Application of Molecular Mechanics to Prediction of Binding Affinities of Estrogen Agonists to the Human Estrogen-α receptor using its X-ray crystallographic structure.

Mikael Gillner, Peter Carlsson
KaroBio AB and Center for Structural Biochemistry, NOVUM,
S-141 57 Huddinge, Sweden.

Paulette Greenidge
KaroBio AB, NOVUM,S-141 57 Huddinge, Sweden.

INTRODUCTION

Estrogens or estrogenic compounds have profound effects on the reproductive system. They include (i) analogues and metabolites of the endogenous hormone 17β-estradiol (E2), (ii) synthetic pharmacological estrogen receptor (ER) agonists and/or antagonists, (iii) environmental estrogens originally produced by the chemical industry that are estrogenic environmental contaminants, (iv) and finally phytoestrogens that are produced by certain plants. The environmental estrogens have been associated with developmental, reproductive and other problems in wildlife and laboratory animals. Some of them are contained in plastics (e.g. bisphenol A) and are produced by the breakdown products of detergents (e.g. nonyl phenol and octyl phenol). In addition, there are 209 possible isomers of (polychlorobiphenyls) PCBs having varying levels of estrogenicity that have been used in electrical transformers and cooling systems. Despite the banning of PCBs, their environmental levels remain high. The pesticide DDT (2,2-bis(p-chlorophenyl)-1,1,1-trichloroethane) is also estrogenic. Although environmental estrogens bind more weakly to the estrogen receptor than endogenous ligands, they tend to have long biological and environmental half lives, bioaccumulating in the fat and tissue of living organisms due to that they are bad substrates for the enzymes involved in the metabolism of xenobiotics and their lipophilicity. This is in contrast to endogenous and plant estrogens, which remain in the blood stream for at most a few hours, before being metabolized by enzymes in the liver to water soluble products, whereupon they are excreted. Thus, due to the possible detrimental effects of environmental estrogens on humans, combined with their widespread occurrence and the costs of assaying large numbers of compounds, a computational screening method would be of value for selection and prioritization of compounds for binding assay. In this paper we explore the utility of molecular mechanics as implemented in the CHARMm software for prediction of ER-ligand binding affinities of environmental estrogens using the crystallographic X-ray structure of the estrogen receptor (1). We have used the data set of Waller et al. that earlier was used for a COMFA study of the same problem (2). The set consisted of of 57 estrogen agonists from 8 structurally different chemical classes.

MATERIALS AND METHODS

The estradiol molecule was removed from the receptor-ligand complex structure, and the waters retained. Hydrogens atoms were added to the X-ray crystallographic structure. Partial atomic charges for the ligands were calculated by fitting the van der Waals surfaces of the molecules to their 6-31G* electrostatic potentials (3), as implemented in Gaussian 94. Sodium and chloride counterions were placed at the maxima and minima of the protein electrostatic

potential near charged amino acid residues as to achieve net neutrality of the system. The C- and N-termini were made neutral.

Four rotamers were generated for each ligand with one or more rotatable bonds. The ligand structures were were translated into the binding site in the same position and orientation as estradiol, and rotated. The ligands were thereafter energy-minimized with CHARMM using the adopted-basis Newton-Raphson algorithm. The binding site whas kept rigid. The all-atom force field and parameters as implemented in QUANTA 96 were used. The non-bonded interactions were cut-off beyond a distance of 15Å; switching (van der Waals) and shifting (electrostatics) functions were turned on between 11 and 14Å. The default heuristic non-bonded list-update method was used. Energy minimizations were continued until the gradient norm was less than 0.05 kcal Å. A distance dependent dielectric function was used. The protein-ligand interaction energies were calculated for each minimized conformation, and the lowest energy so obtained for each ligand was used for correlation to experimental data.

RESULTS AND DISCUSSION

In the crystal structure complex of E2 and human ER-α, the A-ring phenolic hydroxyl makes direct hydrogen bonds to the carboxylate of Glu 353, the guanidinium group of Arg 394, and a water molecule. The D-ring's 17-β hydroxyl makes a single hydrogen bond with His 524. The A-ring is sandwiched between the sidechains of hydrophobic residues on its α- and β- faces, and the D ring has non-polar contacts with hydrophobic residues. The other main interactions of the ligand with the protein are non polar. Analogs of estradiol that have a 3-position moiety that is only capable of acting as a hydrogen bond acceptor and not a donor, have poor binding affinities. This may be due to electrostatic repulsion between the 3-position heteroatom of the ligand and the carboxylate of the Glu 353. The hydrogen bond accepting feature of His 524 as the histidine residue seems to be quite flexible if one compares agonist and antagonist structures, whereas Glu 353 forms a salt bridge with Arg 394 and thus its location is more well defined.

For the present study we have used the X-ray crystallographic structure of the estradiol estrogen receptor-α complex for prediction of ligand binding by molecular mechanichs based scoring functions. The strengt of the prediction was for the purpose of scoring function optimisation assessed by the correlation of the molecular mechanics ligand – protein interaction energy with the binding energies calculated from experimental data. Starting from a simplistic model, we tried to improve the obtained correlation by the use of more sophisticated models with respect to the charges of the ligands. Furthermore, we investigated if correction of the molecular mechanics energy for the solvation and strain energies of the ligands would further improve the correlation.

The experimental data set used was that of Waller et al. earlier used for a COMFA study of estrogen receptor ligand binding (2). This data set is relatively diverse and consists of both environmental, pharmacological and physiological estrogens. The environmental estrogens are represented by plant estrogens as well as environmental contaminants of industrial origins with diverse chemical structures (2). Specifically, the data set contained of 16 PCBs, 6 phenols, 2 pthalates, 3 phytoestrogens, 6 and 7 molecules respectively structurally related to DDT and diethylstilbesterol (DES) respectively, 8 other pesticides, one androgen inhibitor, three plant estrogens and 9 steroids, of which one was excluded since it is an antagonist.

Whether the estrogen receptor provides a hydrogen bond donor or acceptor for the 17-OH group of estradiol, or corresponding hydroxyls of other ligands, is not clear from the structure-affinity relationships of its ligands. In the X-ray crystallographic structure of the estradiol-estrogen receptor complex the 17-O of estradiol is located 3.0 Å from the His-524 ND1. ND1 is probably an acceptor for the hydrogen of the 17-OH group. We have carried out all our molecular mechanichs calculations for both situations, i. e. where the histidine is represented as either a hydrogen bond donor or acceptor. Ethinylestradiol was not included in the dataset

because it is an antagonist, and can not be accomodated in the binding site of the agonist structure. Endosulfan, Kepone and Lindane were excluded from the data set because the programs for solvation energy calculation were not able to handle them.

Initially Gasteiger-Huckel point charges were used for the ligands, wich yielded a correlation coefficient of 0.65 for correlation with the experimental data (with the ligand binding histidine as a donor). When we instead used 6-31G* espfit charges for the ligand atoms a marked improvement of the same correlation was obtained (to 0.77).When the molecular mechanics energies were corrected for strain we did not see any improvement (c.f Table 1). We also tried to correct the molecular mechanics energies with the free energy of solvation of the ligands. To this end we used the solvent continuum model of Clark Still (GB/SA algorithm) (4) as implemented in ETH/Yak and its successor ETH/PrGen. With both of these implementations we obtained a modest improvement for both the uncorrected and strain-corrected molecular mechanics energies (c.f Table 1). The correlations the molecular mechanics energies and experimental binding affinities then obtained were comparable to those earlier obtained by Grootenhuis and van Galen in a study of thrombin inhibitor binding (5).

In conclusion, the molecular mechanics ER protein-ligand interaction energy is predictive with respect to to ER ligand affinities. In combination with a high capacity electronic screening method such as e.g. structure-based 3D pharmacophore searching (as described for the thyroid hormone receptor by Greenidge et al, ms in preparation) the molecular mechanics protein-ligand interaction energy (with or without correction for solvation may with advantage be used for identification of environmental estrogens and/or prediction of their binding affinities for the estrogen receptor.

Table 1. Ligand binding to the estrogen receptor studied by molecular mechanics. The binding of 53 compounds[1] from the data set of Waller et al. (2) to the estrogen receptor X-ray crystallographic structure was assessed by molecular mechanics using the Charmm force-field and compared to the experimental relative binding affinities. 6-31G* ESP point charges were used for all ligands. The numbers represen the cofficients for the correlation of the molecular mechanics energies (without or with correction for ligand strain) with the relative free energies for binding, when the ligand binding site histidine was represented as a hydrogen bond donor or acceptor. The molecular mechanics energies were also corrected for the free energy of solvation of the ligands using the implementation of GB/SA algorithm distributed with the Yak or PrGen programs.

Solvation correction	None	Yak	PrGen
His acceptor	0.76	0.77	0.78
His donor	0.77	0.79	0.81
His acceptor Strain. corr.	0.72	0.74	0.75
His donor Strain. corr.	0.76	0.77	0.79

[1]Endosulfan, Kepone, Lindane och Ethinylestradiol were excluded due to problems with the solvation energy calculations.

REFERENCES

1. Brzozowski, A.M., Pike, A.C., Dauter Z, Hubbard, R.E., Bonn, T., Engstrom, O., Ohman. L., Greene, G..L, Gustafsson, J.A., and Carlquist,M. Molecular basis of agonism and antagonism in the oestrogen receptor *Nature*, **389**, 753-758 (1997).
2. Waller CL, Oprea TI, Chae K, Park HK, Korach KS, Laws SC, Wiese TE, and Kelce WR, Ligand-based identification of environmental estrogens. *Chem Res Toxicol.*, **9**, 1240-1248 (1996).
3. Sing, U. C. S. and Kollman, P. An approach to computing electrostatic charges. *J. Comp. Chem.* **5**, 129-145 (1984)
4. Gray LE Jr, Still, A. Tempczyk, R.C. Hawley and T. Hendrickson, Semianalyticat treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.*, **112**, 6127 (1990).
5. Grootenhuis, P and van Galen, P. J.M. Correlation of binding affinities with non-bonded inter-action energies of Thrombin-inhibitor complexes. *Acta. Cryst.* **D51**: 560-566 (1995).