ELSEVIER

Journal of Molecular Graphics and Modelling 20 (2001) 155–167

Journal of Molecular Graphics and Modelling

www.elsevier.com/locate/JMGM

Homology modeling of the estrogen receptor subtype β (ER-β) and calculation of ligand binding affinities

Robert K. DeLisle, Seong-Jae Yu, Anil C. Nair, William J. Welsh*

Department of Chemistry and Biochemistry, Center for Molecular Electronics, University of Missouri at St. Louis, 8001 Natural Bridge Road, St. Louis, Missouri, MO 63121, USA

Received 5 December 2000; received in revised form 19 June 2001; accepted 19 June 2001

Abstract

Estrogen is a steroid hormone playing critical roles in physiological processes such as sexual differentiation and development, female and male reproductive processes, and bone health. Numerous natural and synthetic environmental compounds have been shown capable of estrogenic effects. This area has been the focus of significant fundamental and applied research due both to the potential detrimental effects of these compounds upon normal physiological processes and to the potential beneficial effects of tissue-selective estrogen agonists/antagonists for the prevention and treatment of numerous diseases. Genomic effects of the active form of estrogen, 17β-estradiol, are mediated through at least two members of the steroid hormone receptor superfamily, estrogen receptor subtype α (ER-α) and estrogen receptor subtype β (ER-β). At the time of this work, the X-ray crystal structure of the ER-α had been elucidated, however, coordinates of the ER-β were not publicly available. Based upon the significant structural conservation across members of the steroid hormone receptor family, and the high sequence homology between ER-α and ER-β (>60%), we have developed a homology model of the ER-β structure. Using the crystal structure of ER-α and the homology model of ER-β, we demonstrate a strong correlation between computed values of the binding-energy and published values of the observed relative binding affinity (RBA) for a variety of compounds for both receptors, as well as the ability to identify receptor subtype selective compounds. Furthermore, using the recently available crystal structure of ER-β for comparison purposes, we show that not only is the predicted homology model structurally accurate, but that it can be used to assess ligand binding affinities.

Keywords: Estrogen; Relative binding affinity; Genomic effects

1. Introduction

The physiological roles of estrogen are extremely diverse and occupy points of critical significance in the normal processes of numerous tissues [1]. Classically considered the female sex hormone, estrogen plays a central role in the normal development and function of the female reproductive system. Estrogen is of central importance in the menstrual cycle and plays a significant role in pregnancy in conjunction with another steroid hormone, progesterone. In both sexes, proper development of bone as well as closure of the epiphyses of long bones are dependent upon estrogen, with significant developmental abnormalities occurring in its absence. Moreover, during menopause when estrogen levels in females decrease, bones can become decalcified, resulting in the serious condition of osteoporosis [2].

Numerous natural and synthetic chemicals have been shown capable of eliciting or antagonizing estrogenic responses in various species [3–8]. Synthetic chemical entities such as DDT, PCBs, and various insecticides and herbicides have been well documented as estrogenic both in vitro and in vivo. Furthermore, these compounds have been implicated in exerting a negative influence on the reproductive capability of wildlife such as birds, fish, reptiles, and amphibians, as well as affecting the normal development and distribution of animal genders [9–13]. The synthetic nature of these compounds has resulted in their persistence within the environment and accumulation through the natural food chain, raising concern over their potential impact on humans and wildlife [14]. The ability of these compounds to exert influence, both agonistic and antagonistic, upon estrogen and other endocrine systems has led to the term endocrine disrupters, and has been the focus of significant research and growing concern [15,16]. In addition to the synthetic estrogenic chemicals, a number of compounds have been found in plant products that also exhibit estrogen-like properties. These phytoestrogens have been shown to modulate the estrogen receptor of numerous species, and have been investigated for

* Corresponding author. Tel.: +1-314-516-5318; fax: +1-314-516-5342.
E-mail address: wwelsh@umsl.edu (W.J. Welsh).

1093-3263/01/$ – see front matter
PII: S1093-3263(01)00115-2

potential harmful effects as well as putative beneficial roles [17–23].

The genomic effects of estrogen are principally mediated through intracellular receptors [24]. Binding of estrogenic compounds to receptor molecules leads to alteration of the transcriptional status of specific genes leading ultimately to cellular alterations. Two different receptors responsible for the effects of estrogen have been characterized — estrogen receptor subtype α (ER-α) [25] and estrogen receptor subtype β (ER-β) [26]. As these two receptors appear to manifest differential tissue distributions, it is reasonable to propose that these receptors may also play differential roles within their respective cellular contexts. In support of this hypothesis, it has been demonstrated that distinct ligand recognition properties exist between the two receptors when assessed for their abilities to interact with, and become transcriptionally activated by, various natural estrogens as well as by endocrine disrupter compounds [27,28].

While crystal structures for ER-α exist and have been published, the crystallographic coordinates for the β isoform [29] were not publicly available at the time of this work. In order to provide a tool for investigation of the structural characteristics of ER-β, and to provide a method useful in the investigation of potential estrogenic compounds, we developed a homology model of the ER-β structure. This model is based upon the high degree of sequence conservation between ER-α and ER-β as well as the apparent strong conservation of structural characteristics across the family of steroid hormone receptors. Using this model, we are able to show a strong correlation between calculated values of the binding energy and experimentally determined values of the relative binding affinity (RBA) for a wide range of estrogenic compounds including steroids, phytoestrogens, and polychlorinated biphenyls. Furthermore, this model is shown to be comparable in both structure and performance in our computational procedure to the crystal structure of ER-β that was recently deposited in the RCSB-PDB 29. The present results support the application of homology modeling to other members of this superfamily of receptors.

2. Experimental procedures

2.1. Sequences and structures

The primary sequence of the human ER-β was obtained from GenBank, available from the National Center for Biotechnology Information (NCBI), accession number BAA24953. An additional sequence was found present in the GenBank database (accession number CAA67555) which exhibits an extended N-terminal sequence. Apart from the additional 53 residues present in the N-terminal, the two sequences are identical. Furthermore, due to our primary interest in the C-terminal ligand binding domain of the molecule, this sequence discrepancy had no impact upon our studies. Structural data for ER-α and ER-β were obtained from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB-PDB). Five independent structures were used to provide overlapping structural data, and consisted of ER-α in complex with estradiol (two structures), diethylstilbestrol, tamoxifen, and raloxifene (RCSB-PDB ID numbers 1ERE, 1A52, 3ERD, 3ERT, and 1ERR, respectively), while a single structure for ER-β was utilized (RCSB-PDB accession #1QKM and #1QKN). Additional structures used for intermolecular comparisons consisted of the progesterone receptor, peroxisome proliferator activated receptor gamma, retinoid X receptor alpha, and retinoic acid receptor gamma (RCSB-PDB ID numbers 1A28, 1PRG, 3LBD, and 1LBD, respectively; Fig. 1).

2.2. Sequence and structural comparisons

Searches for sequences homologous to that of ER-β were conducted using BLASTp and the SwissProt database, both available through NCBI, using default parameters. Further sequence and structural comparisons were performed using the homology module of Insight II via the options described below (Fig. 2). Sequences present in crystal structures were extracted from the RCSB-PDB file information and compared with that of ER-β using the multiple sequence alignment and pairwise sequence alignment operations. Structural comparisons were accomplished using the structural alignment option with default parameters. Subsequent to determination of an appropriate template structure (i.e. ER-α), sequences were aligned via multiple sequence alignment with default parameters. Various parametric adjustments were made in an effort to enhance the alignment; however, no significant improvements could be obtained beyond that resulting from default inputs.

2.3. Homology model construction

Coordinates were assigned to the residues of the ER-β ligand binding domain from the regions of greatest structural conservation among the ER-α structures as well as the additional receptor structures. Those regions which did not retain significant conservation across receptor structures (specifically, loop regions) were modeled by random generation of structural data followed by close comparison with available structures. The natural ligand for the estrogen receptor, 17β-estradiol, was included in the homology structure by superimposing the model with an estradiol — ER-α complex structure (1ERE). Docking of a newly constructed and energy minimized 17β-estradiol molecule was guided by the complexed ligand present in crystal structure data. In addition, a single water molecule has been identified which plays a significant role in mediating the hydrogen bond between the receptor and bound ligands in ER-α [30,31] as well as ER-β [29]. This water molecule also appears to be conserved and function similarly in other steroid receptors (e.g. the progesterone receptor [31,32]), consequently, a water molecule was docked into the ER-β homology model through

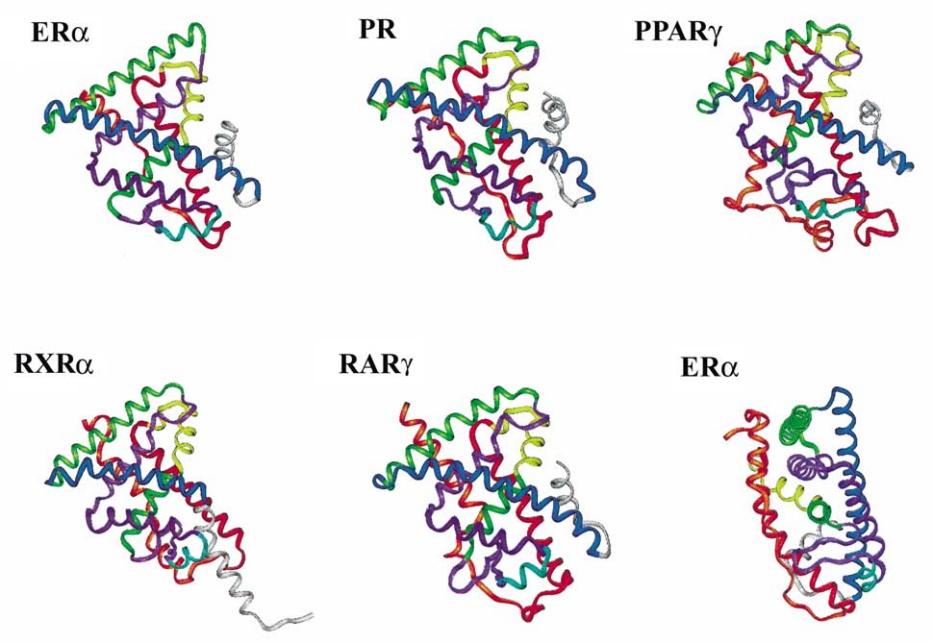

Fig. 1. Structural comparison of steroid hormone receptor crystal structures. Receptors were visualized using Insight II. Structural regions (helices, sheets, and loops) were colored based upon their spatial location in order to allow comparison between independent structures. Abbreviations are as follows with the corresponding Protein Data Bank accession numbers listed in parenthesis: ER-α — estrogen receptor α (1ERE); PR — progesterone receptor (1A28); PPAR-γ — peroxisome proliferator activated receptor γ (4PRG); RAR-γ — retinoic acid receptor γ (2LBD); RXR-α — retinoid X receptor α (1LBD). The ER-α structure at the lower right has been reoriented to allow visualization of the triple-layered α-helical sandwich structure common among all known steroid hormone receptor structures.

```
ER-α(304): nSLAL-SLTADQMVSALLDAEPPILYSEYDPTRPFSEASMMGLLTNLADR
ER-β(256): rELLLDALSPEQLVLTLLEAEPPHVLISR-PSAPFTEASMMMSLTKLADK

ER-α(353): ELVHMINWAKRVPGFVDLTLHDQVHLLE-CAWLEILMIGLVWRSMEHPGK
ER-β(305): ELVHMISWAKKIPGFVELSLFDQVRLLESC-WMEVLMMGLMWRSIDHPGK

ER-α(402): LLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEFVCLKSI
ER-β(354): LIFAPDLVLDRDEGKCVEGILEIFDMLLATTSRFRELKLQHKEYLCVKAM

ER-α(452): ILLNSGVYTFLSSTLKSLEEKDHIHRVLDKITDTLIHLMAKAGLTLQQQH
ER-β(404): ILLNSSMY-PLVTATQDADSSRKLAHLLNAVTDALVWVIAKSGISSQQQS

ER-α(502): ERLAQLLLILSHIRHMSNKGMEHLYSMKCKNVVPLYDLLLEMLDAHRlha
ER-β(453): MRLANLLMLLSHVRHASNKGMEHLLNMKCKNVVPVYDLLLEMLNAHVlrg
```

Fig. 2. Alignment of the ER-α and ER-β primary sequences. Ligand binding domain (LBD) sequences of ER-α and ER-β were aligned using the alignment: pairwise command of Insight II's homology module. Residues in bold type are identical between the two sequences. Boxes indicate residues corresponding to the ligand binding pocket of ER-α as determined by crystal structure analyses. While the entire range of residues shows 60% identity between sequences, regions within the ligand binding pocket are 72% identical. Sequence similarity (retention of residue properties such as charge, hydrophobicity, etc., despite alteration of identity) of binding pocket residues is >90% (residues in lowercase were not considered in sequence or structural analyses).

superimposition as was done for 17β-estradiol. No additional water molecules were considered during computations.

2.4. Homology model optimization

The Discover module of Insight II was used for energy minimization using the CVFF force field parameters for all ligands and protein models. Energy minimizations during initial model development and later during binding-energy calculations were divided into four stages in order to allow the structure to relax gradually into a final conformation, minimizing the possibility of dramatic structural alterations or distortions. In the first stage, only the hydrogen atoms of the model were allowed to reorient. Second, sidechains were allowed freedom of movement, followed by relaxation of backbone atoms while sidechains were held fixed. Finally, the entire structure was allowed full freedom of movement. During initial model development, the docked ligand molecule and water molecule were held fixed until the final minimization stage, at which time they were fully relaxed. Both molecules were allowed full freedom of movement during all stages other than hydrogen-atom reorientation of the binding-energy calculations. Each energy minimization was performed using steepest descent for 500 iterations or until the maximum derivative was <1.0 kcal/(mol Å), followed by conjugate gradient minimization until the maximum derivative was <0.01 kcal/(mol Å). A dielectric constant of 4.0 was used throughout all minimization steps.

2.5. Binding-energy calculations

Values of the binding energy, $\Delta E_{\text{binding}}$, were calculated as the difference between the potential energy of the ligand–receptor complex (E_{complex}) and the potential energy

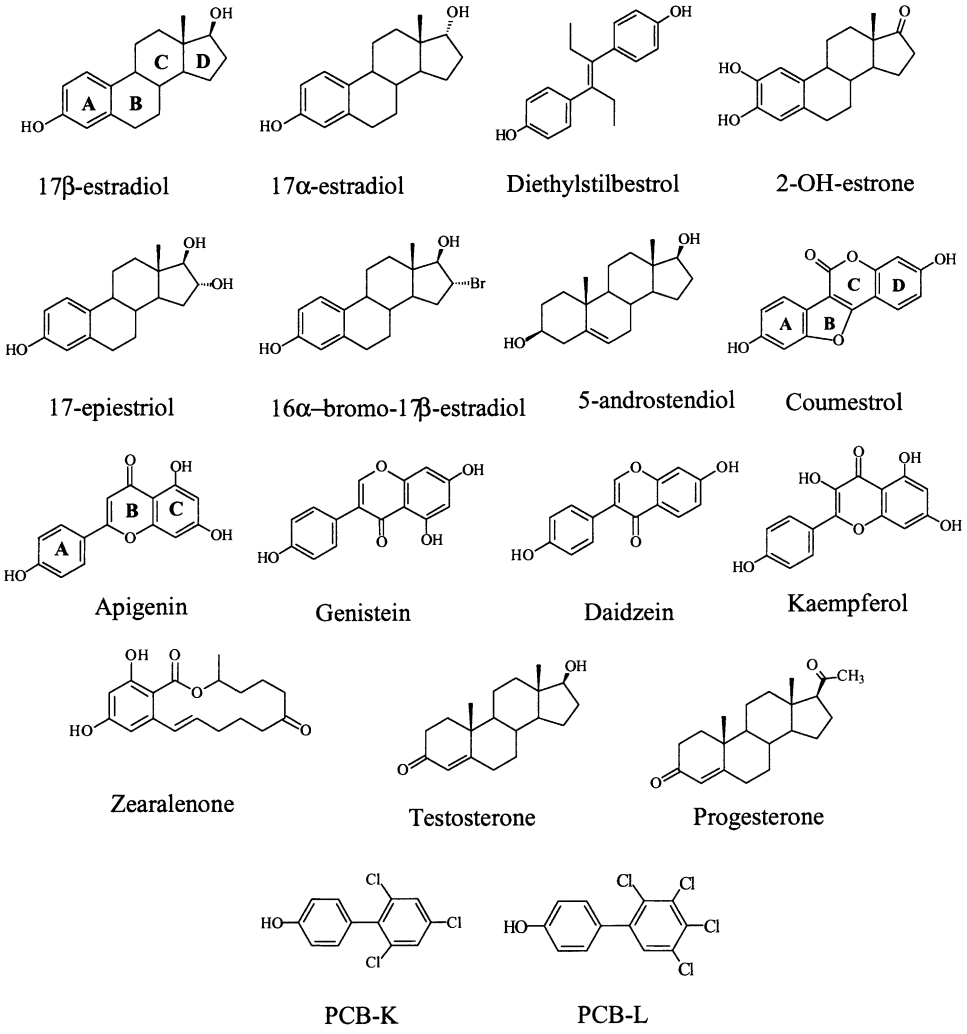

Fig. 3. Structures of compounds used in binding-energy calculation studies. Compounds were generally aligned as they are depicted here with the aromatic A-ring of 17β-estradiol serving as the primary pharmacophore used for alignment. Phytoestrogen compounds depict multiple possible alignments due to the presence of phenolic groups at either end of their structures. Various orientations were attempted and that which resulted in the lowest calculated binding-energy was retained for further analysis.

of the ligand (E_{ligand}) and receptor ($E_{receptor}$) separately.

$$\Delta E_{binding} = E_{complex} - (E_{ligand} + E_{receptor}) \quad (1)$$

Energy values represent the minimum energies resulting from the optimization procedure and parameters described above. The apo-receptor was first minimized by this procedure. Next, each ligand was minimized in isolation using the Discover module of Insight II. Finally, the minimized ligand was docked within the binding pocket of a minimized receptor using the ligand bound crystal structure of ER-α as a guide. Briefly, the ligand bound ER-α crystal structure was aligned with the minimized receptor using the structural alignment option of Insight II's homology module. Aromatic rings of the various ligands were aligned with the A-ring of bound 17β-estradiol, thus maintaining proper orientation of the critical hydrogen-bonding OH group between the two structures. Following the merger of the minimized receptor and the newly aligned compound, minimization proceeded stepwise as described above. This procedure ensures that each minimization was begun at the same starting point for the receptor conformation. In these studies, entropy terms and solvation energies were not considered during calculations. It is reasoned that differences in these terms can be tacitly assumed to cancel out for structurally related molecules, thus ensuring that the relative accuracy of the calculated binding energies ($\Delta E_{binding}$) is not compromised. Further studies would be required to address the effects of solvation/desolvation, entropy, and other factors. In light of the exceptional results obtained from the present approach, it can be reasonably expected that the net effect of these factors would be minimal.

2.6. Estrogenic compounds and experimental data

Activity data in terms of RBA, with respect to estradiol, for a variety of chemical compounds were obtained from Kuiper et al. [28]. RBA was defined by Kuiper et al. as the ratio of concentration of experimental compound required to reduce the specific radioligand binding by 50% (the ratio of estradiol IC_{50} to other ligands' IC_{50}), with the value for estradiol being arbitrarily set to 100 for normalization. Compounds utilized as a data set are illustrated in Fig. 3. Naming conventions were maintained as in Kuiper et al. [28] for the sake of consistency. Specifically, PCB-K refers to 2′,4′,6′-trichloro-4-biphenylol, and PCB-L refers to 2′,3′,4′,5′-tetrachloro-4-biphenylol. Structures were constructed using Sybyl 6.4 based upon structural information obtained from the Merck Index, 10th edition, and energy minimized using MMFF force field parameters. Docking and orientation of steroidal ligands with crystal structures was guided by the crystal structures of ER-α in complex with estradiol (1ERE, and 1A52) and ligands were oriented based upon apparent structural similarity. Orientation of some molecules, however, required the investigation of multiple possible configurations. By virtue of the presence of phenolic rings at either end of the molecule, some compounds (e.g. genistein) might adopt multiple possible orientations within the binding site: one in which the A-ring of the molecule is aligned with the A-ring of 17β-estradiol, and the others in which the C-ring is aligned with 17β-estradiol's A-ring. In such cases, where multiple orientations were possible, orientation was guided by the crystallographic findings of Pike et al. [29] in which the A-ring of genistein aligns with the A-ring of 17β-estradiol within the ligand binding site of ER-β. Ligands were collectively oriented as they are depicted in Fig. 3 with the A-ring of 17β-estradiol serving as the initial orienting pharmacophore for all molecules.

3. Results

3.1. Template selection

Crystal structures for various members of the steroid hormone receptor superfamily exist, however, no fully intact structure has yet been elucidated. Rather, the structures of the DNA binding domain (DBD) and ligand binding domain (LBD) have been solved independently for some members. The structural similarity among members of the steroid hormone receptor LBDs is well known [33]. To illustrate this, crystal structures for the LBDs of PR, PPAR-γ, ER-α, RXR-α, and RAR-γ were visualized in Insight II as α-carbon traces (Fig. 1). The various helices of the receptors are differentially colored based upon their spatial orientation in order to provide contrast between helices and to allow intermolecular comparisons. It is apparent that each of the receptors exhibits the same triple-layered, helical sandwich, and that the spatial orientation of the helices is highly similar despite the significant differences between their independent ligands: PR and ER-α bind steroidal molecules while RAR-γ, RXR-α, and PPAR-γ bind retinoic acid and fatty acid derivatives. This high level of structural conservation across disparate members of this family of receptors strongly supports their use as template molecules for the purposes of ER-β modeling.

It was found that ER-α and ER-β share >60% sequence identity over the ranges present within the various ER-α crystal structures (BLASTp score of 483 when comparing ER-β sequence with itself, and a score of 302 when comparing ER-α and ER-β sequences), while the remaining receptors are not >30% identical to ER-β. From a comparison of the residues present within the ER-α crystal structure to the known ER-β sequence, it was determined that the ER-β sequence would be limited to those residues ranging from glutamic acid 257 through valine 499 which correspond to residues present in the crystal structures of ER-α (serine 305 through arginine 548). Alignment of the ER-α and ER-β ligand binding domain sequences (Fig. 2) reveals a high level of sequence conservation between the two molecules. Additionally, the regions which correspond to the ligand

binding pocket (as determined from X-ray crystal structures) of ER-α are illustrated, and it can be seen that 26 of the 36 residues (72%) within this region are identical residues in ER-β. Based upon this high level of sequence conservation between ER-α and ER-β ligand binding domains, the obvious structural similarity that exists throughout the family of steroid receptors, and the similarity of ligand selectivity between ER-α and ER-β, we chose the ligand binding domain of ER-α as the template molecule for homology modeling of the ER-β ligand binding domain.

3.2. Homology modeling of the ER-β ligand binding domain

The first, and arguably the most critical, step in protein homology modeling is the appropriate alignment of template and experimental sequences. While the ER-α LBD and ER-β LBD sequences hold >60% sequence identity, crystal structures for the ER-α LBD contain gaps in their amino acid sequences due to either poor resolution or lack of order within particular domains [30,31,34]. This is most commonly seen in loop regions in which the conformation is not necessarily fixed, but may in fact be flexible. This deficiency of any single structure is remedied by the availability of five independent crystal structures, allowing overlapping and complementary completion of structural data. To accommodate this approach, the available ER-α LBD structures were first aligned according to structural similarity using the alignment: structure menu command of Insight II's homology module. From this, it can be seen that the independent structures retain a common protein folding pattern with the major helices of the protein occupying nearly identical space (RMSD = 0.49 Å). However, a few points must be taken into consideration. First, loops within the various structures do not have identical conformations. This would be expected as loops are considered to hold loose conformations and may not in fact retain any single conformation at all times. Loops within the various structures do, however, occupy similar regions of crystal space and are not directly involved in ligand binding, which allows more generalized modeling of these structural regions. Second, the terminal helix occurs in three distinctly different conformations among the five available ER-α structures. Most notably, the helix appears to orient itself in diametrically opposed conformations when the receptor is bound to an agonist (estradiol and diethylstilbestrol) versus an antagonist (4-OH-tamoxifen and raloxifene) [31,34]. This observation has been the subject of much speculation and may in fact play an important role in the process of ligand discrimination by the receptor; however, there is no information that suggests either is the preferred conformation or that either conformation results in an altered ligand binding affinity. In addition, one ER-α LBD structure (1A52) shows this helix in an extended conformation projecting away from the bulk of the globule. This anomaly is explained by the presence of a fortuitous disulfide linkage between independent chains during crystal formation, and is considered a non-realistic representation of the spatial positioning of this helix [30].

Having aligned the structures of the template molecules, sequences from the ER-α LBDs were aligned with that of the ER-β LBD by multiple sequence alignment using the homology module of Insight II. Using the default settings, gapping was at a minimum and occurred only in loop regions or at the surface of the molecule, with all gaps being distal to the ligand binding site. From this alignment, structurally conserved regions (i.e. helices, beta sheets, and loops retaining high levels of sequence similarity between sequences) were utilized to assign coordinates to the residues of the ER-β LBD. For the terminal helix, coordinates were assigned from the estradiol-bound ER-α structure (1ERE) which orients this helix in close contact with the protein globule. Determination of coordinates for loop regions was accomplished by random generation within Insight II as described in Section 2.

Since approximately 40% of the residues of ER-β are non-identical to their corresponding residues in ER-α, assigning of coordinates could result in an inappropriate spatial overlap between some amino acid sidechains. When the model was assessed for this possibility, only a single clash was found between lysine 300 and tyrosine 488, both of which are found in areas outside the ligand binding site of the molecule. To rectify these clashes, the sidechains of the offending amino acid were adjusted manually. In addition, the new receptor was assessed for appropriate dihedral angle conformations, particularly at the splice points where random loops were generated. Cases of inappropriate splice site formation were repaired by localized energy minimizations in which loops containing poorly formed angles were allowed to reorient.

Two additional molecules were included in this structure as follows: 17β-estradiol was included by superimposition of the putative ER-β model with a known estradiol-bound ER-α structure (1A52) and docking of the estradiol molecule into the model. This orientation was then verified by comparison of the estradiol docked ER-β model with another estradiol-bound ER-α crystal structure (1ERE). Crystal structures of ER-α have revealed the presence of a single water molecule within the ligand binding site which appears to play a crucial role in mediating various hydrogen bonds between the bound ligand and the receptor [30–32]. In the same manner that estradiol was inserted, the ER-β model and the ER-α structure (1A52) were superimposed and the appropriate water molecule was then identified and inserted into the model.

The resulting model, while containing all residues specific to ER-β and having been refined as described above, required energy minimization which would allow all atoms of the molecule to find appropriate conformations within the context of this structure. This process was divided into four independent phases and conducted using the Discover module of Insight II as described in Section 2.4.

Table 1
Observed vs. computationally predicted log(RBA) values for ER-α[a]

Compound	Calculated binding energy[a]	Predicted log(RBA)	Experimental log(RBA)	Residual
17β-Estradiol	0.000	1.501	2.000	0.499
17α-Estradiol	1.579	1.105	0.845	−0.260
Diethylstilbestrol	−3.132	2.287	2.373	0.085
17-Epiestriol	−2.367	2.095	1.462	−0.633
16α-Bromo-17β-estradiol	0.699	1.326	1.881	0.555
5-Androstenediol	6.903	−0.231	0.000	0.231
Coumestrol	−0.586	1.649	1.301	−0.348
Genistein	6.385	−0.101	0.602	0.730
Daidzein	8.073	−0.524	−1.000	−0.476
Kaempferol	7.662	−0.421	−1.000	−0.579
Zearalenone	3.491	0.626	0.850	0.224
Progesterone	22.480	−4.142	≤2.0	–
Testosterone	14.484	−2.135	≤2.0	–

[a] Values are given in kcal/mol relative to the binding energy calculated for 17β-estradiol.

3.3. Binding-energy calculations are highly correlated with experimental ligand binding affinities for ER-α

Calculation of binding energies has been used for the purpose of predicting ligand–receptor complex formations for a variety of systems [35]. In order to establish the validity of using this method in the present application, we utilized experimental data obtained for a group of compounds, illustrated in Fig. 3 and listed in Table 1, shown to bind ER-α [28]. From this group of compounds, a data set was developed consisting of steroidal molecules, diethylstilbestrol, polychlorinated biphenyls (PCBs), and phytoestrogens which are capable of interacting with ER-α to varying degrees. Compounds were modeled using Sybyl 6.4 (Tripos, Inc., St. Louis, MO) and their geometry optimized structures docked into the ER-α structure (3ERD) by the method of superimposition described above using Insight II (Molecular Sumulations, Inc., San Diego, CA). Orientation of the molecules was accomplished by comparison to the various ligand-bound ER-α crystal structures and construction from the pre-existing ligand within the binding site. Regarding those molecules for which multiple orientations appeared possible, crystallographic information revealing the orientation of genistein within the ligand binding pocket of ER-β was used to guide placement [29]. Binding-energy calculations of two molecules (specifically, 2-OH-estrone and quercitin) docked within the ER-α ligand binding domain resulted in an overestimation of these compounds' binding affinity. It is noteworthy that these compounds are unique among the compounds in this study in that they possess adjacent hydroxyl groups on the benzene ring. This common feature prompts speculation that effects unique to these compounds, and thus, not accounted for in the present calculations, may play a role. For example, 2-OH-estrone and quercitin may possess larger desolvation energy and entropy terms than the other compounds in this series. It was already noted above that binding energies calculated as defined in Eq. (1) assume cancellation of solvation/desolvation terms, thus, this possibility would explain the discrepancy between observed and predicted RBA values in these two cases. In this context, it is worth noting that these two compounds possess the ability to form an intramolecular hydrogen bond between the two adjacent OH groups (particularly within the ER hydrophobic binding cavity). As another possibility, Kuiper et al. [27] postulated that competitive ligand binding data for low-affinity ligands may reflect blockage of the ER entrance to the binding site or interaction with an alternative site which causes a change in the high-affinity ER site. This issue, although beyond the focus of the present study, is a subject for further scrutiny in our laboratory. For the sake of uniformity, 2-OH-estrone and quercitin were omitted from the subsequent analyses. The results of these calculations are listed in Table 1 and illustrated in Fig. 4. It is apparent from these results that the calculated values of $\Delta E_{\text{binding}}$ are highly correlated ($r^2 = 0.82$) with the experimentally obtained RBAs for the selected molecules. In order to further validate our approach, two compounds (testosterone and progesterone) which hold significant structural similarity to 17β-estradiol, but are incapable of binding ER-α at a detectable level were docked within the ligand binding domain and $\Delta E_{\text{binding}}$ calculated. The results, listed in Table 1, along with the larger validation set described above support the validity of this technique for the prediction of ligand binding affinities, and thus, these results serve to validate this methodology as applicable within our experimental system.

3.4. Binding-energy calculations derived from the ER-β homology model show a strong correlation with experimentally derived binding affinities

Having established the utility of binding-energy calculations as a computational measure of actual ER-α-ligand complex formation, we addressed the ability of the ER-β homology model to provide similar information. While it could be expected that a model based upon a known crystal structure is amenable to this method, utilization of a

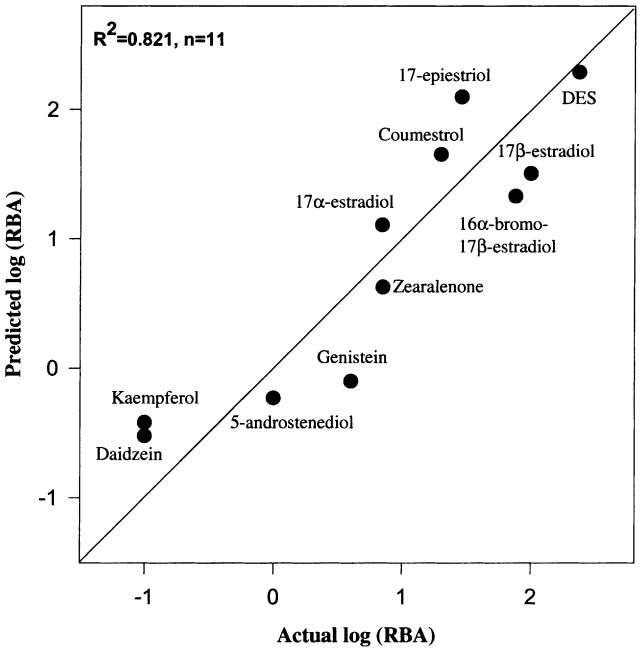

Fig. 4. Experimental vs. predicted ER-α RBA values. Compounds were docked into the ER-α crystal structure (1ERD) using Insight II, and binding energies calculated. Binding energies were normalized, with diethylstilbestrol serving as the reference. Regression analysis was performed relating binding energies and experimentally obtained log(RBA) values. The resulting regression equation was used to predict log(RBA) values for each molecule within the set, followed by graphing of experimental values vs. the predicted values ($r^2 = 0.82$).

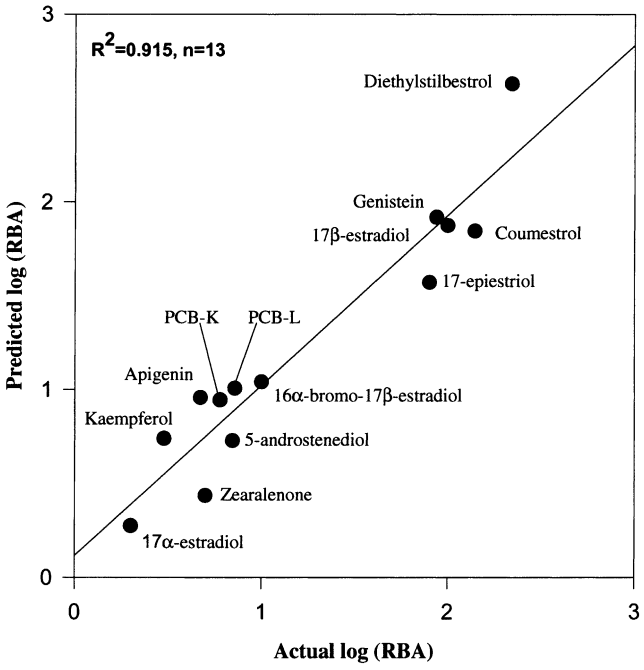

Fig. 5. Experimental vs. predicted ER-β RBA values (ER-β homology model). Compounds were docked into the homology modeled ER-β structure using Insight II, and binding energies calculated. Binding energies were normalized to estradiol and regression analysis performed relating binding energies and experimentally obtained log(RBA) values. The resulting regression equation was used to predict log(RBA) values for each molecule within the set The graph depicts experimental values vs. the predicted values ($r^2 = 0.92$).

homology-based model has not yet been attempted. We employed a similar training-set as that established for ER-α with the additional consideration that the RBA value be >1 (estradiol being set to 100) for any compound within the set This measure was taken to avoid potential problems associated with extremely low or undetectable RBA values falling within the error limits of the model itself. Ligand structures were docked into the ER-β model, after which energy minimizations were performed using the discover module of Insight II. Results are listed in Table 2 and illustrated in Fig. 5. A strong correlation ($r^2 = 0.92$) can be seen between the experimental RBA data and the calculated values of $\Delta E_{\text{binding}}$. In fact, the correlation found here for the ER-β homology model is higher than that found above when using the known crystal structure for ER-α ($r^2 = 0.82$).

Table 2
Observed vs. computationally predicted log(RBA) values for the ER-β homology model

Compound	Calculated binding energy[a]	Predicted log(RBA)	Experimental log(RBA)	Residual
17β-Estradiol	0.000	1.899	2.000	0.101
17α-Estradiol	20.830	0.339	0.301	−0.038
Diethylstilbestrol	−9.830	2.635	2.344	−0.291
17-Epiestriol	3.950	1.603	1.903	0.300
16α-Bromo-17β-estradiol	10.850	1.086	1.000	−0.086
5-Androstenediol	14.930	0.781	0.845	0.064
Apigenin	12.100	0.993	0.778	−0.215
Coumestrol	0.370	1.871	2.146	0.275
Genistein	−0.580	1.942	1.940	−0.002
Kaempferol	14.780	0.792	0.477	−0.315
Zearalenone	18.730	0.496	0.699	0.203
PCB-K	12.760	0.943	0.778	−0.165
PCB-L	11.150	1.064	0.857	−0.207

[a] Values are given in kcal/mol relative to the binding energy calculated for 17β-estradiol.

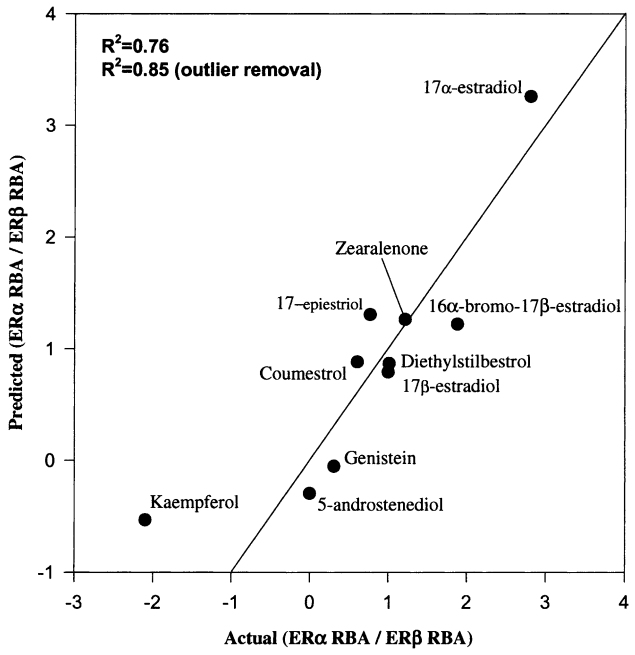

Fig. 6. Ratio of experimental RBA vs. predicted RBA. The ratios of experimental RBA values (ER-α/ER-β) and predicted RBA values (ER-α/ER-β) as determined by binding-energy studies conducted using the ER-α crystal structure and ER-β homology model were calculated for the compounds in common between the two data sets and plotted. The line of identity is illustrated. Prior to removal of the outlier Kaempferol, a regression coefficient (r^2) of 0.76 was obtained, whereas removal gave $r^2 = 0.86$.

3.5. Calculated binding affinities reflect receptor selectivity

In addition to providing a general estimate of ligand affinities for each independent receptor, our results also demonstrate the capability of recognizing receptor selectivity of molecules. Compounds such as 16α-bromo-17β-estradiol and genistein exhibit selectivity for ER-α and ER-β, respectively. This property is significant pharmacologically and physiologically when the differential tissue distribution of the receptors is considered [27]. To assess the degree to which our calculated binding energies reflected the selectivity of ligands for ER-α versus ER-β, we plotted the ratio of experimental RBA values for ER-α and ER-β RBA values versus the ratio of predicted binding affinities for ER-α and ER-β (Fig. 6). The resulting graph depicts a good correlation ($r^2 = 0.76$), indicating that calculated binding energies are capable of predicting and quantifying selectivity between these two receptor subtypes to a reasonable degree. Furthermore, when the single outlier Kaempferol, is removed, the correlation is significantly improved ($r^2 = 0.85$). It is interesting to note that those molecules which depict the greatest degree of receptor selectivity, such as 16α-bromo-17β-estradiol and genistein, which are selective for ER-α and ER-β, respectively, can be easily identified from the computational approach despite the level of error which is inherent to a model of this nature.

3.6. Comparisons with the known ER-β crystal structure

As stated previously, the experimental work described above was initiated and completed prior to the public availability of the ER-β crystal structure. During the final preparation of this manuscript, however, the ER-β crystal structure was elucidated (reference) and the coordinates made publicly available through RCSB-PDB (accession #1QKM and #1QKN). While the authors fully acknowledge that the elucidation of a protein crystal structure supercedes any previously existing predictive model of protein structure, this in no way diminishes the importance of the work described herein. Rather, the existence of a crystal structure for the modeled receptor provides an excellent opportunity to test not only the accuracy of the modeled structure itself, but also the experimental techniques for which the model was developed. Furthermore, considering the diversity and importance of the steroid hormone receptor superfamily in physiology, toxicology, and drug design, successful comparison of the ER-β homology model to the known crystal structure for this receptor would strongly support the use of these techniques, and suggest the likely application of additional techniques, for assessment and analysis of this family of receptors' structure and ligand binding capabilities.

To assess the structural accuracy of the ER-β homology model, the crystal structure and model were first superimposed by structural alignment using the homololgy module of Insight II (Fig. 7). The RMS deviation of the α-carbon backbone was found to be 2.5 Å. It should be noted that the crystal structure is in fact incomplete due to the lack of resolution in particular regions of the receptor, most notably various loop regions were not resolved. Also, as can be seen from the figure (Fig. 7), the terminal helix in the crystal structure and the homology model do not align exactly. This is not surprising considering that the crystal structure of ER-β contains the estrogen antagonist raloxifene, while the template employed for the present homology model was the ER-α crystal structure bound to the natural agonist 17-β estradiol. Various studies have shown that the positioning of this helix may in fact be the mechanism by which steroid receptors are capable of discrimination of ligand activity. Such helix repositioning has been shown with ER-α [31] as well as the retinoid receptor RAR-α [36]. When the RMSD between crystal structure and homology model was assessed in the absence of this helix, it was found to be <1.6 Å. The excellent agreement of α-carbon locations between the two structures in three-dimensional space strongly supports the utility of this homology model. While it is agreed this crystal structure takes precedence over the homology model, the results obtained in comparing the two structures is extremely encouraging regarding modeling of other members of the nuclear hormone receptor superfamily for which crystal structures are not yet publicly available.

As a final validation step, the crystal structure of ER-β was prepared for binding-energy calculations in the same manner as described above for the ER-α crystal structure

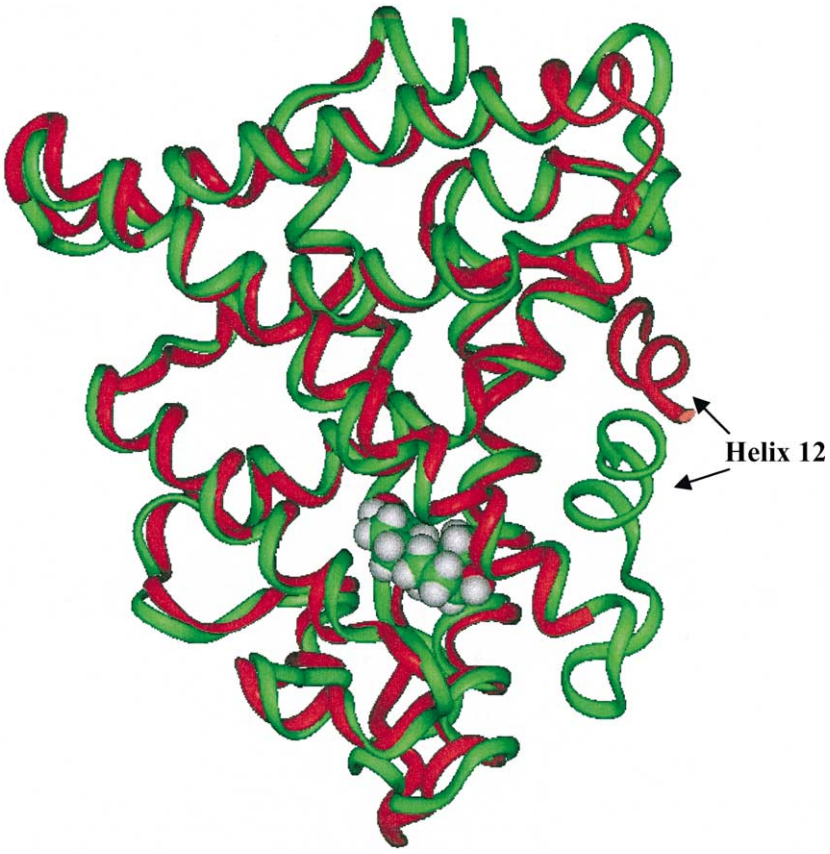

Fig. 7. Structural alignment of the ER-β homology model and the known ER-β crystal structure. The final homology model of ER-β (green) was aligned with the known crystal structure (red) of ER-β bound to raloxifene (RCSB-PDB accession #1QKN) using the alignment: structure command of Insight II's homology module. Structures were visualized as α-carbon traces. Estradiol was included and rendered as a space filling model in order to illustrate the ligand binding pocket of the receptors. The RMS deviation of backbone atoms was determined to be <1.6 Å. Helix 12 is labeled. See the text for a discussion of the orientation of this helix. Also note that the loop joining helix 11 and helix 12 is not present in the crystal structure due to lack of resolution in this domain of the protein.

and the ER-β homology model. Again, ligands were docked into the ligand binding pocket of the receptor based upon crystal structure data of the ER-α and ER-β holo-receptor complexes. The calculated values of $\Delta E_{\text{binding}}$ are highly correlated ($r^2 = 0.95$) with the experimentally determined RBAs for the same set of molecules used with the ER-β homology model (Table 3 and Fig. 8). It is not surprising that this level of correlation is higher than that seen using the

Table 3
Observed vs. computationally predicted log(RBA) values for the ER-β crystal structure

Compound	Calculated binding energy[a]	Predicted log(RBA)	Experimental log(RBA)	Residual
17β-Estradiol	0.000	0.600	0.301	0.299
17α-Estradiol	7.765	1.045	0.845	0.200
Diethylstilbestrol	1.843	0.706	0.778	−0.072
17-Epiestriol	3.047	0.775	1.000	−0.225
16α-Bromo-17β-estradiol	27.851	2.198	2.146	0.052
5-Androstenediol	27.384	2.172	2.340	−0.168
Apigenin	24.053	1.980	1.903	0.077
Coumestrol	25.100	2.041	2.000	0.041
Genistein	22.113	1.869	1.940	−0.071
Kaempferol	−2.181	0.475	0.477	−0.002
Zearalenone	1.802	0.703	0.778	−0.075
PCB-K	7.604	1.036	0.857	0.179
PCB-L	−1.785	0.497	0.699	−0.202

[a] Values are given in kcal/mol relative to the binding energy calculated for 17β-estradiol.

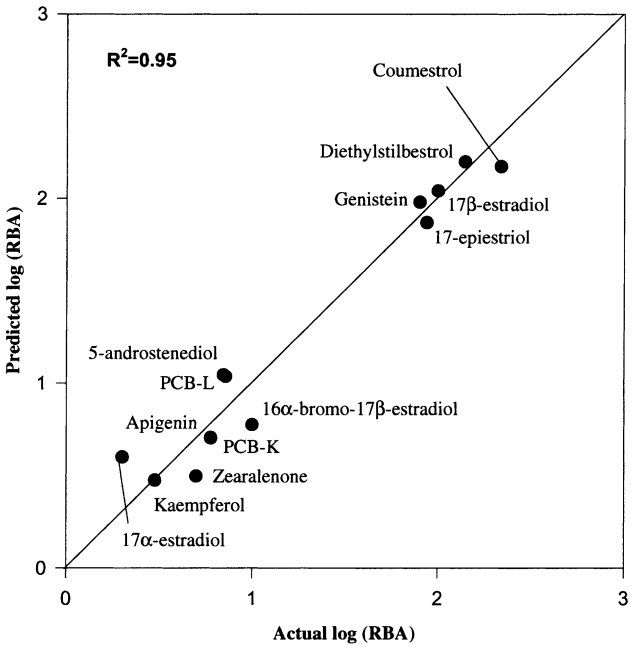

Fig. 8. Experimental vs. predicted ER-β RBA values (ER-β crystal structure). Compounds were docked into the ER-β crystal structure (1QKN) using Insight II, and binding energies calculated. Binding energies were normalized to estradiol and regression analysis performed relating binding energies and experimentally obtained log(RBA) values. The obtained regression equation was used to predict log(RBA) values for each molecule within the set. The graph depicts experimental values vs. the predicted values ($r^2 = 0.95$) (the data markers for apigenin and PCB-K, and 5-androstenediol and PCB-L were not able to be resolved for each pair due to their closeness on the graph).

homology model. Taken independently of other findings, this result further supports the use of calculated $\Delta E_{\text{binding}}$ as a predictive measure of the actual binding ability of various ligands to the receptor. However, when considered in combination with the results obtained for our homology model of ER-β, the result is significant and strongly supports the use of this modeling technique for this family of receptors.

4. Discussion

In an effort to provide a method which can be employed to effectively evaluate the estrogenic activity, as well as the ER-α versus ER-β selectivity, of both natural and synthetic compounds, we have developed a homology-based model of ER-β. Based upon the high level of sequence homology between ER-β and the well characterized ER-α, as well as the overall structural conservation seen throughout the family of steroid hormone receptors [33], ER-α was the obvious choice as the template molecule for homology modeling. ER-α also served as a model to establish the validity of ligand binding energies as a predictive measure for estrogenic RBA values. In the present study, various steroidal and phytoestrogen compounds were docked inside the ER-α structure, and binding energies calculated. Upon comparison with experimental values a strong correlation was observed ($r^2 = 0.82$), thus validating this approach for the system under study.

When the same approach was applied to the ER-β homology model, it was again found that a strong correlation existed between actual RBA values and calculated binding energies for the training-set compounds ($r^2 = 0.92$). In fact, the degree of correlation was higher for ER-β than for ER-α. The utility of this computational approach, and in particular of this ER-β model, is supported by the accurate evaluation of binding affinities for a number of structurally distinct compounds. This result is significant in that it illustrates the extensibility of this model, and further reinforces confidence in its accuracy. Furthermore, it is apparent that compounds which exhibit receptor selectivity can be accurately identified, providing a computational tool for separation of compounds exhibiting mixed activities.

Due to the high level of structural diversity of compounds found to have estrogenic activity (see Fig. 3), it is interesting to consider the orientation of molecules within the ER-α or ER-β ligand binding sites. Pharamcophoric analyses have identified the aromaticity of the estradiol A-ring as a significant contributor to binding capacity [37]. Additionally, the 3-hydroxylation has been implicated through pharmacophoric analyses as well as crystallographic analyses as critical in mediating hydrogen-bonding with amino acid sidechains of the ligand binding site [37]. It is important to point out that a number of the phytoestrogen compounds depict these pharmacophoric properties at either extreme of the molecule suggesting possible multiple orientations within the ligand binding site. In support of this hypothesis, when binding energies for genistein were calculated for either orientation, the values were remarkably close (117 and 122 kcal/mol), suggesting that such ligands may bind in either orientation within the receptor. Curiously, the orientation of genistein reported by Pike et al. [29] represents the higher calculated value, however, the degree of difference between the two $\Delta E_{\text{binding}}$ values suggests the possibility of >1 binding mode for compounds of this nature. Furthermore, the existence of two feasible binding modes would make a favorable entropic contribution to the free energy of binding.

The successful determination of the three-dimensional structure of ER-β LBD bound to the phytoestrogen genistein and the antagonist raloxifene has recently been reported [29], but the corresponding coordinate files were not yet publicly available during the initial phases of this work. Through the development of this homology model, we hoped to provide insight into the structure and function of this receptor and to stimulate further investigation by the scientific community. We are strongly encouraged by the successful comparison of our ER-β homology model with the elucidated crystal structure. In one respect, it was found that the structural prediction itself is sound in that the RMS deviation between the α-carbon backbones of the model and crystal structure was only 1.5 Å. Moreover, the correlation between the actual

RBA values and those predicted based on binding-energy calculations was only slightly lower for the ER-β homology model ($r^2 = 0.92$) than for the corresponding crystal structure ($r^2 = 0.95$). This further supports the validity of the present homology model as well as the binding-energy calculations. We recognize that additional work is needed to enhance this modeling procedure, thus allowing a greater level of predictability. However, the technique of homology modeling should assist both in risk assessment of endocrine disrupting compounds as well as in the discovery of estrogenic and other nuclear hormone receptor binding compounds for myriad therapeutic applications.

The present study has demonstrated that calculated ligand binding energies correlate very highly with experimentally determined RBA values of a wide range of estrogenic compounds for both ER-α and ER-β. Furthermore, we have also shown that this approach is applicable not only to a known crystal structure (ER-α), but also to a homology model (ER-β). This finding is significant in that it adds confidence in the reliability and utility of the homology model for ER-β presented here. Considering the high level of structural and sequence conservation within the steroid hormone receptor superfamily, we propose that the same approach holds promise for application to other members of this receptor group as well as to other receptor families for which such a high level of structural data exists. Furthermore, we feel that successful application of computer-based tools, such as those employed here and in additional studies within our laboratory [38–40], will ultimately reduce the time and cost, and improve the efficiency and reliability of high-throughput screening processes, thus enabling intelligent decisions to be made concerning the efficacy and safety of newly developed chemical compounds.

Acknowledgments

The authors wish to acknowledge useful discussions with Drs. Weida Tong, Roger Perkins, and Daniel M. Sheehan of the National Center for Toxicological Research (NCTR) and with Drs. Thomas Wiese and Thomas Bishop of Tulane University, School of Public Health and Tropical Medicine, New Orleans, Louisana. This work was supported by a grant from the US Environmental Protection Agency STARS (Science to Achieve Results) Program.

References

[1] A.C. Guyton, J.E. Hall, Textbook of Medical Physiology, Ninth Edition, Saunders, London, 1996, p. 1148.
[2] M. Sato, T.A. Grese, J.A. Dodge, H.U. Bryant, C.H. Turner, Emerging therapies for the prevention or treatment of post-menopausal osteoporosis, J. Med. Chem. 42 (1999) 1–24.
[3] K.S. Korach, P. Sarver, K. Chae, J.A. McLachlan, J.D. McKinney, Estrogen receptor-binding activity of polychlorinated hydroxybiphenyls: conformationally restricted structural probes, Mol. Pharmacol. 33 (1988) 120–126.
[4] L.E. Gray Jr., J. Ostby, J. Ferrell, G. Rehnberg, R. Linder, et al., A dose–response analysis of methoxychlor-induced alterations of reproductive development and function in the rat, Fundam. Appl. Toxicol. 12 (1989) 92–108.
[5] P.S. Guzelian, Comparative toxicology of chlordecone (Kepone) in humans and experimental animals, Ann. Rev. Pharmacol. Toxicol. 22 (1982) 89–113.
[6] A.M. Soto, H. Justicia, J.W. Wray, C. Sonnenschein, *p*-Nonyl-phenol: an estrogenic xenobiotic released from modified polystyrene, Environ. Health Perspect. 92 (1991) 167–173.
[7] S. Jobling, T. Reynolds, R. White, M.G. Parker, J.P. Sumpter, A variety of environmentally persistent chemicals, including some phthalate plasticizers, are weakly estrogenic, Environ. Health Perspect. 103 (1995) 582–587.
[8] G.M. Stancel, H.L. Boettger-Tong, C. Chiappetta, S.M. Hyder, J.L. Kirkland, et al., Toxicity of endogenous and environmental estrogens: what is the role of elemental interactions? Environ. Health Perspect. 103 (1995) 29–33.
[9] P.J. Reijnders, Reproductive failure in common seals feeding on fish from polluted coastal waters (published erratum appears in Nature 324 (6096) (1986) 418), Nature 324 (1986) 456–457.
[10] D.M. Fry, C.K. Toone, DDT-induced feminization of gull embryos, Science 213 (1981) 922–924.
[11] J.D. McKinney, C.L. Waller, Polychlorinated biphenyls as hormonally active structural analogues, Environ. Health Perspect. 102 (1994) 290–297.
[12] L.A. Jones, R.A. Hajek, Effects of estrogenic chemicals on development, Environ. Health Perspect. 103 (1995) 63–67.
[13] T. Colborn, F.S. vom Saal, A.M. Soto, Developmental effects of endocrine-disrupting chemicals in wildlife and humans (see comments), Environ. Health Perspect. 101 (1993) 378–384.
[14] J.A. Katzenellenbogen, The structural pervasiveness of estrogenic activity, Environ. Health Perspect. 103 (1995) 99–101.
[15] S.H. Safe, Environmental and dietary estrogens and human health: is there a problem? (see comments), Environ. Health Perspect. 103 (1995) 346–351.
[16] R.J. Kavlock, G.P. Daston, C. DeRosa, P. Fenner-Crisp, L.E. Gray, et al., Research needs for the risk assessment of health and environmental effects of endocrine disrupters: a report of the US EPA-sponsored workshop, Environ. Health Perspect. 104 (1996) 715–740.
[17] K.D. Setchell, S.J. Gosselin, M.B. Welsh, J.O. Johnston, W.F. Balistreri, et al., Dietary estrogens — a probable cause of infertility and liver disease in captive cheetahs, Gastroenterology 93 (1987) 225–233.
[18] A.S. Leopold, M. Erwin, J. Oh, B. Browning, Phytoestrogens: adverse effects on reproduction in California quail, Science 191 (1976) 98–100.
[19] A.L. Murkies, C. Lombard, B.J. Strauss, G. Wilcox, H.G. Burger, et al., Dietary flour supplementation decreases post-menopausal hot flushes: effect of soy and wheat, Maturitas 21 (1995) 189–195.
[20] W.H. Utian, Comparative trial of P1496, a new non-steroidal oestrogen analogue, Br. Med. J. 1 (1973) 579–581.
[21] J.J. Anderson, S.C. Garner, Phytoestrogens and bone, Baillieres Clin. Endocrinol. Metabol. 12 (1998) 543–557.
[22] D.P. Rose, M. Goldman, J.M. Connolly, L.E. Strong, High-fiber diet reduces serum estrogen concentrations in pre-menopausal women, Am. J. Clin. Nutr. 54 (1991) 520–525.
[23] H. Adlercreutz, Phytoestrogens: epidemiology and a possible role in cancer protection, Environ. Health Perspect. 103 (1995) 103–112.
[24] E. Enmark, J.A. Gustafsson, Oestrogen receptors — an overview, J. Int. Med. 246 (1999) 133–138.
[25] S. Green, P. Walter, V. Kumar, A. Krust, J.M. Bornert, et al., Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A, Nature 320 (1986) 134–139.
[26] G.G. Kuiper, E. Enmark, M. Pelto-Huikko, S. Nilsson, J.A. Gustafsson, Cloning of a novel receptor expressed in rat prostate and ovary, Proc. Natl. Acad. Sci. U.S.A. 93 (1996) 5925–5930.

[27] G.G. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Haggblad, et al., Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta, Endocrinology 138 (1997) 863–870.
[28] G.G. Kuiper, J.G. Lemmen, B. Carlsson, J.C. Corton, S.H. Safe, et al., Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta, Endocrinology 139 (1998) 4252–4263.
[29] A.C. Pike, A.M. Brzozowski, R.E. Hubbard, T. Bonn, A.G. Thorsell, et al., Structure of the ligand binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist, EMBO J. 18 (1999) 4608–4618.
[30] D.M. Tanenbaum, Y. Wang, S.P. Williams, P.B. Sigler, Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 5998–6003.
[31] A.M. Brzozowski, A.C. Pike, Z. Dauter, R.E. Hubbard, T. Bonn, et al., Molecular basis of agonism and antagonism in the oestrogen receptor, Nature 389 (1997) 753–758.
[32] S.P. Williams, P.B. Sigler, Atomic structure of progesterone complexed with its receptor, Nature 393 (1998) 392–396.
[33] J.-M. Wurtz, W. Bourguet, J.-P. Renaud, V. Vivat, P. Chambon, et al., A canonical structure for the ligand binding domain of nuclear receptors, Nat. Struct. Biol. 3 (1996) 87–94.
[34] A.K. Shiau, D. Barstad, P.M. Loria, L. Cheng, P.J. Kushner, et al., The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen, Cell 95 (1998) 927–937.
[35] A.C. Nair, S. Miertus, A. Tossi, D. Romeo, A computational study of the resistance of HIV-1 aspartic protease to the inhibitors ABT-538 and VX-478 and design of new analogues, Biochem. Biophys. Res. Commun. 242 (1998) 545–551.
[36] W. Bourguet, V. Vivat, J.M. Wurtz, P. Chambon, H. Gronemeyer, et al., Crystal structure of a heterodimeric complex of RAR and RXR ligand binding domains, Mol. Cell 5 (2000) 289–298.
[37] G.M. Anstead, K.E. Carlson, J.A. Katzenellenbogen, The estradiol pharmacophore: ligand structure-estrogen receptor binding affinity relationships and a model for the receptor binding site, Steroids 62 (1997) 268–303.
[38] W. Tong, R. Perkins, R. Strelitz, E.R. Collantes, S. Keenan, et al., Quantitative structure–activity relationships (QSARs) for estrogen binding to the estrogen receptor: predictions across species, Environ. Health Perspect. 105 (1997) 1116–1124.
[39] W. Tong, R. Perkins, L. Xing, W.J. Welsh, D.M. Sheehan, QSAR models for binding of estrogenic compounds to estrogen receptor alpha and beta subtypes, Endocrinology 138 (1997) 4022–4025.
[40] W. Tong, D.R. Lowis, R. Perkins, Y. Chen, W.J. Welsh, et al., Evaluation of quantitative structure–activity relationship methods for large-scale prediction of chemicals binding to the estrogen receptor, J. Chem. Inform. Comput. Sci. 38 (1998) 669–677.