Published in final edited form as:

Steroids. 2007 March ; 72(3): 247–260.

CoMSIA and Docking Study of Rhenium Based Estrogen Receptor Ligand Analogs

Peter Wolohan and David E. Reichert*

Mallinckrodt Institute of Radiology Washington University School of Medicine, St. Louis, MO 63110

Abstract

OPLS all atom force field parameters were developed in order to model a diverse set of novel rhenium based estrogen receptor ligands whose relative binding affinities (RBA) to the estrogen receptor alpha isoform (ER α) with respect to 17 β -Estradiol were available. The binding properties of these novel rhenium based organometallic complexes were studied with a combination of Comparative Molecular Similarity Indices Analysis (CoMSIA) and docking. A total of 29 estrogen receptor ligands consisting of 11 rhenium complexes and 18 organic ligands were docked inside the ligand-binding domain (LBD) of ER α utilizing the program Gold. The top ranked pose was used to construct CoMSIA models from a training set of 22 of the estrogen receptor ligands which were selected at random. In addition scoring functions from the docking runs and the polar volume (PV) were also studied to investigate their ability to predict RBA ER α . A partial least-squares analysis consisting of the CoMSIA steric, electrostatic and hydrophobic indices together with the polar volume proved sufficiently predictive having a correlation coefficient, r^2 , of 0.94 and a cross-validated correlation coefficient, q^2 , utilizing the leave one out method of 0.68. Analysis of the scoring functions from Gold showed particularly poor correlation to RBA ER α which did not improve when the rhenium complexes were extracted to leave the organic ligands. The combined CoMSIA and polar volume model ranked correctly the ligands in order of increasing RBA ER α , illustrating the utility of this method as a prescreening tool in the development of novel rhenium based estrogen receptor ligands.

Keywords

steroid; docking; estrogen receptor; rhenium; CoMSIA

Introduction

Estrogens are a family of naturally occurring steroid hormones which exert physiological effects in the growth, and maintenance of a wide variety of tissues. As a result of the large role estrogens play in the physiology of many species the estrogen receptors represent a viable and important pharmaceutical target. In particular the estrogen receptors are a target for pharmaceutical agents for hormone replacement in menopausal women, reproductive cancers such as prostate cancer, cancer of the uterus and breast cancer. These effects result from interactions with the estrogen receptor, of which two isoforms are known, ER α and ER β . As

*Address all correspondence to: David E. Reichert, Ph.D., Mallinckrodt Institute of Radiology, Washington University School of Medicine, 510 S. Kingshighway Blvd., Campus Box 8225, St. Louis, MO 63110, phone: (314) 362-8461, fax: (314) 362-9940, e-mail: reichertd@wustl.edu

*This work was funded by the National Institute of Biomedical Imaging and Bioengineering, EB00340.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

members of the nuclear hormone receptor superfamily, they share a common multidomain architecture consisting of a DNA binding domain, a ligand binding domain, and an activation domain.

Pharmaceuticals developed to date can be divided into three distinct categories. The first category acts solely as receptor agonists such as the estrogen receptors natural ligand 17 β -estradiol. The second category includes antagonists such as the compound ICI 164,384. The third category has the ability to act as both agonists and an antagonist depending on the particular isoform targeted. Katzenellenbogen and co-workers have developed several series of nonsteroidal compounds based on substituted furans, pyrazoles and tetrahydrochrysenes^{1–3} which have been shown to exhibit unprecedented estrogen subtype selectivity compared to the classical steroidal compounds. In particular the compound 4-propyl-1,3,5-triphenolpyrazole has been reported to have 400-fold affinity for the ER α isoform of the estrogen receptor.

The development of imaging agents for the *in vivo* assessment of ER receptor status as a method for improving breast cancer staging has a long history. Breast tumors expressing a high level of ER receptors (ER+) can often be treated through hormone therapy with anti-estrogens, while low levels of receptors (ER–) require a more invasive treatment course. Imaging modalities such as PET (positron emission tomography) and SPECT (single photon emission tomography) can be utilized to assess the ER receptor status through the administration of ER targeted compounds labeled with the appropriate radionuclides; such methods are attractive because they simultaneously and noninvasively provide information on all tumor sites unlike biopsy based sampling. As an example 16 α -[¹⁸F]fluoro-estradiol has been used to image primary and metastatic breast tumors and the images shown to have value in predicting the clinical effectiveness of tamoxifen therapy.⁴

Short-lived, cyclotron produced imaging agents have a limited availability due to the required proximity to a cyclotron facility. A more widely applicable choice would be to develop ER targeted agents labeled with the radionuclide ^{99m}Tc; the most widely utilized radionuclide for diagnostic imaging in nuclear medicine. This desirability is a result of several factors, its high specific activity and availability through the ⁹⁹Mo generator, its convenient 6 h half-life, and its emission of a 140 keV γ -ray with 89% abundance. Rhenium, a nonradioactive congener of technetium has been commonly utilized in the development and characterization of technetium containing imaging agents; in addition it has two radionuclides that have been found effective in radiotherapy, ¹⁸⁶Re ($t_{1/2}$ = 3.71 d) and ¹⁸⁸Re ($t_{1/2}$ = 16.8 h).

Several research groups have been actively involved in developing such agents with the Katzenellenbogen group in particular reporting many compounds.^{5–9} In general most attempts have focused on the conjugation of a Tc/Re chelate group onto a steroidal scaffold, although several non-steroidal agents have been reported as well.^{8,10–12} Several examples of such compounds are shown in Figure 1. Many of these large lipophilic complexes suffered from low receptor binding affinity and/or high non-specific binding, and often suffered from low complex stability; thus proving impractical for *in vivo* imaging.

As part of our research in the development and application of tools for the structure based design of imaging agents, we have focused on the prediction of estrogen receptor binding selectivity^{13,14} and on the development of force fields capable of accurately modeling metal based imaging agents.^{15,16} In this study we have combined these two topics by combining the development of a force field capable of modeling a wide variety of Re based estrogen ligands, and the use of molecular docking with full ligand flexibility to study their interactions with the estrogen receptor. Molecular docking studies have become an important method in structure based design, but have not been widely used with metal complexes. Many docking

programs such as Gold, can utilize metal centers but this is generally a metal found in a co-factor or active site catalytic group. We believe that this work is the first example of a docking study where the metal center is incorporated into the ligand being docked, and is treated in the same fashion as the organic portions of the molecule. The successful application of molecular docking as a screening tool for possible metal based imaging agents should lead to a more rapid development of such agents, while analysis of the docked poses can lead to insights into important interactions between the ligand and receptor.

Experimental

Rhenium ligand complexes were modeled as structures using the OPLS-AA/L all atom force field as implemented in the TINKER¹⁷ molecular modeling package running on an SGI Octane2 workstation. The implementation of the OPLS-AA/L force field in TINKER is the full all atom OPLS force field developed by Jorgensen¹⁸ with modified torsional parameters for proteins as described by Kaminsky et al.¹⁹ X-ray structures representative of the classes of novel rhenium based estrogen receptor ligands (Figure 1)^{20–22} were taken from the Cambridge Structural Database²³ (CSD) and manipulated with the program MOLDEN,²⁴ which provides a convenient interface between the CSD and TINKER. The reference data for the rhenium parameters involved a total of eight rhenium containing crystal structures, shown in Figure 2. An additional eight rhenium containing crystal structures, shown in Figure 3, were used as a blind validation set.

Our selection of crystal structures was guided somewhat by the availability of experimental data (i.e. associated relative binding affinities for ER α) to model using our new parameters.^{20–22} As a result we selected structures which contained the tricarbonyl η^5 cyclopentadienyl chelate, the SSS tridentate 3+1 oxorhenium(V) chelate and rhenium analogues of estradiol, representative of the classes of rhenium based estrogen receptor ligands (Figure 1). The reference data chosen consists of several tricarbonyl η^5 cyclopentadienyl rhenium complexes (refcodes MICSAS, ACYPRE, COCPRE, TEBPUL) together with two oxorhenium(V) complexes (refcodes AMONOF and AMOPIB) and two analogues of estradiol (refcodes AMORAV and ZAZNIX).

From MOLDEN a TINKER input file was generated for each of the rhenium complexes. At this point a series of new OPLS-AA/L force field atom types were defined such as a five- and eight-coordinate rhenium atoms; eight-coordinate representing the η^5 coordination of a five-membered ring, together with definitions for coordinating organic ligands. For example looking at tricarbonyl-(η^5 -cyclopentadienyl)-rhenium(I) [COCPRE] an eight-coordinate Re atom type was defined, together with a new cyclopentadienyl carbon Csp^2 , a new carbonyl Csp atom type and finally a new carbonyl Osp atom type. A listing of the newly defined OPLS-AA/L atom types with their associated bond stretching force field parameters are shown in Table 2.

These force field parameters were optimized using the automated parameter development program FFGeneAtor which utilizes a GA. Following the work of Strassner et al.²⁵ a steady state GA using a roulette selector together with a crossover rate of 0.9 and a mutation rate of 0.1 was used for each run. Within the GA the force field parameters were represented by chromosomes of real numbers. Within a population each individual chromosome is constructed of genes that correspond to the missing parameters i.e. bond, angle, torsion and out of plane bending parameters. Initial values for these parameters were taken from an average of the corresponding experimental value following structural analysis of the reference structures from the CSD. Bond lengths were allowed to optimize within a range of ± 0.5 Å of this average while angles were allowed to optimize within a range of $\pm 30^\circ$. Associated force constants were constrained to a range of 0–600 kcal/mol Å² and 0–80 kcal/mol respectively. Coding of the

torsional parameters is more arbitrary because of the sheer number of these missing parameters hence the need to impose flexible boundaries for these terms in order to allow the GA to optimize fully. Torsional angles were allowed to optimize within a range of $(-180 - (+)180^\circ$ and their associated constants were constrained to a range of ± 20 kcal/mol. Only force field bond lengths, angles and dihedrals involving the newly defined atom types were optimized, the original OPLS-AA/L parameters were not altered.

Initial guesses for the coded parameters from the GA were fed into the force field making it now possible to minimize each starting X-ray structure using TINKER's optimize program. Once minimized TINKER's superpose routine was used to calculate the root-mean-squared deviation (rmsd), most commonly in massweighted coordinates, between the two structures. The rmsd is then used to generate a fitness value for each newly developed parameter set which in turn guides the GA. The cycle is repeated for a number of GA cycles which we set to 750 generations. Electrostatic point charges were added to the final force field parameters following the procedure of Bayly et al., using the programs RESP and Jaguar5.0.^{26,27} The input for RESP is the electrostatic potential calculated from a full quantum mechanical optimization of a crystal structure. Crystal structures representative of each of the new force field atom types had their geometries optimized at the density functional level of theory (B3LYP/LACVP*+) using Jaguar.

The end result is a set of newly defined force field parameters whose fitness to both the reference set of structures and a blind validation set of structures can be evaluated based on the average rmsd the goal being to produce a set of robust parameters which accurately reproduce the X-ray structure of a diverse set of related structures, an average rmsd below 0.400 being indicative of accuracy.

Parameter Testing

Eight crystal structures were chosen as a validation set to test the developed rhenium OPLS-AA/L parameters (Figure 3). Again structures in this validation set were selected to be representative of the tricarbonyl η^5 cyclopentadienyl rhenium(I) and the SSS tridentate 3+1 oxorhenium(V) chelates. However we were unable to include direct analogues of estradiol given their limited presence in the CSD. These independent structures were minimized with the developed parameter set and compared to the X-ray crystal structure as a test of accuracy. It is important to note that the structures in the validation sets serve as an independent test of the parameters as they were not utilized in developing the parameters.

Ligand Docking

Preparation of Protein Structure

The crystal structure of ER α in complexation with 17 β -estradiol (1ERE) was extracted from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB-PDB).^{28,29} The structure was read in and manipulated with the program Maestro.³⁰ Missing atoms and residues were added in order to complete the protein chain. Hydrogen atoms were added and then minimized using the OPLS-AA force field and the corresponding partial-charge description while the rest of the structure was held fixed until the maximum derivative was <0.01 kcal/(mol $\text{\AA}$) including a single conserved water molecule which is present in all estrogen protein-ligand complexes and is considered important for binding.^{28,31} The residues which were not resolved in the crystal structure but complete the protein chain were then minimized until the maximum derivative was <0.01 kcal/(mol $\text{\AA}$) while all other atoms were fixed.

Preparation of Ligand Structures

Table 1 lists the experimental binding affinity of each of the ligands relative to that of the natural estrogen 17 β -estradiol which is arbitrarily set at 100 together with a description of their structure. A higher RBA value is indicative of better binding to ER α . Compounds 3a (refcode AMORAV) and 17 (refcode ZAZNIX) were found in the CSD, while compounds 2a and 19 were constructed from these structures respectively. While 17 β -estradiol, extracted from the crystal structure 1ERE, was used as the template from which 11 β -methoxy estradiol (11 β OMeEST) and 11 β -chloromethyl estradiol (11 β MeClEST) were constructed. The A-ring mimic from the crystal structure of (η^5 -tetrakis(Pentafluorophenyl)cyclopentadienyl)-tricarbonyl-rhenium(I) [refcode MICSAS] was used to construct all of the other ligands in Table 1 after first aligning the A-ring mimic with that of 17 β -estradiol extracted from (1ERE). All of the ligands were then minimized using the newly developed OPLS-AA/L force field parameters and the program TINKER until the maximum derivative was <0.001 kcal/(mol Å). The lowest energy conformation of each ligand was then located by using a basin hopping method implemented in the program SCAN which is part of the TINKER suite of programs. In each case all rotatable bonds were selected and the lowest energy conformation was located.

Docking the ligands

In order to perform fully flexible docking studies, it was once again necessary to develop new atom type definitions compatible with Gold. In addition, Gold requires torsional force constants to guide the conformational changes which must occur to dock any complex structure inside the ligand-binding domain of a protein. As a result all of the torsional angles developed for the OPLS-AA/L force field were incorporated into the Gold parameter file. Now with Gold primed to recognize the rhenium compounds and alter their conformational space all of the ligands were docked in the ligand-binding domain (LBD) of both ER α . To do this our previous approach^{13,14} was again utilized and a 16Å cavity was defined around the carbon atom of the terminal methyl group of residue MET421 in the preprocessed crystal structure of ER α in complexation with 17 β -estradiol. Having produced viable poses of the ligands bound to ER α with Gold, the top ranked structures were extracted together with their associated scoring functions for further analysis.

CoMSIA models

CoMSIA models were constructed in the program Sybyl 7.0 using ligands in their top ranked pose after docking in ER α , no further alignment of the ligands was performed.³² A test set consisting of 22 of the 29 ligands was extracted at random leaving a blind test set of 7 compounds. Regarding post alignment of the ligands, models constructed from alignments of the A-ring mimic of 17 β -estradiol proved to be less predictive. Reading the docked poses into Sybyl 7.0 once again required the definition of a new set of compatible atom types. For example when reading in the structure of compound 17 (Table 1) refcode ZAZNIX new Sybyl atom types, similar to the TINKER atom types, were developed to represent an eight-coordinate Re (I) atom type, together with a new cyclopentadienyl carbon Csp², a new carbonyl Csp atom type and finally a new carbonyl Osp atom type. Once the compounds were collected the electrostatic, steric and hydrophobic indices of all of the ligands were calculated by the CoMSIA technique. In addition to the standard indices calculated in CoMSIA, the polar volume and several of the Gold scoring functions, were added to the correlation models to investigate whether their inclusion would aid the predictive nature of our models. These functions were H_Ext and VDW_Ext which refer to the external hydrogen bonding energy function, the external Van der Waals energy function together with the overall Gold fitness function. Table 4 summarizes the results from a partial least squares (PLS) analysis, utilizing the leave-one-out method, of our CoMSIA models for all 22 ligands in the training set. The value q² is a

measure of the external predictive nature of the CoMSIA model, a q^2 greater than 0.50 said to represent a predictive model of use to the drug design process.

Results

Figure 1 illustrates the chemical structures of the different classes of ER ligands used in this study with their pharmacophoric elements highlighted. Obviously the A-ring mimic of estradiol is an important functionality retained in all of the structures. From the crystal structure 1ERE, the natural estrogen 17 β -estradiol interacts with ER α via a hydrogen-bonding network formed from the hydroxy group of the A-ring interacting with ARG394, GLU353 and a single water molecule. The hydroxy group of the D-ring forms a hydrogen bond with HIS524. These hydrogen-bonding interactions form the basis of the favorable binding interaction of 17 β -estradiol with ER α and thus are the core elements for a pharmacophoric model of the ER α binding pocket. In previous studies we have illustrated the docking of the novel non-steroidal ER ligands triphenol pyrazoles and furans.^{13,14} In particular the triphenol pyrazoles have drawn attention because of their reported specificity for the ER α isoform.^{33–35} It is important to note that the origin of this enhanced specificity for ER α comes from the ligands poor binding to ER β which we attributed to a flipped docking pose in ER β .^{13,14} Understanding the origin of the biological discrimination of these novel non-steroidal ER ligands in ER β is of great importance since it may lead to the development of more potent and selective ER ligands.

Obviously the triphenol substituted five membered ring scaffold utilized by Katzenellenbogen *et.al.*, has proven fruitful in the design of ER specific ligands and as expected via the synthesis of mixed Rhenium and Technetium complexes of this scaffold a new set of ER radiopharmaceuticals was developed.²¹ The aryl cyclopentadienyl tricarbonyl rhenium complexes and their organic analogues, aryl cyclopentadienes, form the basis of our study. In order to compare our Re results to previously studied organic ligands, we included several steroidal analogues of estradiol. Table 1 lists the experimental binding affinity of each of the ligands used in this study relative to that of the natural estrogen 17 β -estradiol which is arbitrarily set at 100. A higher RBA value is indicative of better binding. To model these rhenium complexes a new set of force field parameters had to be developed hence given our previous work modeling the ER with the OPLS-AA force field it seemed obvious to develop new rhenium-ligand parameters compatible with this force field.

The unit weighted root-mean-squared deviation (rmsd[U]) of each of the structures in the rhenium(I)/oxorhenium(V) reference (Figure 2) and validation set (Figure 3) are shown in Table 3. As can be seen in Table 3, the GA developed OPLS-AA/L parameters were able to reproduce the X-ray structures with good accuracy; the reference set exhibited an average rmsd of 0.267. A similar level of accuracy was maintained (rmsd[U] = 0.240) when the parameters were used to reproduce the X-ray structures of the eight compounds in the blind validation set (Figure 3). Given that one of the main goals of our work was to model the aryl cyclopentadienyl tricarbonyl rhenium complexes further analysis of how accurately the new parameters modeled the η^5 -(Pentafluorophenyl)cyclopentadienyl-tricarbonyl-rhenium(I) complexes [refcodes MICSAS, ACEFAP, ACEFET and QEWWIX] is warranted. Our new OPLS-AA/L parameters modeled this subset of rhenium complexes quite accurately with an average unit weighted root-mean-squared deviation (rmsd[U]) of 0.334. Given the conformational freedom of the phenyl rings around the cyclopentadienyl core this result is evidence of the accuracy of our parameters. Furthermore within the same framework the accuracy of the modeled rhenium based estradiol analogues [refcodes AMORAV and ZAZNIX], exhibiting a rmsd[U] of 0.293 and 0.330 respectively, is a clear example of the robustness of these new parameters to model such a diverse set of compounds with good accuracy.

With the force field parameters developed and tested the 29 compounds in our study were constructed and modeled as discussed (Table 1). At this point the program Gold was used to dock all of the ligands in the ligand-binding domain (LBD) of ER α using the corrected Gold parameter file containing the compatible rhenium atom types and torsional force constants. The result from a Gold run is a series of viable conformations of the ligand docked inside the LBD of the target protein together with an associated fitness function and other measures of the corresponding protein-ligand interaction energy. As a validation of the accuracy of the docking program Gold and approach used in this study the root-mean squared (rms) deviation of the crystal structure of 17 β -estradiol from IERE was compared against the most favorably ranked conformation of 17 β -estradiol docked with Gold. The rms deviation between the experimental docked conformation and the calculated docked conformation for 17 β -estradiol in ER α was 0.26. Given the low rms deviation between the experimental structure and the calculated docked structure it is reasonable to expect that the program would exhibit a similar accuracy with the other ligands utilized in the study. Indeed Gold was able to locate viable docking conformations, i.e. inside the LBD, of all of the ligands presented to it and in each case the expected A-ring mimic was located in a similar position to that of 17 β -estradiol.

Figure 4a illustrates the top ranked docking pose of 4'-hydroxyphenyl) cyclopentadienyltricarboxyl rhenium(I), 12a, the least active and 3-Ethyl-1,2,3-tri-(4'-hydroxyphenyl)cyclopentadienyltricarboxyl rhenium(I), 12h, the most active of the aryl cyclopentadienyl tricarboxyl rhenium complexes (12a - 12h) respectively. Figure 4b is a closer look at the docked poses produced by Gold illustrating the hydrogen bonding network observed utilizing the program LIGPLOT. Obviously 12a lacks the phenol group in either the R1 or R2 position capable of forming the hydrogen bonding network with HIS524 of the protein but it is also clearly illustrates that the number of hydrophobic interactions is greatly reduced hence they too could be the origin of it's significantly lower potency in ER α . However the lack of ER α potency in these ligands is most probably due to their inability to form both an A-ring mimic hydrogen bonding network with ARG394, GLU353 and D-ring mimic with HIS524 because of their structure. Indeed consider the observed RBA ER α of 12d and 12h 2.50 and 23 respectively. In their docked pose they both retain the phenol groups in the A-ring thus both are able to form the A-ring mimic hydrogen bonding network with ARG394, GLU353 and the conserved water molecule (Figure 4b for 12h) however in 12d the ethyl substituent sits in the pocket where the D-ring mimic of 12h sits and as a result 12d is unable to form the secondary hydrogen bonding network with HIS524 despite the A-ring mimics and tricarboxyl rhenium moiety being virtually overlapped. In addition the series of furans was chosen to enhance our dataset but also to illustrate the importance of the alkyl substituent in the R3 position and it's favorable hydrophobic interactions in the α -face of the LBD specifically with residues PHE404, MET388, PHE425, ILE424 and LEU428 which line the surface of the cavity formed in the α -face of the LBD of ER α . The ethyl group of 15c being the most favorable exhibiting a RBA ER α of 140 while the normal butyl group is least favorable no doubt because of it protruding too deep into this cavity. Indeed it is in this cavity that the tricarboxyl rhenium moiety is docked in the most active complexes of the aryl cyclopentadienyl tricarboxyl rhenium complexes (12d - 12h) while the less active complexes dock in a flipped pose with the tricarboxyl rhenium moiety interacting with the β -face of the LBD (12a - 12c).

An intriguing result comes when comparing the docked poses again of 12h the most potent aryl cyclopentadienyl tricarboxyl rhenium complex and 10h the organic analogue. Obviously from Table 1 the rhenium complex exhibits a 4-fold affinity for ER α over that of its organic analogue. When overlapped in their top ranked docking pose the two structures almost overlap with the A-ring mimic phenols equidistant from the key residues ARG394, GLU353. Indeed the D-ring mimic of 10h is slightly closer to HIS524 ($O_{10h}-N_{HIS524} = 2.35 \text{ \AA}$) then both 17 β -estradiol ($O_{EST}-N_{HIS524} = 2.83 \text{ \AA}$) and 12h ($O_{12h}-N_{HIS524} = 3.69 \text{ \AA}$). One might expect a shorter hydrogen bond to be more favorable for binding and hence lead to a higher affinity.

Yet despite this and with the ethyl substituents in 10h and 12h also almost overlapping it is 12h that exhibits the higher affinity for ER α with an RBA of 23. As a result based on our docking observations this enhanced potency for ER α must be due to the interaction of the tricarbonyl rhenium moiety of 12h which sits in a similar position as the important R3 substituents in the furans and pyrazoles in the α -face of the LBD of ER α . This suggests that the tricarbonyl rhenium moiety in 12h is well tolerated in this docked pose and that the interactions in the α -face of the LBD specifically with residues PHE404, MET388, PHE425, ILE424 and LEU428 lead to favorable binding.

Docking indicates how a ligand might bind inside a protein binding pocket but it will not tell us how active that same compound might be. This is despite the continued effort directed to the development of accurate scoring functions. Indeed as previously reported we found little correlation between a wide range of scoring functions and estrogen receptor ligands^{13,14} Once again we saw no correlation between the RBA ER α and any of the Gold scoring functions from the docking runs each PLS model exhibiting a negative cross-validated correlation coefficient (Table 4). Even after extracting the rhenium based compounds from the training set of 22 compounds the PLS models continued to exhibit no relationship between the RBA ER α and any of the Gold scoring functions, indeed the predictive power of the cross-validated correlation coefficient deteriorated (Table 4: *organics only*). Quantitative Structure Activity Relationships on the other hand provide a reasonable method by which a robust relationship can be developed to describe and predict the activity of such compounds. Previously we utilized the Comparative Molecular Field Analysis CoMFA method to develop PLS models to accurately predict the ER α and ER β of have a series of ER ligands including the triphenyl pyrazoles and furans here we have utilized the Comparative Molecular Similarity Indices Analysis (CoMSIA) method. It was our intention to utilize CoMFA to complement our previous ER work however CoMFA was unable to properly describe the VDW fields of the Re centers. CoMSIA, a newer technique, was able to describe the indexes of the Re complexes correctly. From the CoMSIA results shown in Tables 4 it can be seen that a strong correlation can be found between the calculated CoMSIA molecular steric and electrostatic indices and the observed relative binding affinity, ER α : $r^2 = 0.96$, $q^2 = 0.56$, SEE = 0.20. Inclusion of the hydrophobic similarity index did not improve the PLS model however inclusion of the calculated polar volume enhanced the external predictive ability of the PLS model significantly; ER α : $r^2 = 0.94$, $q^2 = 0.68$, SEE = 0.24. Given that the initial CoMSIA molecular steric and electrostatic indices model is dominated by the electrostatic index which contributes 73% to the model it should not be surprising to see that the polar volume was found to be predictive.

Figure 5 illustrates the standard deviation of the calculated three-dimensional molecular fields from our most predictive ER α CoMSIA model (Table 4). Within these molecular fields the ligand 3-Ethyl-1,2,3-tri-(4'-hydroxyphenyl)cyclopentadienyltricarbonyl rhenium (I) (12h) has been superimposed in it's top ranked Gold pose for easy comparison. Contours of the steric map are shown in transparent yellow and green, those of the electrostatic map are shown in solid red and blue, while those of the hydrophobic map are shown in cyan and purple. Increased biological activity is correlated with: more bulk near green; less bulk near yellow; more positive charge near blue, more negative charge near red, more hydrophobicity near cyan and less hydrophobicity near purple.

As a result these fields could be manipulated in order to increase the activity of the rhenium complexes using 12h as the template upon which to improve. However given the diversity of the data set used in constructing this PLS model utilizing it in the familiar ligand based design role might prove particularly challenging. For ligand based design these methods lend themselves to data sets consisting of closely related backbones. For example for the estrogen receptor a series of derivatives of the antagonist Raloxifene might be used and a novel ligand might be designed by making discrete modifications to the 2-Arylbenzothiophene core. Here

we are mixing not only the core five membered ring but we are also including a significantly diverse functional group in terms of the rhenium chelate and comparing this to the standard steroid core of estradiol. In fact it should be noted that we were more interested in developing CoMSIA models capable of predicting the activity of a diverse set of ER ligands than ligand based design from a single backbone. This can be seen from our analysis of the test set which consisted of the seven compounds, 17, 15d, 4i, 2a, 10e, 12d and 12a which were selected at random (Table 5). These were ranked in order of increasing RBA and their calculated RBA's were predicted using our most predictive CoMSIA PLS model, i.e. the CoMSIA model utilizing the steric, electrostatic and hydrophobic indices together with the calculated polar volume, for comparison. Our goal was to illustrate how our models could be used to prescreen new rhenium based ligands for their affinity to ER α .

Discussion

From Table 5 it can be seen that as a first pass our model was able to accurately predict the most active compound from the test set illustrating the utility of this method as a prescreening tool in the development of novel rhenium based estrogen receptor ligands. In fact our model was able to predict the ranking of the three most active compounds correctly. Given that our models consist of rhenium based ligands and as a result are to our knowledge the first such QSAR model which contains organometallics which have been correctly modeled, to predict the ranking of these compounds in the blind test set was particularly encouraging. Upon further analysis it became clear that the compound (3-Thiapentane-1,5-diotholato)-(17 α -(5-thiolato-1-pentyn-1-yl)-estra-1,3,5(10)trien-3,17 α -diol)-oxo-rhenium(V), 3a, was an outlier given that it's experimental RBA is 14.10 while it's predicted RBA utilizing our CoMSIA model was 0.33 giving it a rank of seventh, the least favorable compound of the seven in the test set. Closer examination of the Gold ranked top pose for 3a suggested a docking pose where the alkyl chain of the SSS chelate protrudes out of the ER α protein through a cavity past HIS524 and ILE424 for reference. This top ranked docking pose was contrary to that found within the ethyl analogue 2a in which the SSS chelate sits beneath the estradiol core in an envelope conformation as in the resolved crystal structure of 3a. While this pose of 3a may well exist it is hard to envision energetically why it would be more favorable to undertake this pose *in vivo* given that it would require energy to contort into this elongated pose and protrude through the protein wall. Indeed Gold found the envelope conformation of 3a to be the second most favorable pose according to the Gold fitness function. Utilizing the second ranked pose for 3a, the RBA was recalculated and the ranking of the seven compounds was repeated. This resulted in a more accurate ranking of the compounds in the blind test set as shown in Table 5 under the heading corrected. Now not only were the most active compounds correctly ranked but in addition 12a was correctly predicted to be the least active compound of the seven.

In conclusion we have utilized a genetic algorithm to develop new force field parameters to model Re(I) and Re(V) organometallic complexes. We have used these new parameters to perform unbiased docking of a series of novel estrogen receptor ligands including several rhenium based ligands utilizing the accurate docking program Gold. Furthermore we have utilized the computational technique CoMSIA in order to study and predict the estrogen receptor binding affinities of this diverse set of compounds. Our developed OPLS-AA/L force field parameters were able to accurately reproduce the crystal structures of the blind validation set and facilitated our further computational studies particularly docking and the electrostatic description of the molecular fields which proved critical to the development of predictive CoMSIA models. Our unbiased docking study highlighted several points of interest including most significantly the favorable location of the tricarbonyl rhenium moiety in the α -face of the ER α LBD. Robust CoMSIA models, consisting of several classes of ER ligands, have been developed and validated within the framework of our original set of ligands. Indeed, we have

shown how these predictive CoMSIA models can be used to focus and prescreen new ligands for their affinity to ER α .

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors wish to acknowledge the National Institute of Biomedical Imaging and Bioengineering, EB00340, for funding this work.

References

1. Mortensen DS, Rodriguez AL, Sun J, Katzenellenbogen BS, Katzenellenbogen JA. Furans with Basic Side Chains: Synthesis and Biological Evaluation of a Novel Series of Antagonists with Selectivity for the Estrogen Receptor Alpha. *Bioorg & Med Chem Lett* 2001;11:2521–2524. [PubMed: 11549460]
2. Stauffer SR, Coletta CJ, Tedesco R, Nishiguchi G, Carlson K, Sun J, Katzenellenbogen BS, Katzenellenbogen JA. Pyrazole Ligands: Structure-Affinity/Activity Relationships and Estrogen Receptor- α -Selective Agonists. *J Med Chem* 2000;43:4934–4947. [PubMed: 11150164]
3. Meyers MJ, Sun J, Carlson KE, Katzenellenbogen BS, Katzenellenbogen JA. Estrogen Receptor Subtype-Selective Ligands: Asymmetric Synthesis and Biological Evaluation of cis- and trans-5,11-Dialkyl-5,6,11,12-tetrahydrochrysenes. *J Med Chem* 1999;42:2456–2468. [PubMed: 10395487]
4. Dehdashti F, Flanagan FL, Mortimer JE, Katzenellenbogen JA, Welch MJ, Siegel BA. Positron emission tomographic assessment of “metabolic flare” to predict response of metastatic breast cancer to antiestrogen therapy. *Eur J Nuc Med* 1999;26:51–56.
5. DiZio JP, Fiaschi R, Davison A, Jones AG, Katzenellenbogen JA. Progesterone-rhenium complexes: metal-labeled steroids with high receptor binding affinity, potential receptor-directed agents for diagnostic imaging or therapy. *Bioconj Chem* 1991;2:353–366.
6. Chi DY, O’Neil JP, Anderson CJ, Welch MJ, Katzenellenbogen JA. Homodimeric and heterodimeric bis(amino thiol) oxometal complexes with rhenium(V) and technetium(V). Control of heterodimeric complex formation and an approach to metal complexes that mimic steroid hormones. *J Med Chem* 1994;37:928–937. [PubMed: 8151620]
7. Hom RK, Chi DY, Katzenellenbogen JA. Heterodimeric bis(amino thiol) complexes of oxorhenium (V) that mimic the structure of steroidal hormones - synthesis and stereochemical issues. *J Org Chem* 1996;61:2624–2631. [PubMed: 11667091]
8. Skaddan MB, Katzenellenbogen JA. Integrated “3+1” Oxorhenium(V) Complexes as Estrogen Mimics. *Bioconj Chem* 1999;10:119–129.
9. Luyt LG, Bigott HM, Welch MJ, Katzenellenbogen JA. 7[α]- and 17[α]-Substituted estrogens containing tridentate tricarbonyl rhenium/Technetium complexes: synthesis of estrogen receptor imaging agents and evaluation using microPET with technetium-94m. *Bioorg & Med Chem* 2003;11:4977–4989. [PubMed: 14604660]
10. Chi DY, Katzenellenbogen JA. Selective formation of heterodimeric bis-bidentate aminothiols oxometal complexes of rhenium(V). *J Am Chem Soc* 1993;115:7045–7046.
11. Jaouen G, Top S, Vessières A, Pigeon P, Leclercq G, Laios I. First anti-oestrogen in the cyclopentadienyl rhenium tricarbonyl series. Synthesis and study of antiproliferative effects. *Chem Commun* 2001:383–384.
12. Bigott HM, Parent E, Luyt LG, Katzenellenbogen JA, Welch MJ. Design and synthesis of functionalized cyclopentadienyl tricarbonylmetal complexes for technetium-94m PET imaging of estrogen receptors. *Bioconj Chem* 2005;16:255–264.
13. Wolohan P, Reichert DE. CoMFA and docking study of novel estrogen receptor subtype selective ligands. *J Comp-Aided Mol Des* 2003;17:313–328.
14. Wolohan P, Reichert DE. Use of binding energy in comparative molecular field analysis of isoform selective estrogen receptor ligands. *J Mol Graph & Mod* 2004;23:23–38.

15. Wolohan P, Yoo J, Welch MJ, Reichert DE. QSAR Studies of Copper Azamacrocycles and Thiosemicarbazones: MM3 Parameter Development and Prediction of Biological Properties. *J Med Chem* 2005;48:5561–5569. [PubMed: 16107156]
16. Wolohan P, Reichert DE. Molecular Modeling of Hexakis(areneisonitrile)technetium(I), Tricarbonyl η^5 cyclopentadienyl Technetium and Technetium(V)Oxo Complexes: MM3 Parameter Development and Prediction of Biological Properties. *J Mol Graph Model*. 2006in press.
17. Ponder, JW. TINKER. 4.2.
18. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J Am Chem Soc* 1996;118:11225–11236.
19. Kaminsky GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and Reparameterization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *J Phys Chem B* 2001;105:6474–6487.
20. Wüst F, Carlson KE, Katzenellenbogen JA, Spies H, Johannsen B. Synthesis and binding affinities of new 17 α -substituted estradiol-rhenium “n + 1” mixed-ligand and thioether-carbonyl complexes. *Steroids* 1998;63:665–671. [PubMed: 9870263]
21. Mull ES, Sattigeri VJ, Rodriguez AL, Katzenellenbogen JA. Aryl Cyclopentadienyl Tricarbonyl Rhenium Complexes: Novel Ligands for the Estrogen Receptor with Potential Use as Estrogen Radiopharmaceuticals. *Bioorg & Med Chem* 2002;10:1381–1398. [PubMed: 11886802]
22. Top S, El Hafa H, Vessieres A, Quivy J, Vaissermann J, Hughes DW, McGlinchey MJ, Mornon J-P, Thoreau E, Jaouen G. Rhenium carbonyl complexes of β -estradiol derivatives with high affinity for the estradiol receptor: an approach to selective organometallic radiopharmaceuticals. *J Am Chem Soc* 1995;117:8372–8380.
23. Allen FH, Kennard O. 3D Search and Research Using the Cambridge Structural Database. *Chemical Design Automation News* 1993;8:31–37.
24. Schaftenaar G, Noordik JH. Molden: a pre- and post-processing program for molecular and electronic structures. *J Comp-Aided Mol Des* 2000;14:123–134.
25. Strassner T, Busold M, Herrmann WA. MM3 parametrization of four- and five-coordinated rhenium complexes by a genetic algorithm - which factors influence the optimization performance? *J Comp Chem* 2002;23:282–290. [PubMed: 11924740]
26. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model. *J Phys Chem* 1993;97:10269–10280.
27. Jaguar. 5.0. Schrödinger, Inc; Portland, OR:
28. Brzozowski AM, Pike ACW, Dauter Z, Hubbard RE, Bonn T, Engström O, Öhman L, Greene GL, Gustafsson J-Å, Carlquist M. Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature* 1997;389:753–758. [PubMed: 9338790]
29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Research* 2000;28:235–242. [PubMed: 10592235]
30. Maestro Molecular Modeling Interface. 5.1.020. Schrödinger, Inc; Portland, OR:
31. Pike ACW, Brzozowski AM, Hubbard RE, Bonn T, Thorsell A-G, Engstrom O, Ljunggren J, Gustafsson J-A, Carlquist M. Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. *EMBO J* 1999;18:4608–4618. [PubMed: 10469641]
32. Tripos SYBYL. 7.0. Tripos Inc; St Louis:
33. Stauffer SR, Coletta CJ, Tedesco R, Nishiguchi G, Carlson K, Sun J, Katzenellenbogen BS, Katzenellenbogen JA. Pyrazole Ligands: Structure-Affinity/Activity Relationships and Estrogen Receptor- α -Selective Agonists. *J Med Chem* 2000;43:4934–4947. [PubMed: 11150164]
34. Mortensen DS, Rodriguez AL, Sun J, Katzenellenbogen BS, Katzenellenbogen JA. Furans with Basic Side Chains: Synthesis and Biological Evaluation of a Novel Series of Antagonists with Selectivity for the Estrogen Receptor Alpha. *Bioorg & Med Chem Lett* 2001;11:2521–2524. [PubMed: 11549460]

35. Stauffer SR, Huang YR, Aron ZD, Coletta CJ, Sun J, Katzenellenbogen BS, Katzenellenbogen JA. Triarylpyrazoles with Basic Side Chains: Development of Pyrazole - Based Estrogen Receptor Antagonists. *Bioorg & Med Chem* 2001;9:151–161. [PubMed: 11197335]
36. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: A Program to generate schematic diagrams of protein - ligand interactions. *Protein Engineering* 1995;8:127–124. [PubMed: 7630882]

Figure 1.
Estrogen receptor ligand classes used in this study with their pharmacophoric elements highlighted.

Figure 2. Reference compounds used to develop the Re(V)-oxo and Re(I)- η^5 cyclopentadienyl tricarbonyl parameters.

Figure 3. Compounds used to validate the Re(V)-oxo and Re(I)- η^5 cyclopentadienyl tricarbonyl parameters.

(a) 12a in ER α

(b) 12h in ER α

(a) 12a in ER α

(b) 12h in ER α

Ligand bond
 Non-ligand bond
 Hydrogen bond and its length

Non-ligand residues involved in hydrophobic contact(s)
 Corresponding atoms involved in hydrophobic contact(s)

Figure 4.

Figure 4a. Unbiased best fit configuration of of ReCPs (a) 12a in ER α and (b) 12h in ER α with the protein residues of interest highlighted.

Figure 4b. Schematic of H-Bonding network of ReCPs (a) 12a in ER α and (b) 12h in ER α generated using LIGPLOT.³⁶

Figure 5.

CoMSIA derived molecular field maps from the most predictive ER α CoMSIA model with 12h in it's docked ER α pose illustrated. Contours of the steric map are shown in transparent yellow and green, those of the electrostatic map are shown in solid red and blue, while those of the hydrophobic map are shown in cyan and purple. Increased biological activity is correlated with: more bulk near green; less bulk near yellow; more positive charge near blue, more negative charge near red, more hydrophobicity near cyan and less hydrophobicity near purple.

Table 1
Experimental relative binding affinities (RBA) of 17 β -estradiol analogues.

ID	Ligand	R-Group				RBA ER α
		R1	R2	R3	R4
	<i>estradiols</i>				
17 β -Estradiol		H	H			100
11 β OMeEST		OCH ₃	H			86
11 β MeClEST		ClCH ₂	H			260
17 (ZAZNIX)		H				15
19		ClCH ₂				172
2a		H				7.4
3a (AMORAV)		H				14.1
	<i>cyclopentadienes</i>				
10c		Ar'	H	Me		1.2 $\pm$ 0.2
10d		Ar'	H	Et		0.37 $\pm$ 0.06
10e		Ar'	H	Ar'		3.8 $\pm$ 1.7
10f		Ar'	Me	Me		0.12 $\pm$ 0.02
10g		Ar'	Et	Et		8.6 $\pm$ 0.4
10h		Ar'	Ar'	Et		4.9 $\pm$ 0.6
	<i>Re tricarboxyl cyclopentadienes</i>				
12a		H	H	H		0.020 $\pm$ 0.0
12b		H	H	Ar'		2.4 $\pm$ 0.1
12c		Ar'	H	Me		1.8 $\pm$ 0.4
12d		Ar'	H	Et		2.5 $\pm$ 0.07
12e		Ar'	H	Ar'		10.5 $\pm$ 0.0
12f		Ar'	Me	Me		3.6 $\pm$ 0.9
12g		Ar'	Et	Et		4.3 $\pm$ 1.1
12h		Ar'	Ar'	Et		23 $\pm$ 6
	<i>Furans</i>				
15a		Ar'	Ar'	Me	Ar'	40 $\pm$ 6.5
15b		Ar'	Ar'	Et	Ar'	140 $\pm$ 38

ID	Ligand	R-Group				RBA ER α
		R1	R2	R3	R4	
15c		Ar'	Ar'	Pr	Ar'	100 $\pm$ 14
15d		Ar'	Ar'	Bu	Ar'	21 $\pm$ 0.6
4f		Ar'	Ar'	Et	Ar'	36 $\pm$ 6
4g		Ar'	Ar'	Pr	Ar'	49 $\pm$ 12
4h		Ar'	Ar'	<i>i</i> -Bu	Ar'	75 $\pm$ 6
4i		Ar'	Ar'	<i>n</i> -Bu	Ar'	14 $\pm$ 4

[Ar' = p-(OH)-C₆H₄, Me = CH₃, Et = CH₂CH₃, Pr = (CH₂)₂CH₃, Bu = (CH₂)₃CH₃]

Determined by competitive radiometric binding assay, where the RBA of 17 β -Estradiol is arbitrarily set at 100. Values represent the average of multiple determinations.

Table 2
Bond Stretching OPLS-AA/L Parameters for Re(I) and Re(V)oxo atom types.

OPLS-AA/L atom type and force field identifier	Bond stretching parameter	
	bond length (Å)	force constant (kcal/mol Å ²)
Re(V)		
Re(203)=O(206)	1.72	414.67
Re(203)-S _{sp} ² (205)	2.26	205.42
Re(203)-S _{sp} ³ (212)	2.40	483.68
Re(I)		
Re(204)-C _{cyclopentadienyl} (207)	2.22	240.41
Re(204)-C _{carbonyl} (208)	1.94	226.50

Table 3

Comparison of Calculated OPLS Structure to X-ray for Re(V)-Oxo and Re(I)-Tricarbonyl η^5 Cyclopentadienyl Reference Set and Validation Set.

Reference Set		Validation Set	
REFCODE	rmsd [U] (Å) OPLS	REFCODE	rmsd [U] (Å) OPLS
AMORAV	0.293	AKOQOG	0.212
AMONOF	0.236	AMOPAT	0.238
AMOPIB	0.355	LUYXAE	0.427
MICSAS	0.444	ACEFAP	0.350
ACYPRE	0.233	ACEFET	0.387
COCPRE	0.095	QEWVIX	0.156
TEBPUL	0.149	TEBPOF	0.205
ZAZNIX	0.330	VODXAN	0.137
Average	0.267		0.240

rmsd [U] refers to the unit weighted root-mean-squared deviation.

Table 4
Results from several PLS models for the ER α training set.

PLS Model	r ²	q ²	SEE	F	PC	% Contribution			
						S	E	H	Other
<i>Gold docked conformers</i>									
Log RBA v CoMSIA (S+E)	0.96	0.56	0.20	73	2	27	73		
Log RBA v CoMSIA (S+E+H)	0.93	0.48	0.23	85	3	16	49	35	4
Log RBA v CoMSIA (S+E+H) + PV	0.94	0.68	0.24	62	4	15	38	31	15
Log RBA v GoldH_Ext	<0.1	−0.18	0.85	<1	1				100
Log RBA v GoldVDW_Ext	<0.1	−0.17	0.86	<1	1				100
Log RBA v Gold_Fitness	<0.1	−0.20	0.86	<1	1				100
<i>Organics only</i>									
Log RBA v GoldH_Ext	<0.1	−0.32	0.98	1	1				100
Log RBA v GoldVDW_Ext	<0.1	−0.30	0.97	1	1				100
Log RBA v Gold_Fitness	0.30	−0.02	0.82	1	1				100

r^2 refers to the non-validated correlation, q^2 to the cross-validated correlation, F is a measure of statistical significance, SEE refers to the standard error of estimate, S refers to the steric field, E refers to the electrostatic field, H refers to the hydrophobic field, Other refers to the contribution of any other descriptor used in the PLS model, GoldH_Ext, GoldVDW_Ext and Gold_Fitness refer to the addition of the external hydrogen bonding, external Van der Waals and fitness function all from the Gold docking run.

Table 5

Ability of the Log RBA v CoMSIA (S+E+H) + PV model to rank the ligands in the test-set in terms of their RBA ER α .

Ligand	Exp.		Pred. (S+E+H) + PV		Corrected (S+E+H) + PV	
	RBA	Rank	RBA	Rank	RBA	Rank
17 (ZAZNIX)	172	1	129	1	129	1
15d	21	2	120	2	120	2
4i	14	3	56	3	56	3
2a (AMORAV)	14	3	0.33	7	35	4
10e	3.8	5	18	5	18	6

Ligand	Exp.		Pred. (S+E+H) + PV		Corrected (S+E+H) + PV	
	RBA	Rank	RBA	Rank	RBA	Rank
12d	2.50	6	22	4	22	5
12a	0.02	7	0.91	6	0.91	7

Exp. refers to experimental data, Pred. refers to predicted data and Corrected refers to the predicted values with the corrected 3a docking pose.