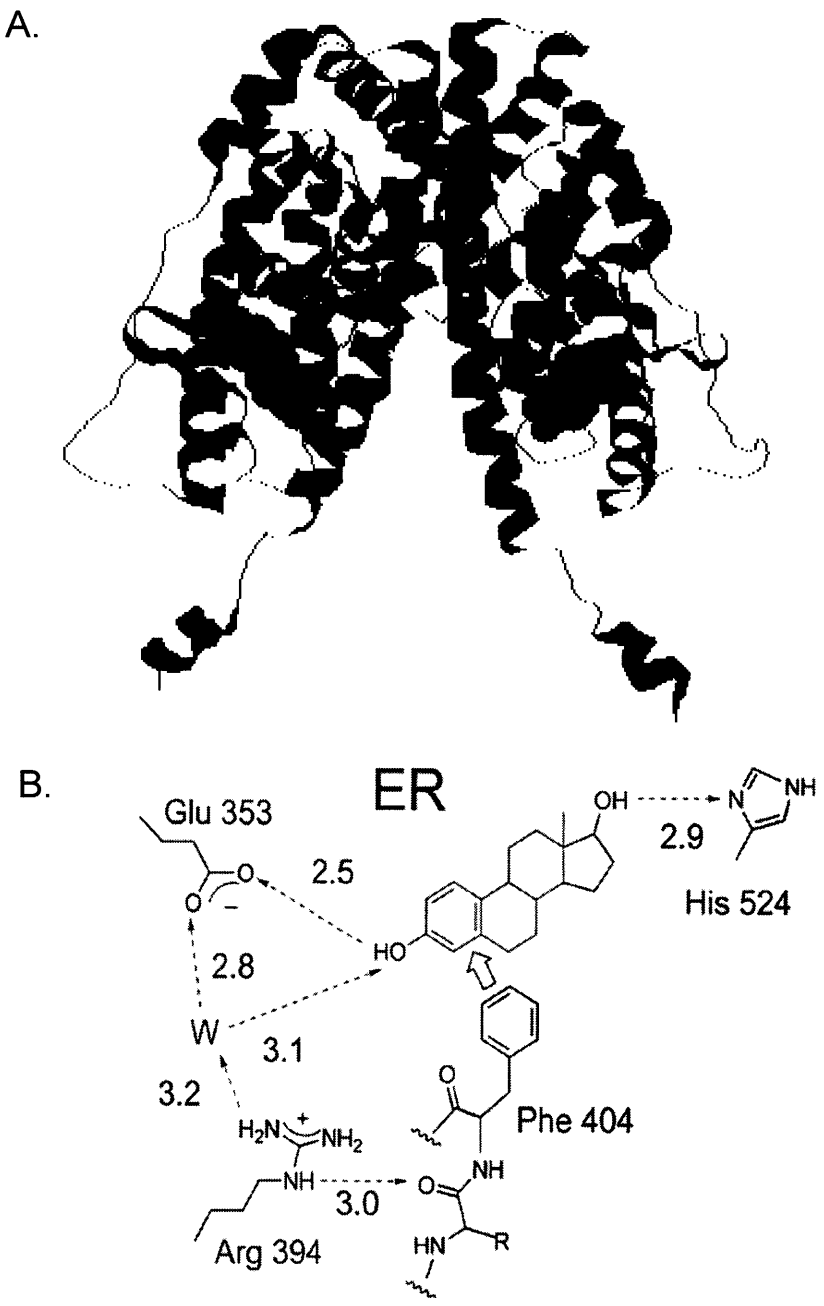


Figure 3. Crystal structure of hER-LBD and the ligand binding pocket

A. The ribbon representation of the ligand binding domain of the estrogen receptor. The structure shown is a homodimer of the protein containing 2 molecules of estradiol The ligand is represented as red space-fill model. The structure is rendered in Rasmol from the crystallography data published by the Sigler group (12). B. The ligand binding pocket of the hER. The interactions between estradiol and the residues of the protein are illustrated. The D-ring hydroxyl group of estradiol forms a hydrogen bond with His524. The phenyl hydroxyl group forms hydrogen bonds with Arg394 via a water molecule and directly with Glu353 of the protein. The figure was borrowed from Tanenbaum, et al. (12).

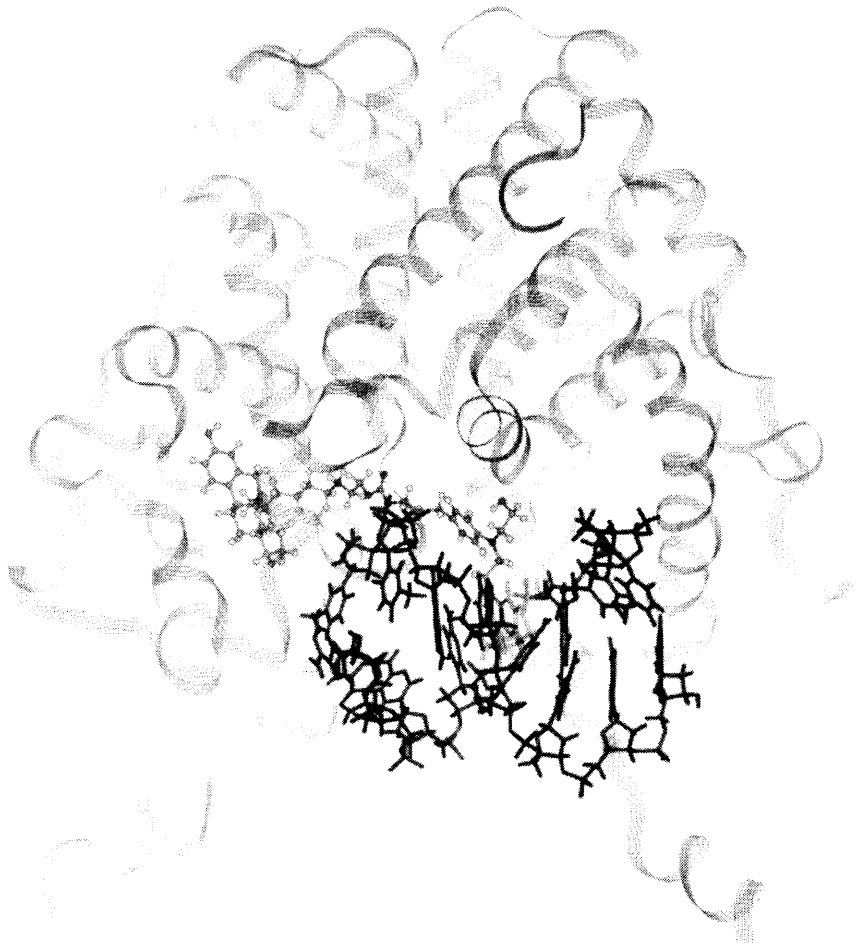


Figure 4. Model of E2-C6NC2-modified DNA bound to hER-LBD
The structure of the DNA modified by the drug, E2-C6NC2, was docked into the binding pocket of the hER-LBD. The structure of the protein used was solved by the Sigler group (12). The model was obtained by performing a energy-minimization in water using molecular modeling program, InsightII. The protein backbone is rendered as ribbon, the drug molecule is shown as a ball-and-stick figure and the DNA helix is represented as sticks.

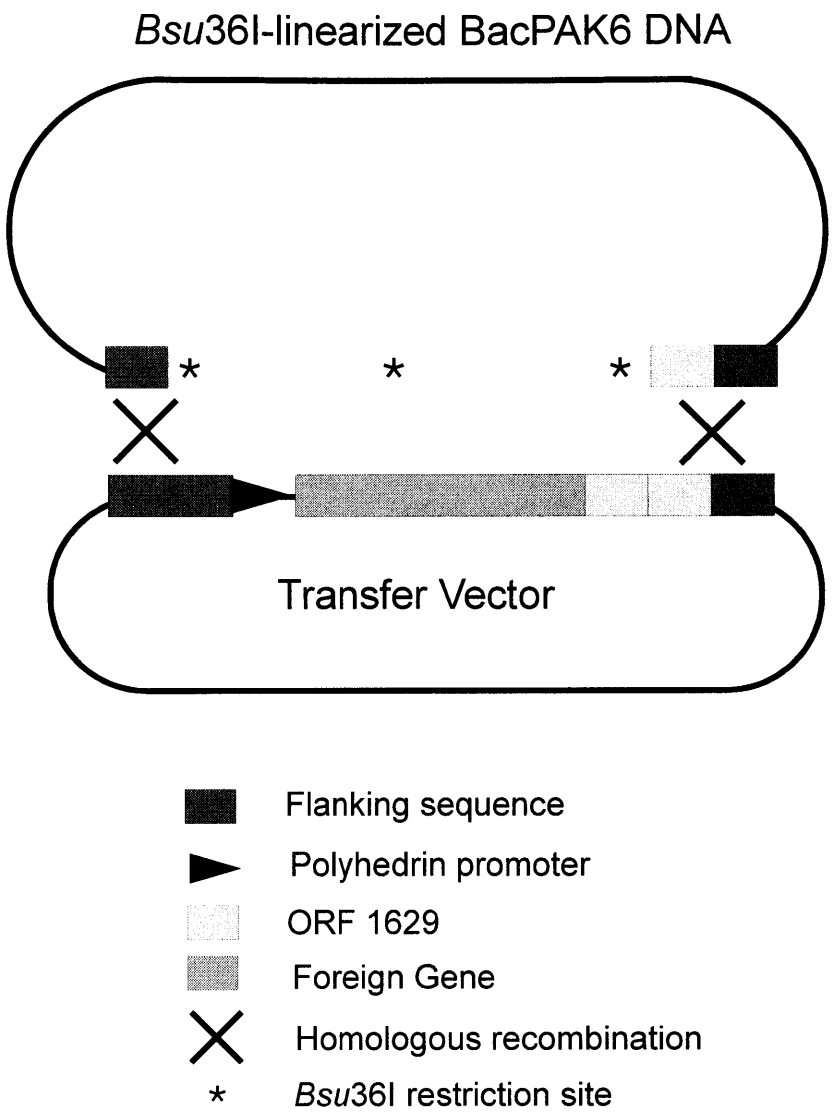


Figure 5. AcMNPV genome and co-transfection scheme

The gene of interest is inserted into the baculovirus transfer vector (pBacPAK8/9, pSK277/278) containing the intact ORF1629 gene, polyhedrin promoter and the flanking sequences. The plasmid is co-transfected with Bsu36I-digested AcMNPV viral DNA (BacPAK6, Clontech). Homologous recombination following the co-transfection into Sf21 cells restores the ORF1629 and inserts the polyhedrin promoter and the foreign gene, producing a viable virus. The genes of interest for this study were hER and the ligand binding domain of hER. Figure was adopted from Merrington, et al. (18).

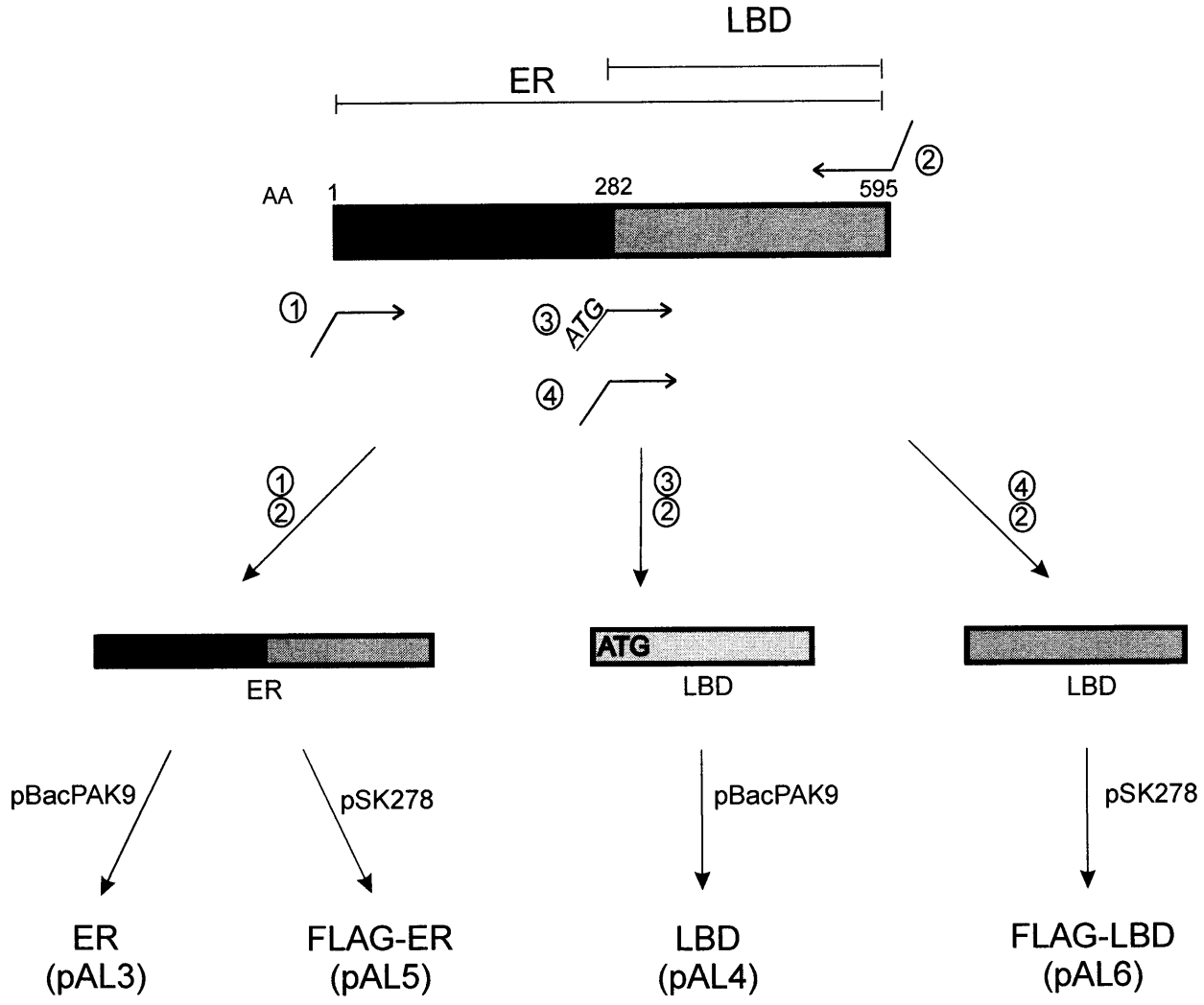


Figure 6. Cloning scheme of ER and LBD in baculovirus expression vectors

A PCR reaction with primers ALO1(①) and ALO2(②) gave the full-length estrogen receptor (ER) gene (1.8 kb The fragment was ligated into the baculovirus expression vectors pBacPAK9 and pSK278 to give pAL3 and pA respectively. pAL3 was the construct for the un-tagged ER and pAL5 was the construct for FLAG-ER. Similar PCR reaction with ALO3(③) and ALO2 gave the ligand binding domain (LBD) fragment (0.94 kb) containing th start codon (ATG). The fragment in pBacPAK9 yielded pAL4 which was the construct for the un-tagged LBD. The LBD fragment (0.94 kb) produced by the PCR reaction with ALO4(④) and ALO2 was inserted into pSK27 give pAL6, the construct for FLAG-LBD.

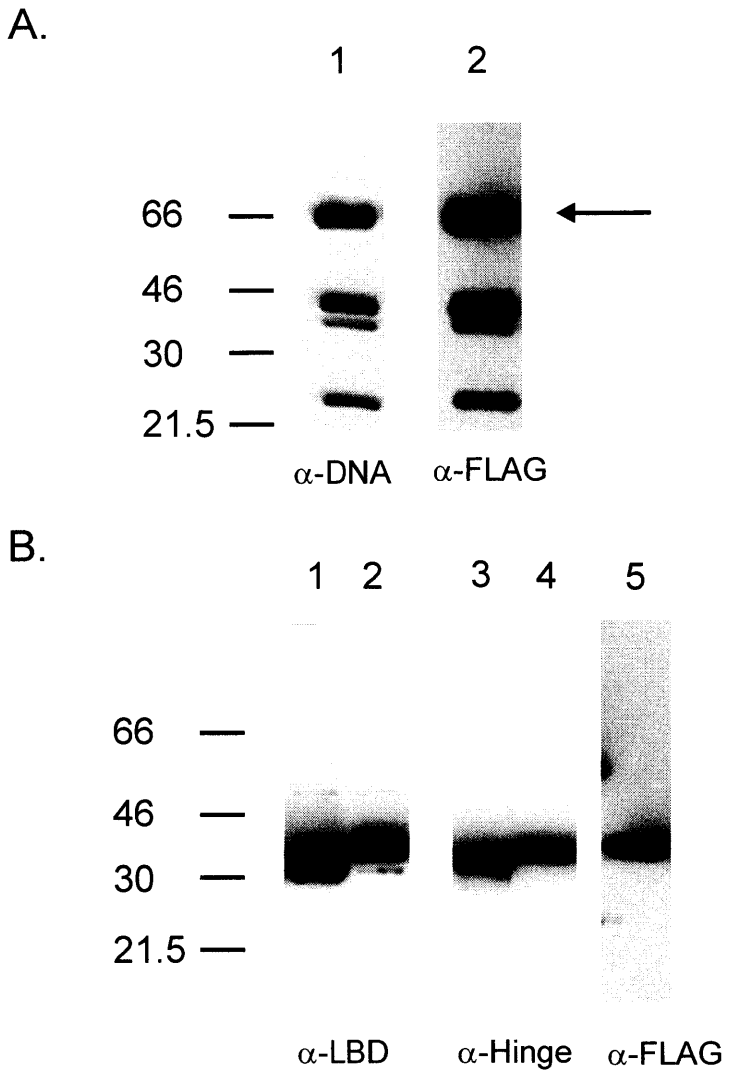


Figure 7. Western blot analysis of the recombinant proteins

A. The whole cell extract of one of the clones (clone #2) expressing FLAG-ER was resolved on a SDS-PAGE gel. It was blotted against the anti-ER antibody raised against the DNA binding domain (lane 1) and against the anti-FLAG antibody (lane 2). The lines indicate the position of the molecular marker standards and the arrow indicates the 67 kDa FLAG-ER.

B. The whole cell extracts of the cells expressing LBD (lanes 1 and 3) and FLAG-LBD (lanes 2, 4 and 5) were separated on a SDS-PAGE gel side by side. They were blotted against the anti-ER antibody raised against the ligand binding domain (lanes 1 and 2) and against the hinge region (lanes3 and 4). FLAG-LBD whole cell extract was also blotted against the anti-FLAG antibody (lane 5). The lines indicate the positions of the molecular weight standards.

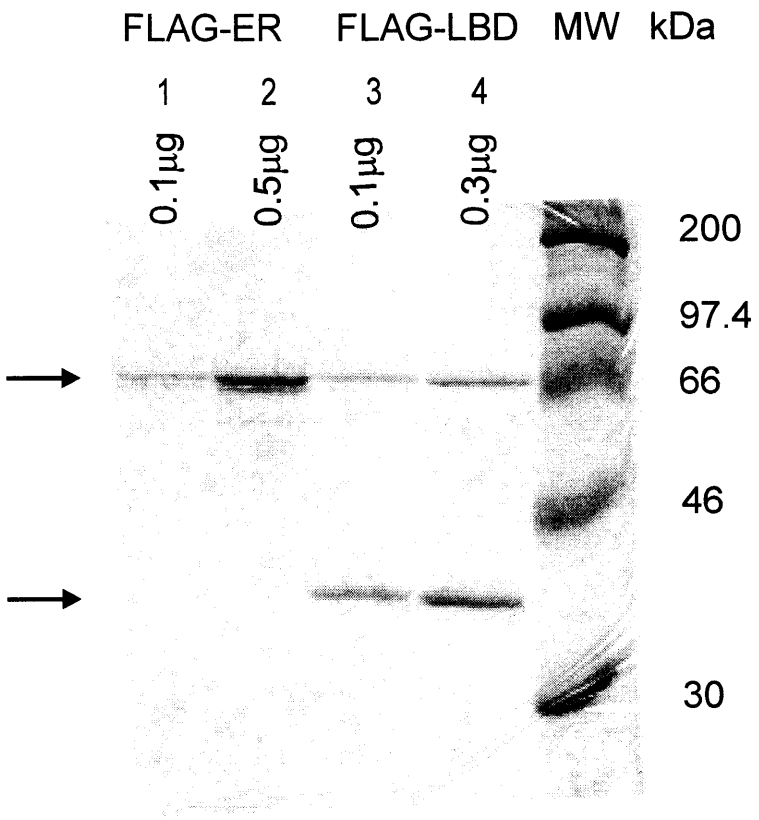
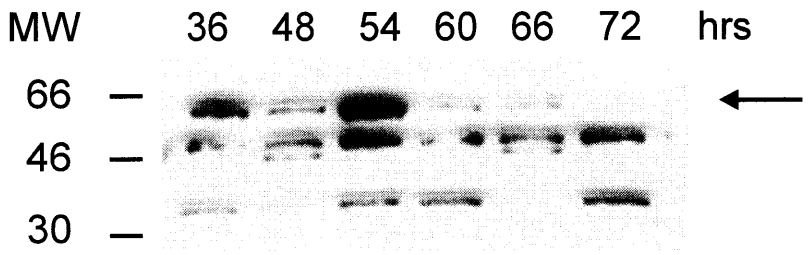


Figure 8. Purified FLAG-ER and FLAG-LBD

The purified proteins were evaluated by SDS-PAGE analysis. The protein bands were visualized by Coomassie staining. Lanes 1 and 2 were FLAG-ER and lanes 3 and 4 were FLAG-LBD. The numbers above the panel were the amount of the desired proteins in each lane, estimated by comparing to BSA solutions of known concentrations. The numbers on the right show the molecular weight marker sizes. The arrows on the left indicate the positions of the proteins, where the upper arrow points to FLAG-ER and the lower to FLAG-LBD.

A. FLAG-ER



B. FLAG-LBD

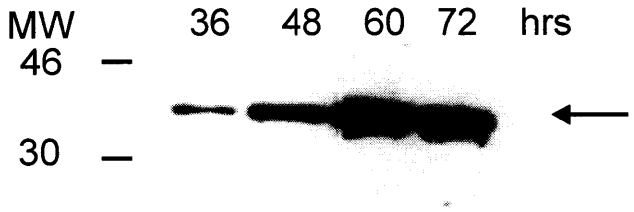


Figure 9. Time course expression of the recombinant proteins

The Western blots of the FLAG-ER (Panel A) and the FLAG-LBD (Panel B) are shown above. The numbers above the panels indicate the infection time in hours. The cells harvested at the designated time post-infection were made into whole cell extracts and the extracts were resolved on SDS-PAGE gels. They were blotted against the anti-ER (raised against LBD) antibody. The lines represent the position of the molecular weight standards and the arrows denote the FLAG-ER and FLAG-LBD proteins.

Figure 10

A.

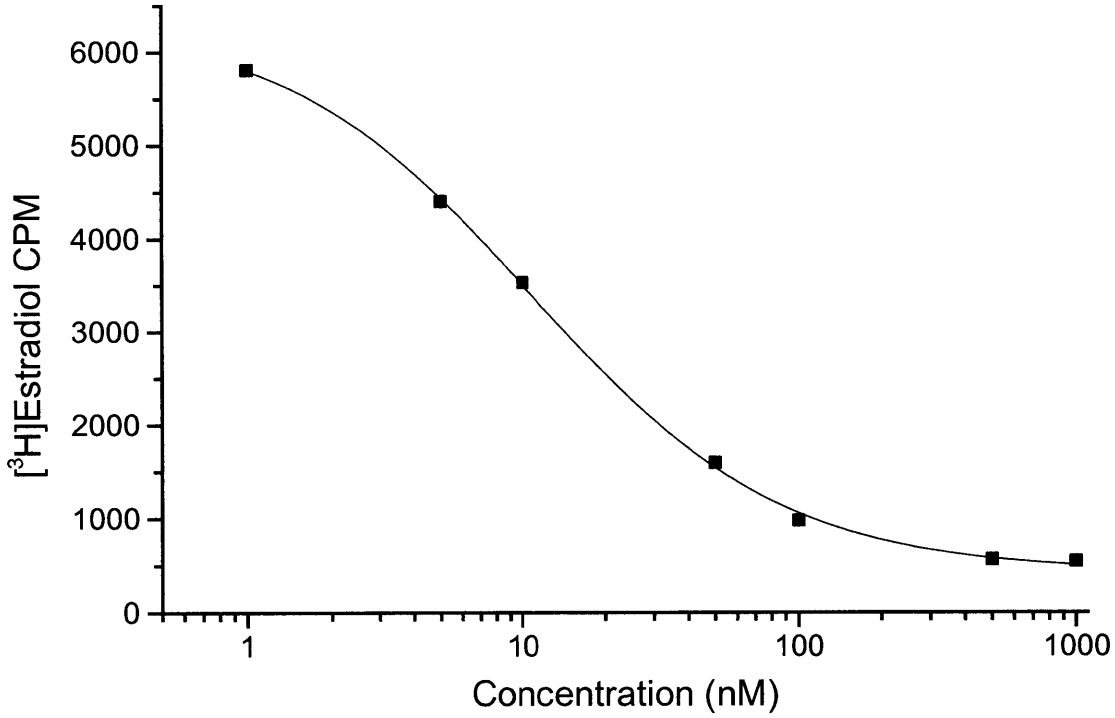

6000
5000
4000
3000
2000
1000
0
[3H]Estradiol CPM
1
10
100
1000
Concentration (nM)

Figure 10

B.

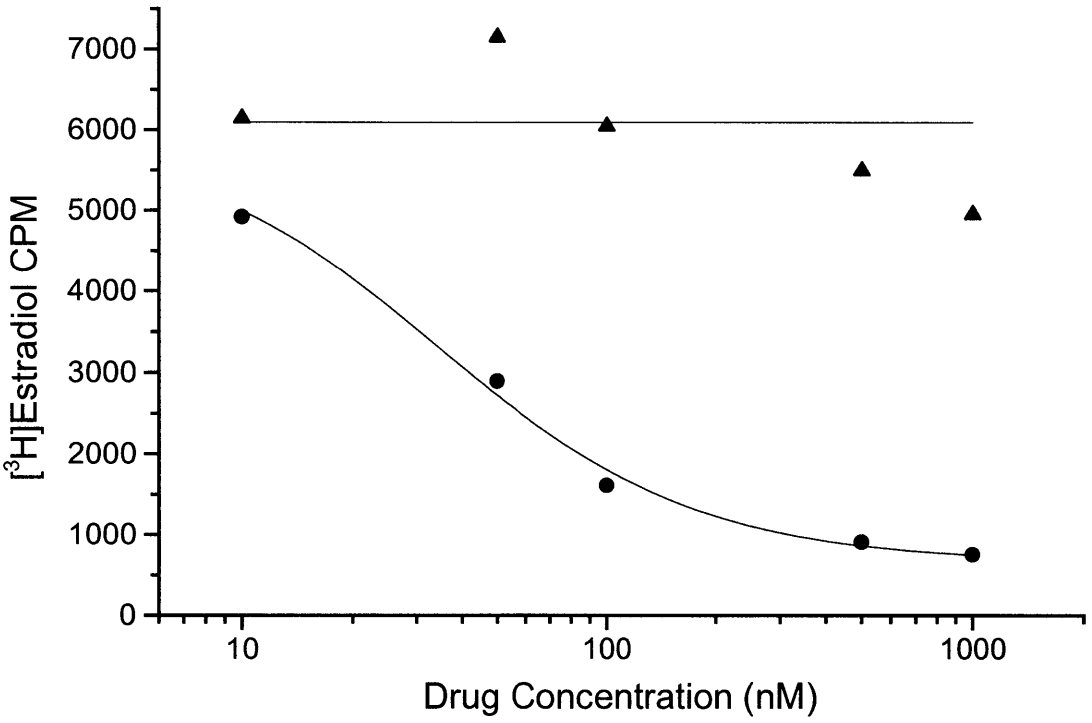

7000
6000
5000
4000
3000
2000
1000
0
[3H]Estradiol CPM
10
100
1000
Drug Concentration (nM)

Figure 10

C.

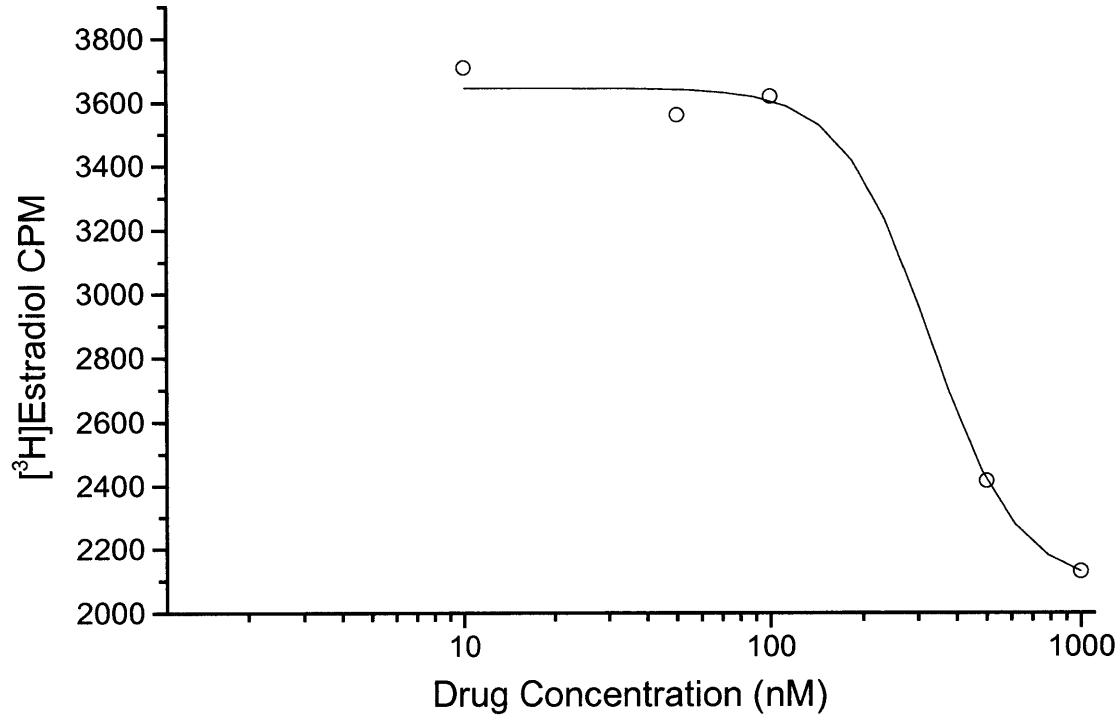


Figure 10. FLAG-LBD competition assays
The binding of different ligands to the FLAG-LBD is shown. The amount of radiolabeled estradiol bound to the protein was plotted as a function of unlabeled ligand concentration. The competing ligand was unlabeled estradiol (A), Hexene (B, closed circle), AM (B, closed triangle) or Heptene (C).

Figure 11

A.

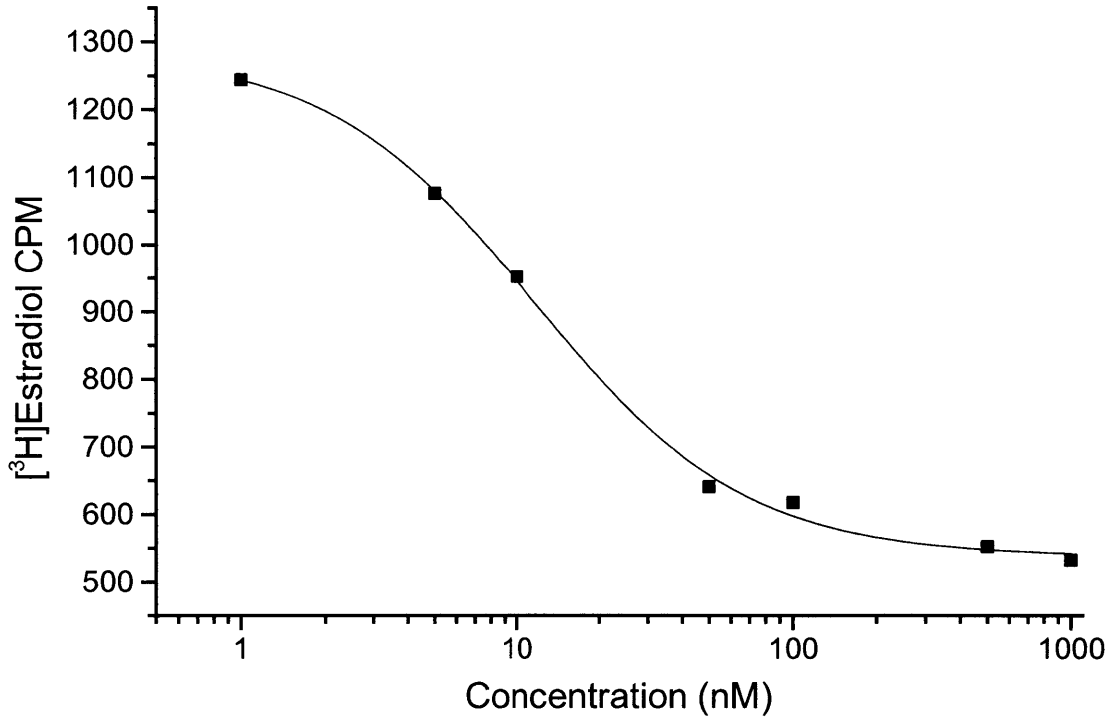

1300
1200
1100
1000
900
800
700
600
500
[3H]Estradiol CPM
1
10
100
1000
Concentration (nM)

Figure 11

B.

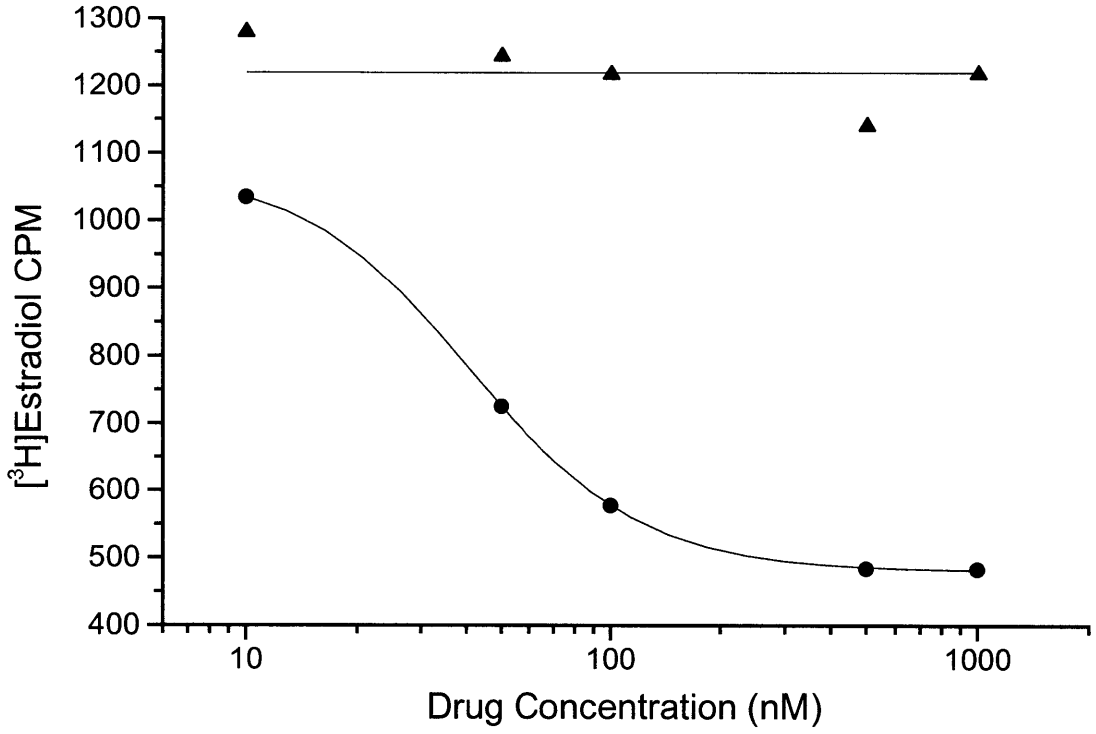


Figure 11. FLAG-ER competition assays
The binding of different ligands to the FLAG-ER is shown. The amount of radiolabeled estradiol bound to the protein was plotted as a function of unlabeled ligand concentration. The competing ligand was unlabeled estradiol (A), Hexene (B, closed circle), or AM (B, closed triangle).

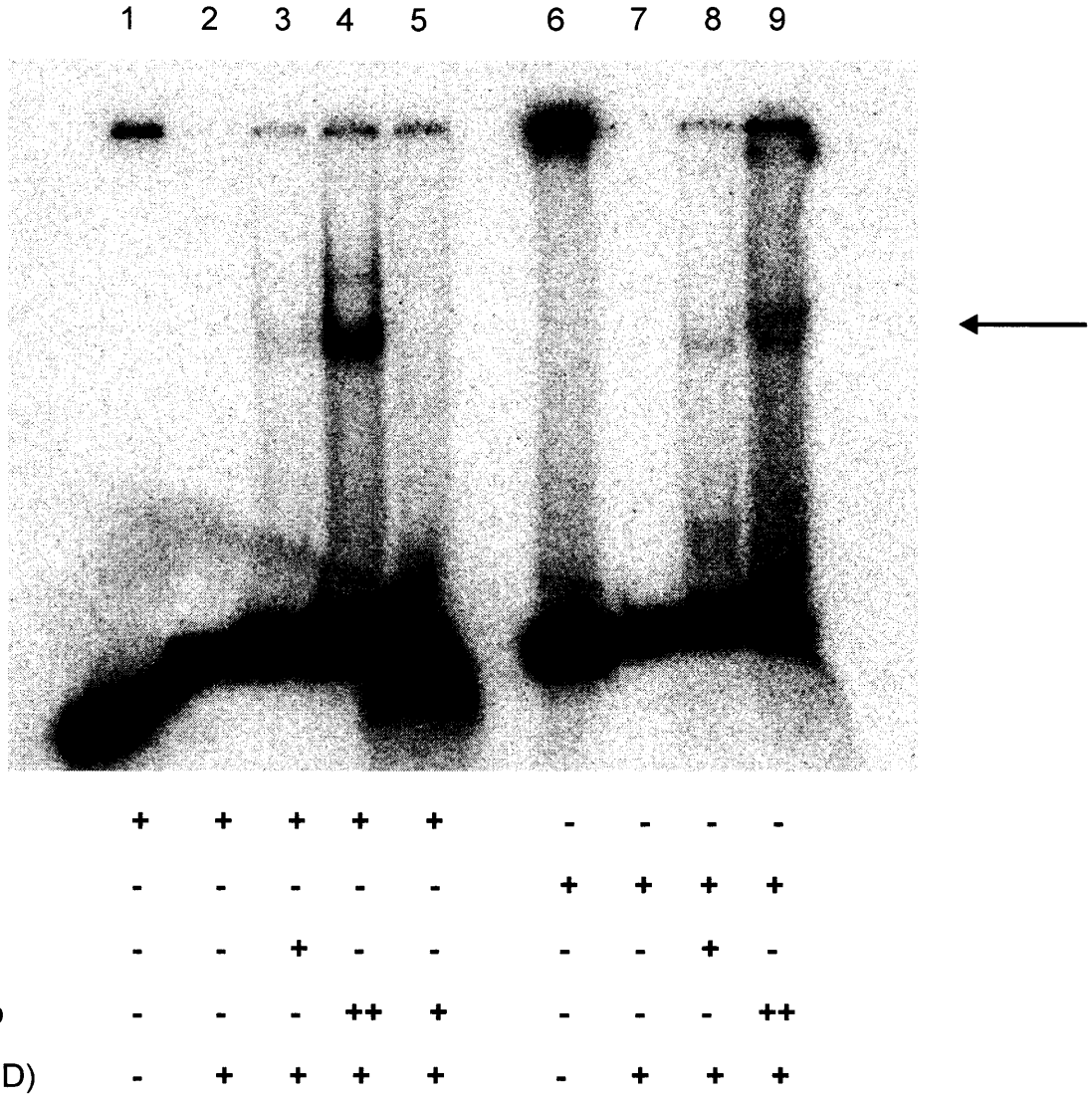


Figure 12. Gel shift mobility assay

The association of the protein and the DNA duplex via estrogen receptor selective drugs was demonstrated. The DNA duplex with radiolabeled C1 oligo (C1D) was used for lanes 1 through 5 and DNA duplex with radiolabeled C2 oligo (C2D) was used for lanes 6 through 9. Lanes 1 and 6 were negative controls containing only DNA duplexes whereas lanes 2 and 7 were controls with DNA duplex and FLAG-LBD protein. Phenylindole (PI) was added to the reactions in lanes 3 and 8. For Heptene (Hep), + indicates Heptene monoadduct on the DNA (lane 5) and ++ represents multiple adducts of Heptene formed on DNA duplexes (lanes 4 and 9). The arrow denotes the shifted bands representing the DNA-drug-protein complexes.

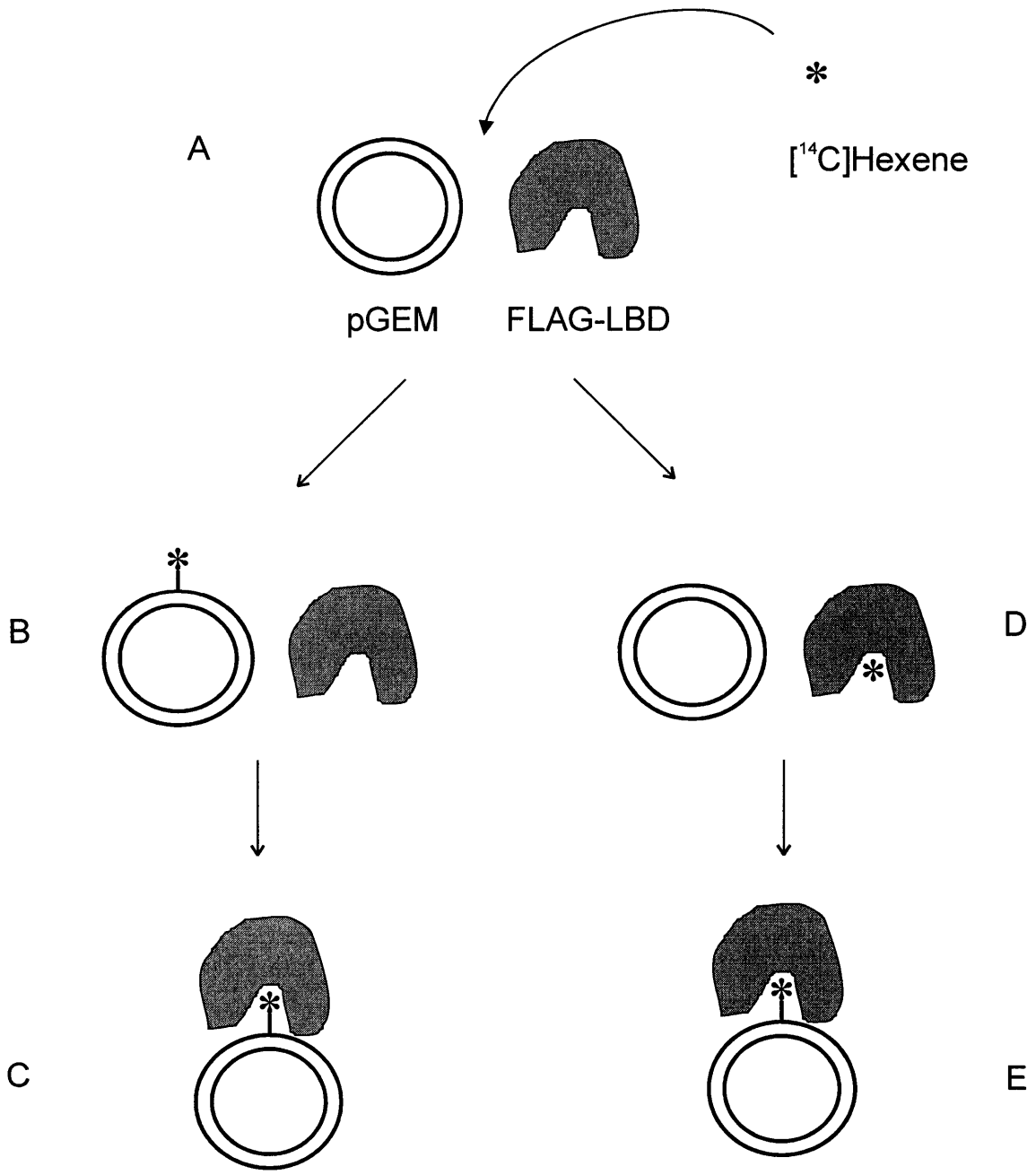


Figure 13. Possible association kinetics of the protein-drug-DNA complex
The possible pathways of interactions among FLAG-LBD, [^{14}C]Hexene and pGEM are shown. It is shown that the protein binds to the drug when the drug is already bound to the plasmid (B to C), but it is not known if the drug is capable of binding the plasmid after it has associated with the protein (D to E). Also, it is not known if the drug associates with the protein or the DNA preferentially.

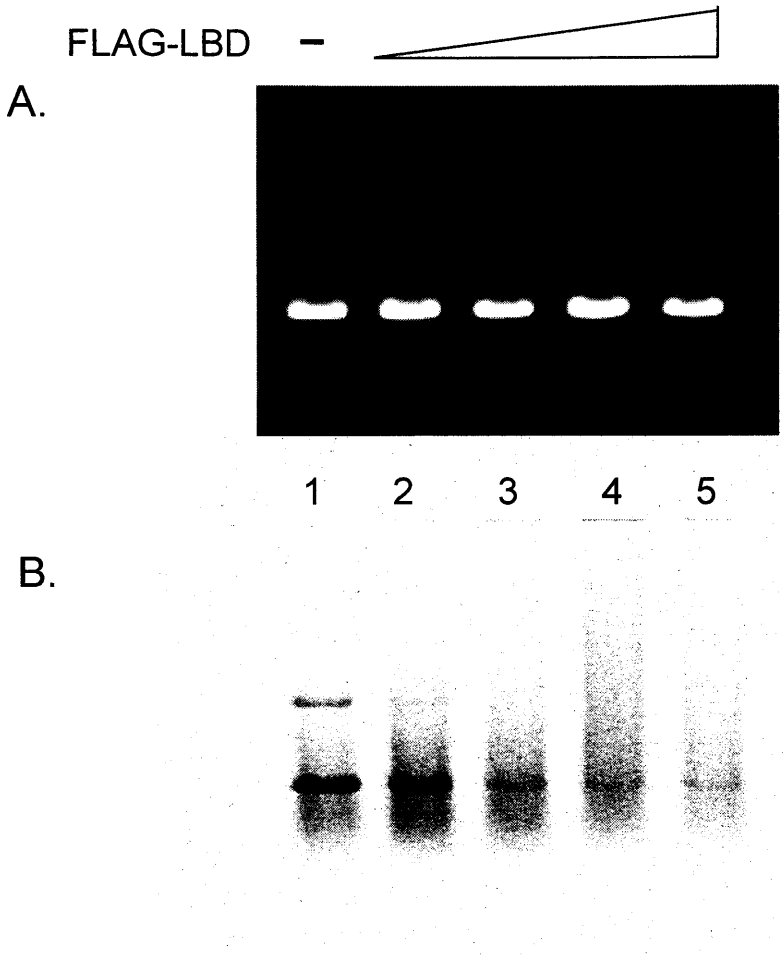


Figure 14. Association kinetics of the protein-drug-DNA complex
The association reaction carried out with different amount of FLAG-LBD protein was resolved on a 0.8% agarose gel. The plasmid DNA is visualized with UV light (A) and the [^{14}C]Hexene adducted to the DNA was visualized by phosphorimager analysis (B). The amount of protein in each lane was 0, 0.15, 0.3, 0.45 and 0.6 μg, which was 0, 1.5, 3, 4.5, 6 fold of the number of the drug adducts estimated to be in each reaction. The phosphorimager screen was exposed for 36 hours.

Figure 15

A.

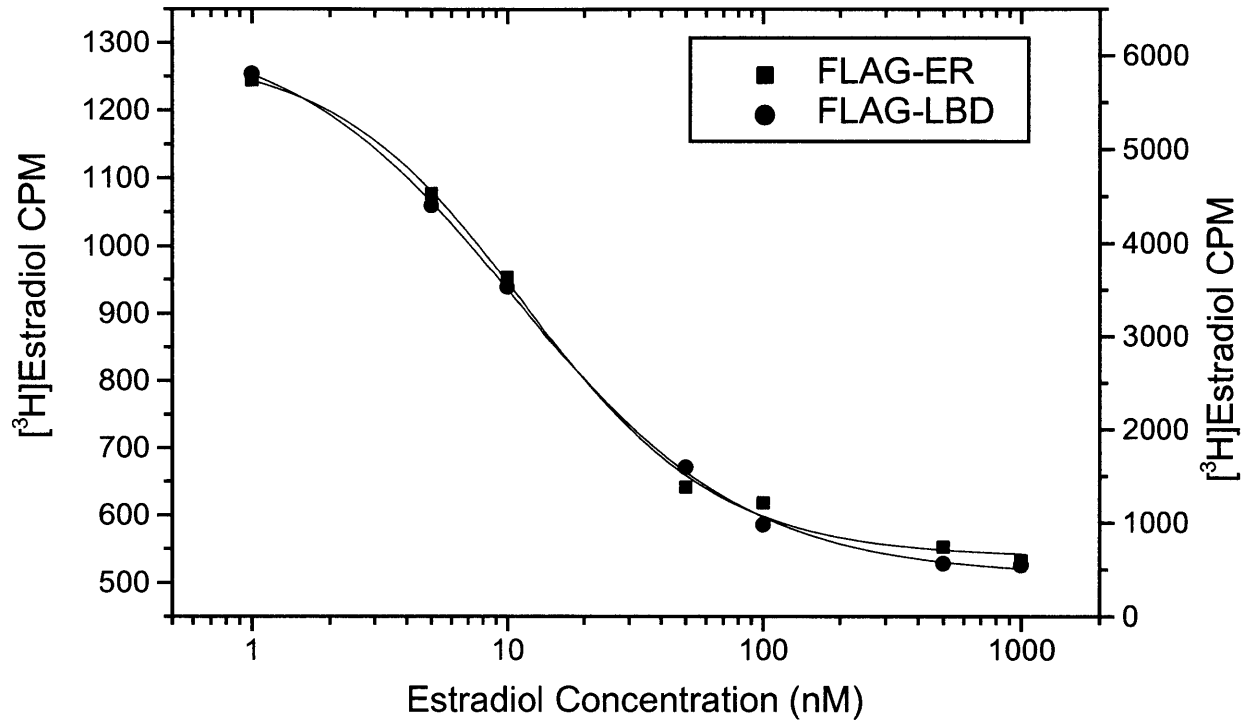

FLAG-ER
FLAG-LBD
[3H]Estradiol CPM
[3H]Estradiol CPM
Estradiol Concentration (nM)
1300
1200
1100
1000
900
800
700
600
500
6000
5000
4000
3000
2000
1000
0
1
10
100
1000

Figure 15

B.

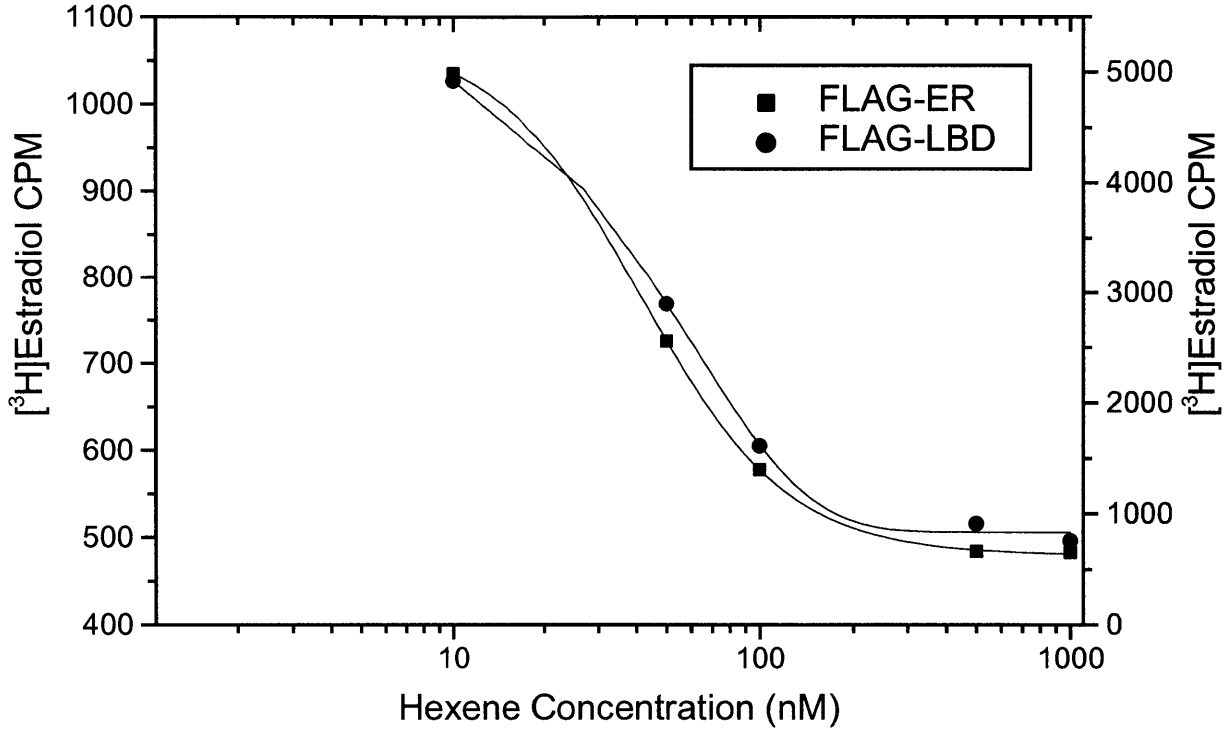


Figure 15. Comparison of FLAG-ER and FLAG-LBD's ligand binding capacity
A. Figure 10A and Figure 11A are overlapped on a double y-axis graph to show the resemblance of the two curves indicating the ability to bind estradiol. The left y-axis is FLAG-ER and the right y-axis is FLAG-LBD. Although the span of the [^{3}H]Estradiol CPM values are different, the shape and the inflection points are almost identical (Inflection point of FLAG-ER: 11.7 nM and of FLAG-LBD: 10.9 nM).
B. Overlap of Figure 10B and Figure 11B. As in panel A, these curves compare the two proteins' binding capacity for Hexene (Inflection point of FLAG-ER: 41.2 nM and of FLAG-LBD: 34.1 nM)

TABLES

Protein		Baculovirus	Plasmid	
Name	Size (kDa)		Name	Size (kb)
FLAG-ER	67	ALBV1	pAL5	7.3
FLAG-LBD	33	ALBV2	pAL6	6.5
ER	67	ALBV3	pAL3	7.3
LBD	33	ALBV4	pAL4	6.5

Table 1. Plasmid and baculovirus used for protein expression

The names and the sizes of the plasmids from which the baculoviruses were derived are listed. Also the names and the sizes of the proteins expressed by the baculoviruses are listed.