An In Silico Prediction of ER-alpha Agonism for Quats and Phthalates

Andrew Tantisukarom

ABSTRACT

Endocrine Disruptors are a growing concern as the female population is experiencing precocious maturation and more seriously, a greater incidence of breast cancer. Endocrine disruptors, or mimickers, can be found naturally in the environment but are most prevalent in man-made subtances. In this study, I investigated two classes of compounds- Ammonium Quaternary Compounds and Phthalates- that are less represented in the literature in respects to Estrogen Receptor-alpha (ER-a) agonism, which is linked to breast cancer. I employed Computational Toxicology, an in silico approach using state-of-the-art structure analyses modeling, specifically MDL-QSAR. The study collectively examined 70 quats and phthalates and predicted that 10 compounds showed significant ER-a agonism.

KEYWORDS

Computational Toxicology, Quaternary Ammonium Compound, Breast Cancer, MDL-QSAR, Endocrine Disruptor

INTRODUCTION

Endocrine disrupters are hormonally active agents that mimic the function of endogenous (bodily) hormones of the endocrine system, commonly referred to as the hormone system. Endocrine systems are found in all mammals, birds, fish, and many other types of living organisms. Hormones are involved in just about every biological function (Jesperson). Via the chemical system of hormones, the endocrine system regulates all biological processes in the body from conception through adulthood and into old age, including the development of the brain and nervous system, the growth and function of the reproductive system, as well as metabolism and blood sugar levels (FDA 2010).

Risks of exposure to endocrine disruptors include adverse effects to the developmental, reproductive, neurological, and immune functions of humans. (National Library of Medicine 2009). Some notable examples of endocrine disrupting compounds include: Diethylstilbestrol (DES), a prescribed prenatal drug that was halted by the U.S Food and Drug Administration in 1971 because it was linked to a rare vaginal cancer (U.S. EPA 2010); Polychlorinated biphenyl (PCBs) production in the U.S. stopped in 1977 and later exports and imports in 1979 because of suspected harmful health and environmental effects; and more notoriously, Dichlorodiphenyltrichloroethane (DDT) was banned by the U.S. EPA in 1972 because it posed high risks to the environment and human health (ExtoxNet 2002). Currently, the population at greatest risk is females, as a greater number of the youth reach maturation sooner and as breast cancer rates have risen. The incidence of breast cancer was estimated to be approximately 30% greater between 1975 and 2000 (Szabo 2006). This is a result of endocrine disruption of estrogen receptors, which respond to hormones responsible for female sexual development (National Cancer Institute 2006).

There are two distinct estrogen receptors in the human body: Estrogen Receptor-alpha (ER-a) and Estrogen Receptor-beta (ER-b). Both bind to the hormone Estrogen as well as other chemicals that act as agonists, receptor stimulants, or antagonists, receptor repressors. The two receptors have distinct localizations and concentrations within the body. Structural differences between the two allow for a wide range of diverse and complex bodily processes (McClure 2001).

A wide range of substances, both natural and man-made, are thought to cause endocrine disruption. Endocrine disruptors may be found in many everyday products including plastic bottles, metal food cans, detergents, flame retardants, food, toys, cosmetics, and pesticides (National Library of Medicine 2009). Two classes of compounds used prevalently in industry and in day-to-day activities are Quaternary Ammonium Compounds ("Quats") and Phthalates. Quats are commonly found in surfactants, disinfectants, and antistatic agents which are present in shampoo, spermicidal jellies, liquid fabric softeners, and dryer anticling strips. The food industry uses quats in conjunction with bleach products as a sanitizing agent. In addition, some are used in surgery as anesthetics (Rahn 1974). Phthalates are commonly found in flooring, cosmetics, soft toys, medical equipment (i.e. medical tubing, catheters and blood bags) and even air fresheners (Environmental Working Group 2007).

Specifically, I am attempting to determine the estrogen-stimulating capacity, or agonism, of two chemical classes: Quats and Phthalates. I will be employing Computational Toxicology, an in silico, or computer-based, modeling method to investigate ER-a, as experiments have shown a strong link between receptor agonism and breast cancer. I am utilizing predictive chemical structure analyses models for ER-a to assess the threat of Quats and Phthalates to the female population in respect to breast cancer.

METHODS

The methods employed are based on Computational Toxicology principles of modeling and predicting chemical activity on a specific biological endpoint. Computational Toxicology utilizes Quantitative Structure Activity Relationship (QSAR) modeling software. MDL-QSAR was the software of choice. Figure 1 is a flow chart of how the model was built.

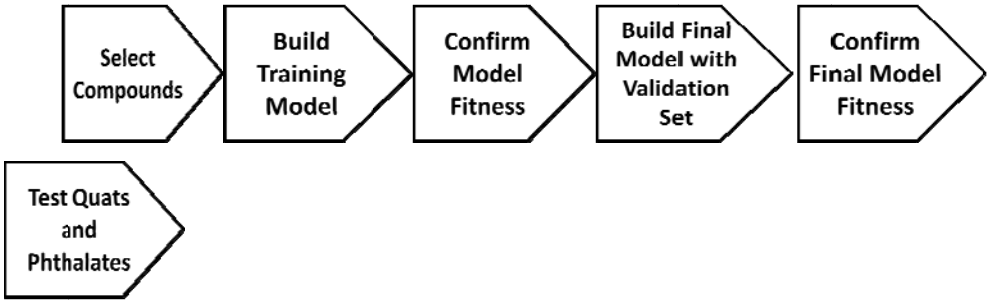

Figure 1. Model construction flow chart. The methods employed through MDL-QSAR are sequenced.

I used the peer-reviewed articles by Escande et al. (2006), Fitzpatrick et al. (2007), Kostelac et al. (2003), Moein et al. (2008), Overk et al. (2005), and Pillon et al. (2005) and the chemical databases PubChem and bindingdb.org to research compounds active on ER-a, the biological endpoint of interest. The compounds needed to have a corresponding EC50 or pEC50 (-logEC50) value, a common toxicological potency measure, as MDL-QSAR predicts in terms of EC50s. These compounds were used to create a highly predictive model and for simplicity will be referred to as Group A. During this initial research, I also chose the compounds I wished to test- Quaternary Ammonium Compounds (Quats) and Phthalates. These compounds will be referred to as Group B. After, I found the SDFs, or Structure Data Format, a computer binary code that describes one specific compound, for Group A and Group B. If PubChem, a chemical database where SDF files are archived, did not have a valid SDF, then the compound was discarded.

A Training Set of 69 compounds (Appendix A) and SDFs from Group A was created to be initially exported into MDL-QSAR. This Training Set was the basis for the final MDL-QSAR model. After arriving at a Training Model (corresponding to the Training Set) deemed "statistically very significant" by randomly adjusting model parameters, I verified its capabilities using a Validation Set (Appendix B), which consisted of the 35 remaining compounds from Group A not included in the Training Set. After obtaining predicted pEC50 values for the Validation Set, converting them back to EC50 values, and graphing them against the known, scientifically-determined EC50s, I obtained an R^2 value of greater than 0.70, indicating that the model was strong enough to predict agonistic activity of compounds relative to ER-α. If the results for the Validation

Set did not yield an R^2 value greater than 0.70, I readjusted the model parameters for the Training Model and revalidated the model.

Upon creating a strong predictive model (Figure 4) I uploaded the Unknown Set consisting of Group B compounds and ran them through the model. After obtaining the predicted pEC50 values and converting them back to EC50s, I scaled the results according to the distribution of values. I distributed the spread between the lowest and highest EC50 predicted by the model on a log scale. 17-beta-estradiol, a known ER-α agonist, displayed "Strong Agonism" and was used to confirm the reasonability of the scale.

RESULTS

A total of 104 compounds (Group A) were used to build a predictive model through MDL-QSAR.

Training model

The Training Model consisted of 69 compounds with the regression equation y=0.6773x-0.6512 and an R-squared value of 0.7131 (Figure 2). Figure 3 shows the fitness of the Training Model's regression equation. Upon attaining a Training Model with an R-squared greater than 0.7131, it was able to be verified by the Validation Set.

Validation and final model

The 35 compounds of the Validation Set are incorporated into the final, verified MDL-QSAR model (Figure 4). The final predictive model is described by the regression equation: pEC50 = -0.37719*SHBint8_Acnt - 0.21029*SsCH3_acnt + 0.21145*SssCH2_acnt + 11.5*SsOH_acnt + 0.49269*SssO_acnt - 3.9102*SHsOH - 3.523592. The R-squared value of the model is 0.7872 with a standard error estimation of 0.5327, F-statistic of 35.05 and a P-value of 0.00. The model was determined by MDL-QSAR to be well described by the regression equation and thus statistically very significant with high predictive power.

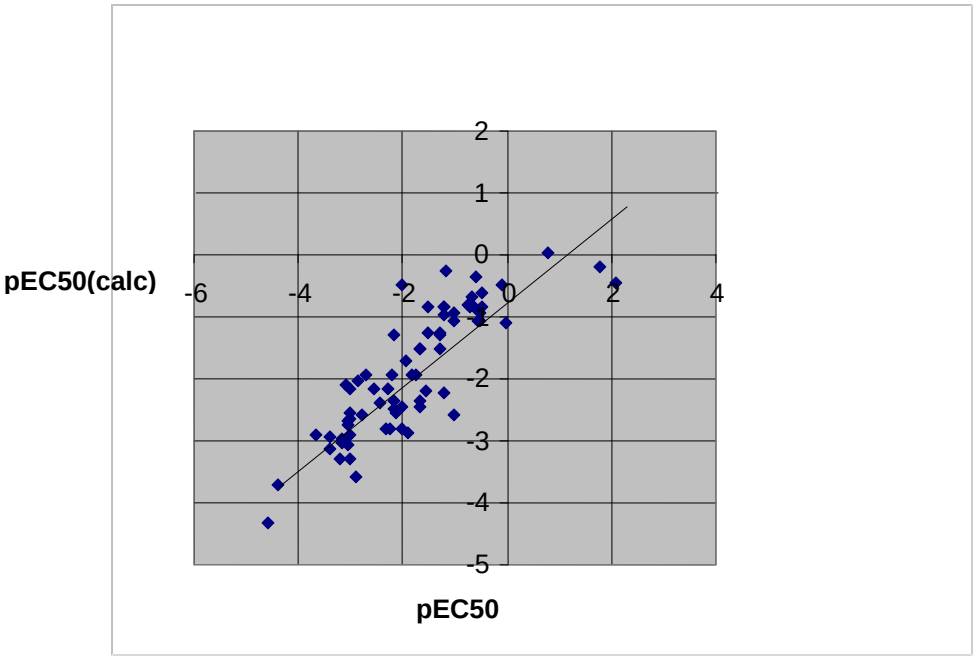

Figure 2. MDL-QSAR Training Model regression plot. The regression plot compares model-derived pEC50s against the actual pEC50s for the Training Model's 69 compounds and depicts a positive linear regression.

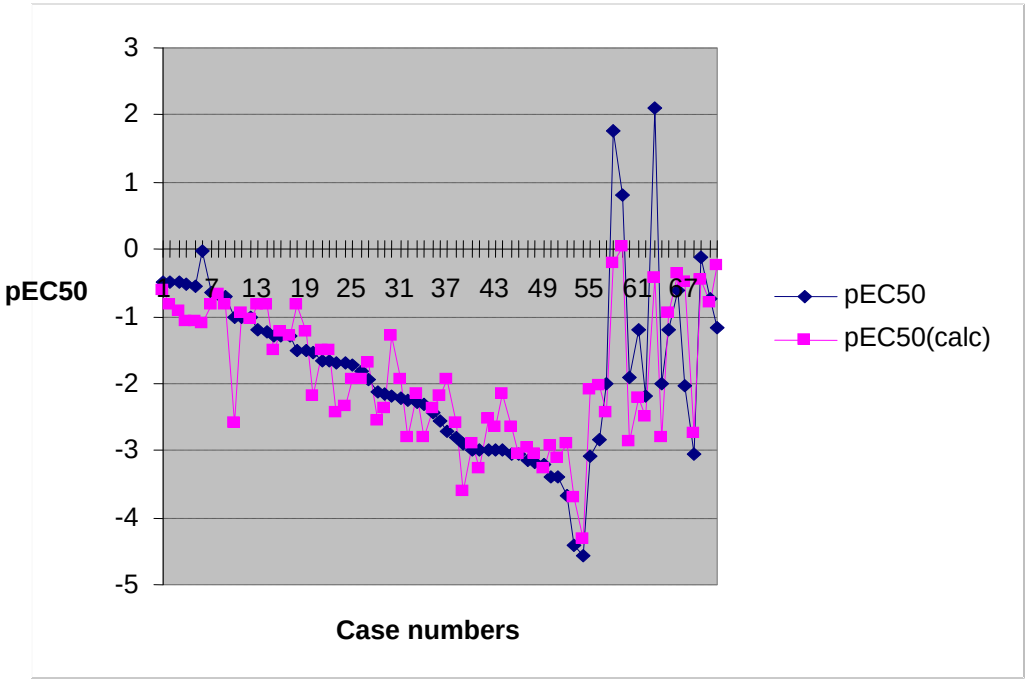

Figure 3. Training Model fitness. The plot represents the fitness of the regression equation for the Training Model by juxtaposing model-derived and actual values of pEC50.

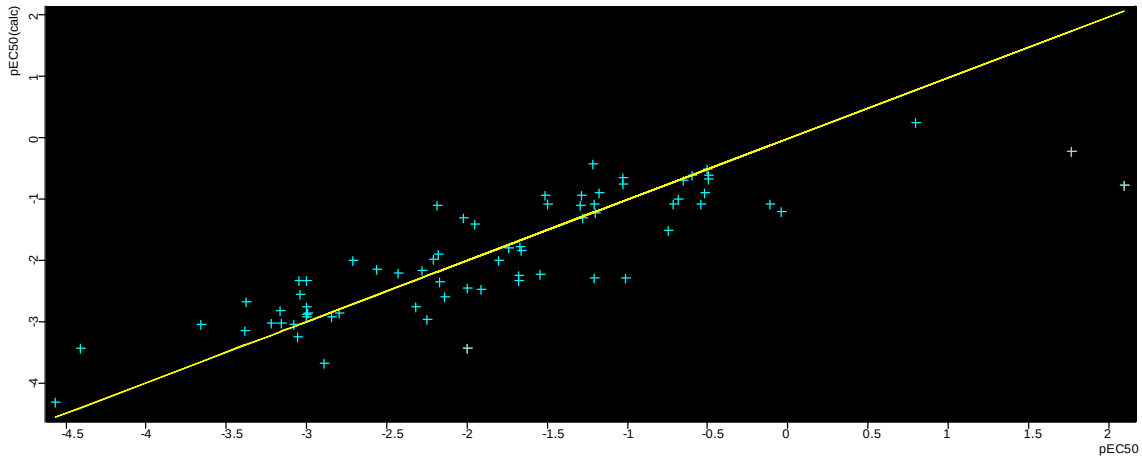

Figure 4. Final MDL-QSAR Testing Model. Above is a picture of the final verified MDL-QSAR model used to predict the agonism for the quats and phthalates.

A total of 49 quats and 21 phthalates were run through the MDL-QSAR model. Table 1 and Table 2 list each quat and phthalate with its model-derived pEC50 and corresponding EC50, respectively. The scale of agonism was color-coded with the color red, yellow, green, blue, and indigo (Table 3). Table 1 and Table 2 are a compilation of the findings arranged in order of the scale in Table 3: 17 quats and 6 phthalates were indigo, displaying "Negligible Agonism;" 8 quats and 4 phthalates were blue, displaying "Mild Agonism;" 8 quats and 4 phthalates were green, displaying "Moderate Agonism;" 7 quats and 3 phthalates were yellow, displaying "Strong Agonism;" most importantly, a total of 7 quats and 3 phthalates were red, displaying "Very Significant Agonism."

Compound Name	**Predicted pEC50**	**EC50**
Acital Carnatine	-4.1242	13310.67
Acroflavone	-3.2579	1810.92
Bathanachol	-4.4018	23975.18
Bibenzonium bromide	-3.0153	1027.417
Bretillium	-3.6667	5546.305
Corbachol	-4.1003	16880.46
Carnetine	-3.3498	2805.611
Citicholine	-5.6837	59121.78
Metha-choline chloride	-4.2305	18956.13
Neurine	-4.4228	32030.08
Paraquat	-3.7998	5001.051
Purvenium	-3.2015	2014.713
Quionapyramine	-3.9221	5037.01
Suxa-methonium chloride	-3.189	1773.445
Thiazenamium	-3.0015	1326.21

Tri-methylglycine	-4.3121	20946.21
Trigonelleine	-4.0145	13173.18
Benzothonium chloride	-2.2046	160.1769
Deca-methonium	-2.4593	168.6831
Danatonium	-2.8964	602.231
Di-quat	-2.9536	668.4451
Propanthalin bromide	-2.9284	676.7651
Quaternium-15	-2.7892	942.5443
Tieamonium bromide	-2.6132	380.371
Tri-ethylcholine	-2.5765	335.192
Dimethldodecyl-benzyl ammonium	-1.2332	17.10803
Carbethopendecinium bromide	-1.3052	25.41947
Centrimonium	-1.0262	22.39554
Clidnium bromide	-1.3273	59.7368
Clofillium	-1.2121	32.56167
Methscopolemine	-1.7668	55.37992
Sanguinarine	-1.299	42.06807
Tieamonium iodide	-1.6672	36.0994
Cetyl-pyridnium chloride	-0.33874	5.242502
Dequalinium	-0.74732	3.527037
Didecyl-dimethyl ammonium chloride	-0.55888	2.284915
Gallamene	-0.64182	3.333671
Miltephosine	-0.60083	3.196276
Pinavarium	-0.07914	1.115679
Tridihexethyl	-0.79281	8.00219
Atracurium	1.5469	0.02839
Bradophen	1.1450	0.07161
Cis-atracurium	1.6298	0.04492
Dimethy-ldioctadecyl ammonium chloride	3.4045	0.00045
Doxacurium Chloride	1.332	0.07021
Pipacuronium bromide	0.58043	0.31869
Rocuronium	1.4391	0.01292

Table 1. Predicted EC50 values for quats. The predicted pEC50 and corresponding EC50 is recorded and arranged according to scale.

Compound Name	**Predicted pEC50**	**EC50**
Di-methyl	-4.0014	15003.71
Di-ethyl	-3.6305	4504.225
Diallyl	-3.6305	4504.325
Di-n-propyl	-3.232	1610.2
Di-butyl	-2.7513	563.205
Disobutyl	-3.5837	4574.453
Di-n-pentyl	-2.3006	199.2316

Butyl benzyl	-2.6516	602.31
Diisohexyl	-2.974	541.2012
Diisoheptyl	-2.1947	196.337
Cyclohexyl-butyl	-1.8224	69.83932
Butyl decyl	-1.3468	27.65134
Diisononyl	-1.3202	22.67388
Diisodecyl	-1.3202	22.67388
Dicyclohexyl	-0.96691	9.433125
Dicyclohexyl	-0.9669	9.433125
Diisoundecyl	-0.39981	2.89429
Diundecyl	0.5163	0.298812
Ditridecyl	1.5279	0.033684
Diisotridecyl	0.62231	0.313041

Table 2. Predicted EC50 values for phthalates. The predicted pEC50 and corresponding EC50 is recorded and arranged according to scale.

Predicted EC50	**Predictions**
0-1	Very Significant Agonism
1.1-10	Strong Agonism
10.1-100	Moderate Agonism
100.1-1000	Mild Agonism
1000+	Negligible Agonism

Table 3. Designated scaling for predicted EC50s. Strength of Agonism is arranged according to a log scale.

DISCUSSION

Quats and Phthalates are two classes of compounds prevalently used on a daily basis and are potentially harmful endocrine disruptors. Because ER-a agonism has been linked to breast cancer, the study assessed ER-a agonism for Quats and Phthalates by creating a predictive model through MDL-QSAR.

The model provided a statistically significant prediction of the agonistic activity of Quats and Phthalates. The model found that 7 quats and 3 phthalates exhibited significant agonism. The findings are not evidence for cause, however they highlight the risk of these compounds in everyday use.

Quaternary ammonium compounds

Atracurium, Cisatracurium (trade name Nimbex), Doxacurium Chloride, Pipecurium Bromide (trade name Arduan), and Rocuronium (trade name Zemuron in the U.S. and Esmeron in most other countries) are neuromuscular-blocking drugs that are

used as muscle relaxants Wikipedia 2010). They are commonly used as anesthesia during surgery and thus are found in hospital settings. Exposure is limited to a specific population, so danger is relatively minimal since the majority of people are rarely anesthetized.

Benzoxonium Chloride (trade name Bradophen) is used as an antiseptic and disinfectant in many cleaning products. Bradophen is distributed under at least ten different names and is generally distributed in South Asian countries such as China, India, Singapore, Thailand, Hong Kong, Malaysia, and Taiwan. Medically, Benzoxonium Chloride functions as an antiseptic for skin, mucous membranes, and is also used in the treatment of dermatological diseases (MIMS USA 2004). Due to its common use as a disinfectant in cleaning products, Bradophen poses a significant threat as a carcinogen and should be studied further to determine in vivo ER-a agonism.

Dimethyldioctadecylammonium chloride (also called distearyldimonium chloride) functions as a surfactant and a detergent. In the house, it can be found in various fabric softeners, cosmetics, and hair conditioners where it is used for its antistatic effects. It is also found in many different popular lotions (Jergens, Vaseline, and Aveeno), soaps (Oil of Olay), face wash (Clearasil), and bug sprays (OFF Skintastic) (Household Products Database). Due to the prevalent use of this compound in products that are used daily, there is a great cause for concern and further studies should be done immediately to determine in vivo ER-α agonism and the actual carcinogenicity of Dimethyldioctadecylammonium chloride.

Phthalates

Diisotridecyl phthalate and Ditridecyl phthalate (commercially known as JAYFLEX DTDP) are plasticizers, which are chemicals that increase the flexibility or fluidity of a particular substance. Both are used as plasticizers for such materials as polyvinyl chloride, film, leathers, electric wires and cables (U.S. Department of Human and Health Services 2000). Only under conditions of high temperatures would these compounds be harmful, as most materials containing it are stable under normal conditions. Diisotridecyl phthalate and Ditridecyl phthalate are thus generally not harmful and poses minimal risk of carcinogenicity.

Diundecyl phthalate (commercially known as JAYFLEX L11P-E) is used in wire and cable insulation for automotive and communications applications (Exxon Mobil Chemical 2000). Because most wires and cables are usually wrapped and unexposed, exposure to the compound is minimal. Diundecyl phthalate poses a negligible threat as a carcinogen.

Closer investigation revealed that from the 10 compounds displaying "Very Significant Agonsim" only 2 actually posed significant threat, given the usage and exposure. Keep in mind the basic principle of toxicology that dose determines toxicity. Certain substances such as anesthesia are used sparingly and infrequently; however substances such as cosmetic products (lotions and shampoos) are used frequently, if not daily. Though the body might not absorb it through the point at which it was administered, these substances are easily transferred from one locale on the body to another and could easily be ingested or inhaled.

Limitations, implications, and future studies

The study used Computational Toxicology principles of investigation. Computational Toxicology modeling relies on statistics and thus shares its limiting factors. The number of compounds used to build the predictive model is completely arbitrary but generally speaking, a greater number of compounds build a better model. More compounds would have built a stronger predictive model. In addition, throughout the process I readjusted model parameters arbitrarily, ultimately striving for an R-squared value of 0.70 or greater. If I had achieved an R-squared greater than 0.7872, the model would have greater predictive power. Most importantly with any software modeling, the prediction can only be as strong as the algorithms the software has been programmed with. Other QSAR software would most likely have yielded different results, given the same parameters, compounds, and data.

Quantity is just as important as quality. Because endocrine disruptors cannot be characterized structurally, it is important to model as many different types of compounds. The reality is that certain types of compounds are more studied than others, increasing the likelihood of these types of compounds being archived and being used for modeling. QSAR models analyze and compare structural characteristics meaning that a greater presence of certain types of compounds could steer the model and skew predictions. The

problem is that not every compound has been archived and translated to SDF and thus cannot be modeled. A greater effort must be made to build these chemical databases, especially considering that thousands of new chemicals are added to the market annually.

It is imperative that compounds be investigated and archived so more endocrine disruptors are uncovered. Breast cancer is one of the most preventable cancers today, however, it would be better to prevent harmful exposure to endocrine disruptors that are linked to the cancer.

Given certain limitations, Computational Toxicology modeling proves useful. In fact its capabilities are vast. Studying other attributes of chemical systems such as binding affinity could help draw new and further conclusions about ER-a. Because of its versatility it is able to investigate various types of chemical attributes, given a specific biological endpoint.

CONCLUSION

From the compounds tested, 7 quats and 3 phthalates had dangerously low EC50 values. Upon research, only 2 quats proved to be of any significant danger. Products containing these compounds should be temporarily banned, as they exhibit strong agonism for ER-a. It is important to be aware of the specificity of the study and not lose sight that other biological attributes should be considered before making any definitive conclusions Quats and Phthalates, especially as they relate to breast cancer.

REFERENCES

Escande, A., A. Pillon, N. Servant, J. Cravedi, F. Larrea, P. Muhn, J. Nicolas, V. Cavaillès, and P. Balaguer. 2006. Evaluation of Ligand Selectivity Using Reporter Cell Lines Stably Expressing Estrogen Receptor Alpha or Beta. Biochemical Pharmacology 71: 1459-1469.

Fitzpatrick, L., T.J. Berrodin, S.F. Jenkins, D.M. Sindoni, D.C. Deecher, and D.E. Frail. Effect of Estrogen Agonists and Antagonists on Induction of Progesterone Receptor in a Rat Hypothalamic Cell Line. Endocrinology 140(9).

Jesperson, Kathy. Endocrine Disruptors-What are they doing to you? On Tap-West Virginia University. Accessed April 19, 2010.

<http://www.nesc.wvu.edu/ndwc/articles/OT/WI03/endocrine_disrupt.html>

Kostelac, D., G. Rechkemmer, and K. Briviba. 2003. Phytoestrogens Modulate Binding Response of Estrogen Receptors r and *â* to the Estrogen Response Element. Journal of Agricultural and Food Chemistry 51: 7632 -7635.

McClure, W. Estrogen Receptor Structures & Functions. 2001. Carnegie Mellon University-Biological Sciences Department. Accessed April 24, 2010.

<http://www.bio.cmu.edu/Courses/BiochemMols/ER/ERIntro.html#overview>

Moein, M.R., S. I. Khan, Z. Ali, S.A.M. Ayatollahi, F. Kobarfard, S. Nasim, M.I. Choudhary, and I.A. Khan. 2008. Flavonoids from *Iris songarica* and their Antioxidant and Estrogenic Activity. Letter 74: 1492 -1495.

Overk, C.R., P. Yao, L.R. Chadwick, D. Nikolic, Y. Sun, M.A. Cuendet, Y. Deng, A. S. Hedayat, G.F. Pauli, N.R. Farnsworth, R.B. van Breemen, and J.L. Bolton. 2005. Comparison of the *in vitro* Estrogenic Activities of Compounds from Hops (*Humulus lupulus*) and Red Clover (*Trifolium pretense*). Journal of Agricultural and Food Chemistry 51: 6246-6253.

Pillon, A., A. Boussioux, A. Escande, S. Ait-Aissa, E. Gomez, H. Fenet, M. Ruff, D. Moras, F. Vignon, M. Duchesne, C. Casellas, J. Nicolas, and P. Balaguer. 2005. Binding of Estrogenic Compounds to Recombinant Estrogen Receptor-α: Application to Environmental Analysis. Environmental Health Perspectives 113(3).

Rahn, O. and W. P. Van Eseltine. 1974. Quaternary Ammonium Compounds**.** Annual Review of Microbiology. 1: 173-192.

Szabo, L. U.S. breast cancer rates drop dramatically. 2006. USA Today. Accessed May 1, 2010. <http://www.usatoday.com/news/health/2006-12-14-breast-cancer_x.htm>

Jayflex Plasticizers. 2000. Exxon Mobil Chemical. Accessed April 26, 2010. <www.exxonmobilchemical.com/OXO_Jayflex_L11P-E_NA_en-FPS.pdf>

NTP-CERHR Expert Panel Report. 2000. U.S. Department of Human and Health Services. p. 17**.**

Questions About Endocrine Disruptors. 2002. ExtoxNet. Accessed April 20, 2010.

<ace.orst.edu/info/extoxnet/faqs/pesticide/endocrine.htm>

DrugInfo: Benzoxonium Chloride. 2004. MIMS USA. Accessed March 20, 2010 <http://www.mimsonline.com/Page.aspx?menuid=mng&name=benzoxonium%20chloride&brief=false#Actions>

Understanding Cancer Series: Estrogen Receptors/SERMs. 2006.

National Cancer Institute. Accessed February 1, 2010. <http://www.cancer.gov/cancertopics/understandingcancer/estrogenreceptors/Slide8>

Chemical Families: Phthalates. 2007. Environmental Working Group. Accessed April 10, 2010. <http://www.ewg.org/chemindex/term/480>

Tox Town: Endocrine Disruptors. 2009. National Library of Medicine. Accessed April 20, 2010. <http://www.toxtown.nlm.nih.gov/text_version/chemicals.php?id=65>

Target: Estrogen Receptor-alpha. The Binding Database. Accessed February 15, 2010. <bindingdb.org>

EDKB: Endocrine Disruptors. 2010. FDA. Accessed April 20, 2010. <http://edkb.fda.gov/endodisrupt.html>

Endocrine Disruptors: What are endocrine disruptors? 2010. U.S. EPA. Accessed April 20, 2010. <http://www.epa.gov/endo/pubs/edspoverview/whatare.htm>

PubChem. 2010. National Center for Biotechnology Information. Accessed January 15, 2010. <http://pubchem.ncbi.nlm.nih.gov>

Category: Quaternary Ammonium Compounds. Wikipedia. Accessed January 5, 2010. <http://en.wikipedia.org/wiki/Category:Quaternary_ammonium_compounds>

Household Products Database. Dimethyldioctadecylammonium chloride. Accessed April 1,2010.<http://householdproducts.nlm.nih.gov/cgi-bin/household/brands?tbl=chem&id=1000&query=Dimethyldioctadecylammonium+chloride+&searchas=TblChemicals>

Appendix A

Training Set Compounds		
Compound	**EC50**	**pEC50**
Lasofoxifene	3.1	-0.49
tetrahydroisoquinoline derivative, 20d	3.1	-0.49
LY2066948 analogue, 17	3.18	-0.5
tetrahydroisoquinoline derivative, 19d	3.3	-0.52
tetrahydroisoquinoline derivative 19b	3.5	-0.54
raloxifene	2.4	-0.38
tetrahydroisoquinoline derivative, 20h	4.5	-0.65
LY2066948 analogue, 9	4.78	-0.68
tetrahydroisoquinoline derivativce, 20e	5.2	-0.72
4-hydrotamoxifen	10.3	-1.01
LY2066948	10.7	-1.03
tetrahydroisoquinoline derivative, 17a	10.7	-1.03
tetrahydroisoquinoline derivative, 20i	16.1	-1.21
tetrahydroisoquinoline derivative, 18a	16.5	-1.22
tetrahydroisoquinoline derv, 19j	19	-1.28
SERBA-1	19.4	-1.29
tetrahydroisoquinoline deriv, 20j	19.8	-1.3
piperazine deriv, 47	31.5	-1.5
serba-2	32.5	-1.51
pip deriv, 30	34.9	-1.54
pip deriv, 38b	46.1	-1.66
pip deriv, 38a	46.6	-1.67
pip deriv, 38e	31	-1.49
Genistein	48	-1.68
benzthiophene compound, 11	48	-1.68
Piperazine Derivative, 33	55	-1.74
Piperazine Derivative, 31	64	-1.81
Pyrazolo[1,5-a]pyrimidine, 12a	90	-1.95

Appendix B

Validation Set Compounds		
Compound	**EC50**	**pEC50**
17-B estradiol	0.22	0.6576
Diethylstilbestrol	0.06	1.2218
tetrahydroisoquinoline derivative, 19f	0.6	0.2218
LY326315	1.14	-0.057
tetrahydroisoquinoline derivative, 26a	1.5	-0.176
tetrahydroisoquinoline derivative, 19c	1.6	-0.204
tetrahydroisoquinoline derivative, 19a	1.8	-0.255
tetrahydroisoquinoline derivative, 20c	1.9	-0.279
Raloxifene	2.4	-0.38
tetrahydroquinoline derivative, 19e	2.6	-0.415
tetrahydroisoquinoline derivative, 20a	2.6	-0.415
tetrahydroisoquinline derivative, 20b	2.9	-0.462
tetrahydroisoquinoline derivative, 19h	2.9	-0.462
tetrahydroisoquinoline derivative, 19g	3	-0.477
3-arylquinazolinethione, 1ba	2793	-3.446
3-arylquinazolinone, 1aa	2918	-3.456
3-arylquinazolinone, 1ae	3112	-3.493
3-arylquinazolinone, 1ak	3143	-3.497
3-arylquinazolinone, 1aas	4261	-3.63
3-arylquinazolinone, 1aan	4521	-3.655
3-arylquinazolinone, 1aao	4624	-3.665
3-arylquinazolinethione, 1bf	4859	-3.687
3-arylquinazolinone, 1ad	5137	-3.711
3-arylquinazolinone, 1ac	5305	-3.725
3-arylquinazolinone, 1am	8317	-3.92
3-arylquinazolinone, 1aae	9268	-3.967
3-arylquinazolinone, 1aw	9317	-3.969

benzthiophene compound, 4	138	-2.14
benzthiophene compound, 5	148	-2.17
Piperidine Derivative, 38c	153	-2.18
benzthiophene compound, 7	162	-2.21
benzthiophene compound, 2	178	-2.25
benzthiophene compound, 6	191	-2.28
benzthiophene compound, 1	209	-2.32
benzthiophene compound, 9	269	-2.43
benzthiophene compound, 8	363	-2.56
Piperazine Derivative, 32	512	-2.71
Tamoxifen	622	-2.79
3-arylquinazolinethione, 1bb	776	-2.89
3-arylquinazolinone, 1aac	983	-2.99
Pyrazolo[1,5-a]pyrimidine, 12b	1000	-3
benzthiophene compound, 10	1000	-3
benzthiophene compound, 12	1000	-3
benzthiophene compound, 13	1000	-3
thiogenistein, 13	1125	-3.05
3-arylquinazolinone, 1ag	1143	-3.06
3-arylquinazolinone, 1ah	1427	-3.15
3-arylquinazolinethione, 1be	1462	-3.16
3-arylquinazolinethione, 1bd	1662	-3.22
3-arylquinazolinone, 1aaj	2392	-3.38
3-arylquinazolinethione, 1bc	2406	-3.38
3-arylquinazolinone, 1ai	4576	-3.66
4,4'-DDE	25600	-4.41
Phenol Red	36900	-4.57
3-hydroxybenzo[a]pyrene	1200	-3.08
9-hydroxybenzo[a]pyrene	700	-2.85
Benzo[a]pyrene	100	-2
Estradiol	0.017	1.77
Estriol	0.16	0.796
Biochanin A	82	-1.91
Coumestrol	16	-1.2
Daidzein	150	-2.18
Ethynylestradiol	0.008	2.097
Zuclomiphene	100	-2
Mestranol	15.8	-1.2
Estradiol benzoate	3.98	-0.6
Tibolone	105.1	-2.02
Isoxanthohumol	1100	-3.04
3-arylquinazolinone, 1ao	9344	-3.971
3-arylquinazolinone, 1at	10082	-4.004
Equol	3500	-3.544
Alpha-Zearalanol	0.135	0.8697
Zearalenone	0.313	0.5045
Delta-5-androstenediol	13.1	-1.117
2,4'-DDE	2740	-3.438