HHS Public Access

Author manuscript

Chemosphere. Author manuscript; available in PMC 2021 June 17.

Published in final edited form as:

Chemosphere. 2016 February ; 144: 2238–2246. doi:10.1016/j.chemosphere.2015.11.004.

3D QSAR studies of hydroxylated polychlorinated biphenyls as potential xenoestrogens

Patricia Ruiz[a,*], **Kundan Ingale**[b], **John S. Wheeler**[a], **Moiz Mumtaz**[a]

[a]Computational Toxicology and Methods Development Laboratory, Division of Toxicology and Human Health Sciences, Agency for Toxic Substances and Disease Registry, 1600 Clifton Road, MS-F57, Atlanta, GA, 30333, USA

[b]VLife Sciences Tech. Pvt. Ltd., Plot No-05, Survey No 131/1b/2/11, Ram Indu Park, Baner Road, Pune, 411045, India

Abstract

Mono-hydroxylated polychlorinated biphenyls (OH-PCBs) are found in human biological samples and lack of data on their potential estrogenic activity has been a source of concern. We have extended our previous *in silico* 2D QSAR study through the application of advance techniques such as docking and 3D QSAR to gain insights into their estrogen receptor (ERα) binding. The results support our earlier findings that the hydroxyl group is the most important feature on the compounds; its position, orientation and surroundings in the structure are influential for the binding of OH-PCBs to ERα. This study has also revealed the following additional interactions that influence estrogenicity of these chemicals (a) the aromatic interactions of the biphenyl moieties with the receptor, (b) hydrogen bonding interactions of the p-hydroxyl group with key amino acids ARG394 and GLU353, (c) low or no electronegative substitution at para-positions of the p-hydroxyl group, (d) enhanced electrostatic interactions at the meta position on the B ring, and (e) co-planarity of the hydroxyl group on the A ring. In combination the 2D and 3D QSAR approaches have led us to the support conclusion that the hydroxyl group is the most important feature on the OH-PCB influencing the binding to estrogen receptors, and have enhanced our understanding of the mechanistic details of estrogenicity of this class of chemicals. Such *in silico* computational methods could serve as useful tools in risk assessment of chemicals.

Keywords

2D QSAR; 3D QSAR; OH-PCBs; In silico modeling; kNN MFA

1. Introduction

Human exposure to the environmentally persistent polychlorinated biphenyls (PCBs) as documented by human biomonitoring, mostly as mixtures, is widespread (ATSDR, 2000; ATSDR, 2012). The potential impact on public health from individual PCB congeners, their

*Corresponding author. pruiz@cdc.gov (P. Ruiz).

Conflicts of interest

The authors declare no conflict of interest.

metabolites and/or their mixtures is a concern because these chemicals are known to be toxic to laboratory animals and wildlife (ATSDR, 2000; ATSDR, 2012; Ruiz et al., 2013; Machala et al., 2004; Kamata et al., 2009; Gregoraszczuk et al., 2008; Arulmozhiraja et al., 2005). They are also extremely persistent leading to potential exposure risk for many years at PCB contaminated sites. Mixtures of PCBs in the environment contain a variety of the 209 individual PCB congeners, each congener having its own characteristics regarding toxicity, chemistry and environmental fate. Defining toxicity of specific PCB mixtures in the environment continues to challenge toxicologists, risk assessors, and public health professionals because of the limitless number of mixtures, the changing mixture profile and the difficulty in assessing the risks posed by the individual congeners, their metabolites and the potential for interactions between them (Ruiz et al., 2013).

Co-planar and non-coplanar PCBs are metabolized *in vivo* to hydroxyl and sulfur containing metabolites (Takeuchi et al., 2011; ATSDR, 2000; Grimm et al., 2015). As many as 837 mono-hydroxylated PCB congeners (OH-PCBs) whose molecular structure encompass 1 to 9 chlorine atoms, can be found in PCB mixtures (Arulmozhiraja et al., 2005; Grimm et al., 2015). Such metabolites and their mixtures have been found in several tissues and are known to exert potential adverse effects during critical windows of embryonic and fetal development (Guvenius et al., 2002; Soechitram et al., 2004). They are known to cause their effects through various mechanisms, but the main postulated mechanism has been estrogenic or antiestrogenic activity in mammals (ATSDR, 2000; ATSDR, 2012; Arulmozhiraja et al., 2005). Thus, exposure to the OH-PCBs continue to attract attention because of their potential endocrine disruption activity (Park et al., 2009; Antunes-Fernandes et al., 2011; Machala et al., 2004; Grimm et al., 2015; Ruiz et al., 2013).

Unfortunately, major gaps still exist in the traditional toxicological databases of these chemicals, hence adverse health effects from OH-PCBs exposure in humans are poorly characterized and understood. To help fill these missing data gaps, a growing number of *in silico* tools and QSAR models are being developed to identify and evaluate appropriate surrogates for these “data-poor” chemicals (Tunkel et al., 2005; Nigsch et al., 2009; Mekenyan et al., 2003; Oberg, 2006). Through these efforts, chemical structural features can be identified that may lead to the development of robust predictive models and specific selection criteria for use of such models (Matthews and Contrera, 2007; Liu et al., 2006; Lill et al., 2005; Cassano et al., 2010; Cronin, 2002; Netzeva et al., 2005; OECD, 2004; Mombelli and Devillers, 2010).

In our previous study, we evaluated the published estrogenic activity of OH-PCBs reported using two-hybrid assays yeast cells containing the human estrogen receptor ERα (Arulmozhiraja et al., 2005). We developed a 2D QSAR prediction model based on molecular descriptors for a data set of 71 OH-PCBs covering a range of molecular properties The results showed that the position of the hydroxyl substitution, polarizability, and *meta* adjacent unsubstituted carbon pairs at the phenolic ring contribute towards greater estrogenic activity for these chemicals (Ruiz et al., 2013).

The principal objectives of this study are to: (i) use the x-ray crystal structure of the estrogen receptor (ERα) to evaluate binding patterns of the OH-PCBs with respect to human estrogen

receptor; (ii) obtain consensus features from the QSAR and binding/docking studies that influence the binding; and (iii) model quantitatively the relationship between the molecular descriptors and estrogenicity activity of OH-PCBs by using 3-D QSAR to determine predictivity and identify the most important molecular descriptors influencing the estrogenicity of OH-PCBs.

2. Materials and methods

2.1. Data set

We are extending our previous 2D QSAR study using the estrogenic activity of OH-PCBs (Ruiz et al., 2013). These data set, as described before, consisted of estrogenic agonist activities of OH-PCBs (Table 1) that were measured with a reporter gene assay using a yeast strain (*Saccharomyces cerevisiae* Y190) that incorporates the human estrogen receptor ERα and the coactivator TIF2 (Ruiz et al., 2013; Arulmozhiraja et al., 2005). The estrogenic activity of active OH-PCBs was recorded as REC20 (the concentration of test compound showing 20% of the activity of 10^{-10} M estradiol (E2)) and percentage activity relative to estradiol ((calculated from REC20 E2/REC20 test compound)* 100), the remaining OH-PCBs were listed as non active. Non-active chemicals were defined with values equivalent to the lowest REC20 activity plus two standard deviations, or the percentage activity relative to estradiol divided by two (score of 3.7). For modeling purposes, OH-PCB activity was converted to negative logarithm values. Thus, the active OH-PCBs had scores <3.7, while the non-active had equal or > than 3.7. The compounds used for the current study are as shown in Table 1.

2.2. In silico modeling

2.2.1. Geometry optimization—The 3D QSAR studies of 71 hydroxylated PCB congeners were carried out using VLife molecular design suite (VLifeMDS, 2010) software version 4.3 (MDS 4.3). This suite allows flexibility to study molecular fragments of interest and screen molecules for specific activities. Three-dimensional structures of OH-PCBs were constructed and their geometries optimized to make the conformations with least potential energy using Merck Molecular Force Field (MMFF) and MMFF charge for the atom using a distance-dependent dielectric constant of 1.0 and convergence criteria (rms gradient) of 0.001 kcal/mol.

2.2.2. Docking studies—Toxicology and drug discovery processes require a rapid assessment or screening of small molecules that potentially bind to biological receptors. Towards this computer models are being used for virtual screening of large chemical data bases. If the 3-D structure or model of the target is available techniques such as docking studies are employed via a computer generated representation of a small molecule into a target structure like an active site of the enzyme or a receptor in a variety of positions, conformations and orientations.

For our docking studies, we used the Protein database (PDB) of Estrogen receptor (PDB ID: 3ERT) “Human Estrogen Receptor alpha ligand-binding domain in complex with 4-hydroxytamoxifen.” We started by cleaning the 3-D structure of the receptor (adding or

deleting atoms or substructures) and used the Generalized Rigid PLP based (GRIP) docking using VLife MDS 4.3, to dock the OH-PCBs in the receptor. A flexible ligand approach with Piecewise Linear Pairwise Potential (PLP) scoring function was applied for docking using 15° as the angle of rotation at grid points. GRIP docking employs the PLP scoring function in a novel way for fast and accurate capturing and prediction of ligand–receptor interactions in the active site of proteins. GRIP docking methodology innovatively incorporates the PLP scoring function that includes ligand–receptor interactions of hydrogen bonding (donor–acceptor), repulsions (donor–donor, acceptor–acceptor) and dispersion (involving non-polar group interactions) types (Ajmani et al., 2009).

2.2.3. Alignment of molecules—All molecules in the data set were aligned by a template-based method where a template of the estrogen receptor is built by including common substructures in the series. The most highly bioactive energetically stable conformation in the series is chosen as a reference molecule on which other molecules are aligned.

2.2.4. 3D QSAR—Aligned molecules were subjected to generation of a common rectangular grid. The steric and electrostatic and hydrophobic interaction energies were computed at the lattice points of the grid using a methyl probe of charge +1. These interaction energy values were considered for relationship generation and utilized as descriptors to decide nearness between molecules. The term descriptor is utilized in the following discussion to indicate field values at the lattice points.

The compounds were divided in optimal training and test sets by using the sphere exclusion algorithm which accounts for the descriptor space of the dataset for the distribution. In the present study, molecular field analysis coupled with k-nearest neighbor (kNN) was applied to obtain 3D QSAR models. The calculated steric, electrostatic and hydrophobic field descriptors were used as independent variables and -log ER values were used as dependent variables to derive the 3D QSAR models.

2.2.5. Elimination of outliers—In the course of obtaining a reasonable 3D QSAR model which can provide a robust interpretation of interactions of the OH-PCBs with the estrogen receptor, we systematically eliminated outliers from the dataset. In a skewed data set, like the one in the current study, outliers are carefully evaluated as visual outliers might be in the right tail of the skewed data. We eliminated one molecule at a time and generate new QSAR models using remaining molecules in the training set. The model whose predictive ability drastically improves after elimination of a particular molecule (outlier) was selected as the best model.

3. Results and discussion

3.1. Docking studies

The docking scores for the 71 OH-PCBs in the binding site using the estrogen receptor ranged between −54.64 and −34.8. These docking scores were compared with the score for the co-crystal ligand 4-hydroxytamoxifen (−101.57). Thus the obtained docking scores suggest that the OH-PCBs have fair binding affinity towards the ER receptor. The

interactions of co-crystal ligand 4-hydroxy tamoxifen are as shown in Fig. 1a, the key interactions between 4-hydroxy tamoxifen and the human estrogen receptor can be noted as a hydrogen bonding interaction at GLU353 and charge interactions with ASP351 along with hydrophobic interaction at LEU346, THR347, ALA350, MET388, LEU391, MET421, ILE 424, LEU428, LEU 525.

The docking poses of the OH-PCBs were analyzed with respect to interactions and alignment towards the co-crystal ligand. The co-crystal and the top scoring poses overlay in the binding site (Fig. 1b). The red stick co-crystal ligand and the best docking pose of the compound # 58 (2,2′,4′,6′-tetrachlorobiphenyl-4-ol) in the green ball and stick have good overlay in the binding site. The best docked pose shows hydrogen bonding interaction with GLU353 as in Fig. 1c.

The best docked pose is compound # 58 which is the most active molecule from the dataset (ER activity = −0.23); the presence of key interactions with GLU353supports the high estrogenic binding of compound # 58. The docking results of individual interactions of the compounds along with docking scores provide a justification for their activity. As compared with the 2D QSAR models, it can be seen that compound #58 lacks *meta* chlorine atom and has only one para hydroxy group. This observation is in accordance with the results of the 2D QSAR and 3D QSAR studies suggesting that having electrostatic group (chloro) at *meta* position on the phenyl ring is conducive for toxicity. Table 1 indicates the best docking scores for the 71 compounds under study along with key interacting residues.

All the compounds show aromatic Pi stacking interaction with TRP383. Further analysis of the most active molecules and their best binding poses suggest that most of the compounds have strong hydrogen bond interactions with GLU353 and ARG394; these H-bonding interactions clearly define a salient distinction between the active and inactive compounds. Results suggest that the presence of hydrogen bonding interactions of hydroxyls group on the OH-PCBs potentiate the estrogenic activity. The overlay of best poses of compounds #52, #56, #58, # 59 and #63 are shown in Fig. 2. The analysis of key interactions of the molecules suggests that the aromatic interactions have major contributions towards the binding of the OH-PCBs with the receptor, while the H-bond interactions with the receptor potentiate the estrogenic activity of the compounds.

3.2. 3D QSAR

3D QSAR models for the estrogenic activity were generated for the dataset of 71 molecules. Different models were generated using kNN approaches for model building, due to skewness in the data the generated models show higher standard errors. To estimate outliers in the dataset, we eliminated one by one the training set molecules with residuals more than 1 log unit. If the elimination of a compound improved the prediction accuracy and resulted in significant reductions in the standard error of estimate, it was considered an outlier. We found that elimination of compound #51 enhanced the cross validation correlation coefficient, prediction accuracy, and reduced the standard error significantly. The generated QSAR model and the best model were reported in Table 2. Further elimination of compounds was not performed in order to avoid over-fitting of the models and increasing the probability of chance correlation.

The models clearly indicate that eliminating compound 51 from the training set enhanced the model accuracy, predictive ability, and reduced the errors of estimate. The change in model descriptors after elimination of compound 51 suggests that most probably the QSAR model 1 was a chance correlated model.

The kNN model describes the optimum structural feature for the estrogenicity activity of the 49 molecules training set and 21 molecules test set (Table 2). The compounds in the training and test set are indicated in Table 3. The model is able to classify 82% of variation in the biological activity of molecules in the training set and 84% of the molecules in the test set. Low standard errors of estimate for q^2 and pred_r^2 suggest the goodness of the model. The terms E_15, E_186, E_407, E_303, S_95 and E_152 are the electrostatic and steric field energy of interactions between probe (CH_3^{+1}) and compounds. The selected terms can be visualized around the molecules with optimum ranges in Fig. 3.

The fitness plot using the actual activity on the X-axis and predicted activity on the Y-axis can be seen in Fig. 4a. The radar plot highlighting the difference between actual and predicted activities for the molecules in the training and test sets are shown in Fig. 4b and c respectively. The plots shows actual activity (red line) and the predicted activity (blue line), and provide information on model behavior regarding the prediction of active and inactive molecules. The active molecules are the data points close to the center and vice versa. The plots suggest that the model captures features that separate active and inactives in both training and test sets. The actual and predicted activities along with calculated residuals are as given in Table 3.

The terms selected in the QSAR model are electrostatic interactions at grid point 15, 186, 407, 303, and 152 along with steric interaction at grid point 95. The optimum ranges of the interactions suggest that the electrostatic interactions at grid points 186, 303 and steric interactions at 95 are favorable while electrostatic interactions at 407 and 152 are completely unfavorable for estrogenic activity of the compounds (Fig. 3).

The electrostatic interactions at grid point E_15 and corresponding *meta*-position on the phenyl ring A are partly favorable to estrogenic activity of the compounds. The descriptor range suggests not only that the presence of highly electronegative groups at the meta position will enhance the electrostatic interactions and render the compound toxic, but that low electronegative groups will also have toxic effects. The analysis can be verified from the fact that the top active compounds #52, #58 and #59 are unsubstituted at the meta-position of the A phenyl ring. The grid points 186 and 303 in the model are highly favorable for estrogenic activity, which corresponds to the meta-position on the B ring (Fig. 3). The optimum range for the interactions for both points suggest that adding more electronegative substitution at the B ring renders compounds more toxic.

The electrostatic interactions at grid point 407 suggest that the electrostatic interactions around the B ring at ortho and *meta* positions are unfavorable for estrogenic activity. The grid point E_152 surrounding the para position on ring A (substituted mostly by hydroxyl groups) suggest that electrostatic interactions are unfavorable for estrogenic activity. A closer look at placement of the grid point suggests that the orientation of the hydroxyl

hydrogen is affecting the H-bonding interaction and is also captured by the grid point 152. The conformational analysis of docking results and the position of 3D QSAR grid point suggest that a coplanar hydroxyl group on a phenolic ring is a key feature of the OH-PCBs binding to the receptor. The only steric term in the model S_95 contributes favorably for the estrogenic activity. This term corresponds to the *ortho* and *meta* positions on the ring A suggesting that steric groups at ortho positions are favorable while the meta-positions have electrostatic constraints.

The results from these docking and 3D QSAR studies are in agreement with our previous 2D QSAR study of the estrogenic activity of these OH-PCBs. The 2D results showed that the position of the hydroxyl substitution, polarizability, and *meta* adjacent unsubstituted carbon pairs at the phenolic ring contribute towards greater estrogenic activity for these chemicals. The docking results combined with 3D QSAR results also suggest similar features; the presence of a *para* hydroxyl group with *meta* un-substituted/low electronegative substitutions is favorable for estrogenic activity. In addition to our earlier results, the following additional key interactions have been highlighted throughout the current study that enhance our understanding of the estrogenic activity of the OH-PCBs:

- Aromatic interactions of the biphenyl moieties with the ER receptor.
- Hydrogen bonding interactions of the *p*-hydroxyl group with key amino acid ARG394 and GLU353,
- Presence of no or low electronegative substitution at *para*-positions of the *p*-hydroxyl group.
- Enhanced electrostatic interactions at *meta* position on the B ring.
- Co-planarity of the hydroxyl group on the A ring.

The current study is based on the yeast assay, and the model predictions are specific to the assay type. However a statistical correlation can be established between predictions of different assay types to extend and extrapolate the prediction of this model to other assay types (see also Ruiz et al., 2013). Some models can be generalized to different assays depending on the closeness of the assays and their hierarchical position within the toxicity pathway. The performance of a QSAR model to predict other chemicals depends on the structural similarity of the chemicals in the database used to develop the model. The applicability domain of these models defines their advantages and limitations. Our current QSAR model can be used to predict the estrogenic activity of an external chemical set (other metabolites such as di-hydroxy PCBs, hydroquinone or benzoquinone), and the prediction confidence evaluated based on if they are within the applicability domain of the model. Thus, the model predictions will be reliable if they meet the preset criteria. However, if the preset criteria are not met the model can be fine-tuned by including these new metabolites in the database and making them part of the training and test data set of the model. QSAR model development is a reiterative process wherein new experimental data help create models and the improved models, in turn, guide experimental research to fill data gaps. These dynamic processes are therefore followed until the model developers and users are satisfied. Since an individual model cannot be considered to possess complete applicability domain or total accuracy, consensus or comparative QSAR analysis, as we present here, is a

logical approach for employing several models and thus broadened the applicability of such models.

In conclusion, this study has integrated robust docking and 3D QSAR models to develop a novel approach for evaluation of the estrogenicity of OH-PCBs that lack experimental toxicity data.

The availability of crystal structure for the ER enabled the application of docking methods. Using the conventional docking methods, the docked compounds reveal key interactions between the OH-PCBs and the ER receptor. This comparative 3D QSAR analysis that describes the relationship between the structural interactions and estrogenic activity of OH-PCBs was performed using the kNN method for model building. The consensus of docking and 3D QSAR models have provided valuable insights of additional molecular features that influence the estrogenic activity of OH-PCBs. The docking and 3D QSAR results were in good agreement with experimental results. The consensus suggests that the hydroxyl group is the most important feature on the compound. Its position, orientation and surroundings in the structure are influential for the binding of OH-PCBs to ER. The results obtained from the current study also provide information on additional structural features on the OH-PCBs that render these compounds toxic.

Laboratory measured values are needed to develop *in silico* models such as QSAR models and are chosen over estimated values. In silico modeling is an evidence-based predictive method being evaluated by regulatory agencies for risk assessment and supports scientific decision making. These methods offer a rapid and cost-effective first-pass screening capability to assess toxicity when conventional toxicology data are limited or lacking, particularly to identify compounds for further testing. Recent laws such as the European Union's REACH regulation, support the acceptance of these methods and their use by the regulatory and public health communities in the mitigation of potential hazardous exposures that could compromise the quality of human and environment health (Worth et al., 2007; OECD, 2005; OECD, 2006). Due to the large number of OH-PCBs and the limited understanding of their dose response relationship for risk assessment, research addressing these data gaps are required. This can be achieved through strategic and critically designed laboratory and experimental studies. Consequently, the data generated will be used for the development of QSAR models which are required to have an appropriate range of similar structures and toxicity endpoints. Modeling and experimental approaches complement each other and help develop accurate tools for risk assessment of these chemicals and other emerging chemicals that lack toxicity data.

Acknowledgments

The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Agency for Toxic Substances and Disease Registry. Mention of trade names is not an endorsement of any commercial product.

References

Agency for Toxic Substances and Disease Registry, 2000. ATSDR, Toxicological Profile for Polychlorinated Biphenyls, U.S. Department of Health and Human Services. Public Health Services, Atlanta, USA.

Agency for Toxic Substances and Disease Registry, 2012. ATSDR, Addendum Toxicological Profile for Polychlorinated Biphenyls, U.S. Department of Health and Human Services. Public Health Services, Atlanta, USA.

Ajmani S, Karanam S, Kulkarni SA, 2009. Rationalizing protein–ligand interactions for PTP1B inhibitors using computational methods. Chem. Biol. Drug Des 582–595. [PubMed: 19824894]

Antunes-Fernandes EC, Bovee TFH, Daamen FEJ, Helsdingen RJ, Van Den Berg M, Van Duursen MBM, 2011. Some OH-PCBs are more potent inhibitors of aromatase activity and (anti-) glucocorticoids than non-dioxin like (NDL)-PCBs and $MeSO_2$-PCBs. Toxicol. Lett 206, 158–165. [PubMed: 21782008]

Arulmozhiraja S, Shiraishi F, Okumura T, Iida M, Takigami H, Edmonds JS, Morita M, 2005. Structural requirements for the interaction of 91 hydroxylated polychlorinated biphenyls with estrogen and thyroid hormone receptors. Toxicol. Sci 84, 49–62. [PubMed: 15601674]

Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M, Bigoni D, Benfenati E, 2010. CAESAR models for developmental toxicity. Chem. Cent. J 4 (Suppl. 1).

Cronin MTD, 2002. The current status and future applicability of quantitative structure-activity relationships (QSARs) in predicting toxicity. Altern. Lab. Anim 30 (Suppl. 2), 81–84. [PubMed: 12513655]

Gregoraszczuk EL, Rak A, Ludewig G, Gasinska A, 2008. Effects of estradiol, PCB3, and their hydroxylated metabolites on proliferation, cell cycle, and apoptosis of human breast cancer cells. Environ. Toxicol. Pharmacol 25, 227–233. [PubMed: 21783862]

Grimm FA, Hu D, Kania-Korwel I, Lehmler HJ, Ludewig G, Hornbuckle KC, et al., 2015. Metabolism and metabolites of polychlorinated biphenyls. Crit. Rev. Toxicol 45, 245–272. [PubMed: 25629923]

Guvenius DM, Hassanzadeh P, Bergman A, Noren K, 2002. Metabolites of polychlorinated biphenyls in human liver and adipose tissue. Environ. Toxicol. Chem 21, 2264–2269. [PubMed: 12389902]

Kamata R, Shiraishi F, Nakajima D, Takigami H, Shiraishi H, 2009. Mono-hydroxylated polychlorinated biphenyls are potent aryl hydrocarbon receptor ligands in recombinant yeast cells. Toxicol. Vitro 23, 736–743.

Lill MA, Dobler M, Vedani A, 2005. In silico prediction of receptor-mediated environmental toxic phenomena-application to endocrine disruption. Sar. QSAR Environ. Res 16, 149–169. [PubMed: 15844448]

Liu H, Papa E, Gramatica P, 2006. QSAR prediction of estrogen activity for a large set of diverse chemicals under the guidance of OECD principles. Chem. Res. Toxicol 19, 1540–1548. [PubMed: 17112243]

Machala M, Blaha L, Lehmler H-J, Pliskova M, Majkova Z, Kapplova P, Sovadinova I, Vondracek J, Malmberg T, Robertson LW, 2004. Toxicity of hydroxylated and quinoid PCB metabolites: inhibition of gap junctional intercellular communication and activation of aryl hydrocarbon and estrogen receptors in hepatic and mammary cells. Chem. Res. Toxicol 17, 340–347. [PubMed: 15025504]

Matthews EJ, Contrera JF, 2007. In silico approaches to explore toxicity end points: issues and concerns for estimating human health effects. Expert Opin. Drug Metab. Toxicol 3, 125–134. [PubMed: 17269899]

Mekenyan O, Dimitrov S, Schmieder P, Veith G, 2003. In silico modelling of hazard endpoints: current problems and perspectives. Sar. QSAR Environ. Res. 14, 361–371. [PubMed: 14758980]

Mombelli E, Devillers J, 2010. Evaluation of the OECD (Q)SAR application toolbox and toxtree for predicting and profiling the carcinogenic potential of chemicals. Sar. QSAR Environ. Res 21, 731–752. [PubMed: 21120759]

Netzeva TI, Worth A, Aldenberg T, Benigni R, Cronin MTD, Gramatica P, Jaworska JS, Kahn S, Klopman G, Marchant CA, Myatt G, Nikolova-Jeliazkova N, Patlewicz GY, Perkins R, Roberts D, Schultz T, Stanton DW, Van De Sandt JJM, Tong W, Veith G, Yang C, 2005. Current status of

methods for defining the applicability domain of (quantitative) structure-activity relationships. The report and recommendations of ECVAM workshop 52. Altern. Lab. Anim 33, 155–173. [PubMed: 16180989]

Nigsch F, Macaluso NJM, Mitchell JBO, Zmuidinavicius D, 2009. Computational toxicology: an overview of the sources of data and of modelling methods. Expert Opin. Drug. Metab. Toxicol 5, 1–14. [PubMed: 19236225]

Oberg T, 2006. Virtual screening for environmental pollutants: structure-activity relationships applied to a database of industrial chemicals. Environ. Toxicol. Chem 25, 1178–1183. [PubMed: 16629159]

OECD, 2005. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment, Series on Testing and Assessment, Paris, France.

OECD, 2006. Report on the Regulatory Uses and Applications in OECD Member Countries of (Quantitative) Structure-activity Relationship [(Q)SAR] Models in the Assessment of New and Existing Chemicals, Series on Testing and Assessment. OECD, Paris, France.

OECD, 2004. Report from the Expert Group on (Quantitative) Structure-activity Relationships [(Q)SARs] on the Principles for the Validation of (Q)SARs, Series on Testing and Assessment, Paris, France.

Park H-Y, Park J-S, Sovcikova E, Kocan A, Linderholm L, Bergman A, Trnovec T, Hertz-Picciotto I, 2009. Exposure to hydroxylated polychlorinated biphenyls (OH-PCBs) in the prenatal period and subsequent neurodevelopment in eastern Slovakia. Environ. Health Perspect. 117, 1600–1606. [PubMed: 20019912]

Ruiz P, Myshkin E, Quigley P, Faroon O, Wheeler JS, Mumtaz MM, Brennan RJ, 2013. Assessment of hydroxylated metabolites of polychlorinated biphenyls as potential xenoestrogens: a QSAR comparative analysis*. Sar. QSAR Environ. Res 24, 393–416. [PubMed: 23557136]

Soechitram SD, Athanasiadou M, Hovander L, Bergman A, Sauer PJJ, 2004. Fetal exposure to PCBs and their hydroxylated metabolites in a Dutch cohort. Environ. Health Perspect 112, 1208–1212. [PubMed: 15289169]

Takeuchi S, Shiraishi F, Kitamura S, Kuroki H, Jin K, Kojima H, 2011. Characterization of steroid hormone receptor activities in 100 hydroxylated polychlorinated biphenyls, including congeners identified in humans. Toxicology 289, 112–121. [PubMed: 21843587]

Tunkel J, Mayo K, Austin C, Hickerson A, Howard P, 2005. Practical considerations on the use of predictive models for regulatory purposes. Environ. Sci. Technol 39, 2188–2199. [PubMed: 15871254]

VLifeMDS, 2010. Molecular Design Suite. VLife Sciences Technologies Pvt. Ltd., Pune, India. www.vlifesciences.com.

Worth AP, Bassan A, De Bruijn J, Gallegos Saliner A, Netzeva T, Patlewicz G, Pavan M, Tsakovska I, Eisenreich S, 2007. The role of the European Chemicals Bureau in promoting the regulatory use of (Q)SAR methods. Sar. QSAR Environ. Res 18, 111–125. [PubMed: 17365963]

HIGHLIGHTS

- We extended our previous 2D-QSAR study by applying docking and 3D-QSAR modeling.
- 3D QSAR models are presented and validated for transparency and reliability.
- 3D QSAR modeling results are in agreement with our previous 2D-QSAR study.
- Additional key interactions have been highlighted throughout the current study.
- These results enhance our understanding of the estrogenic activity of the OH-PCBs.

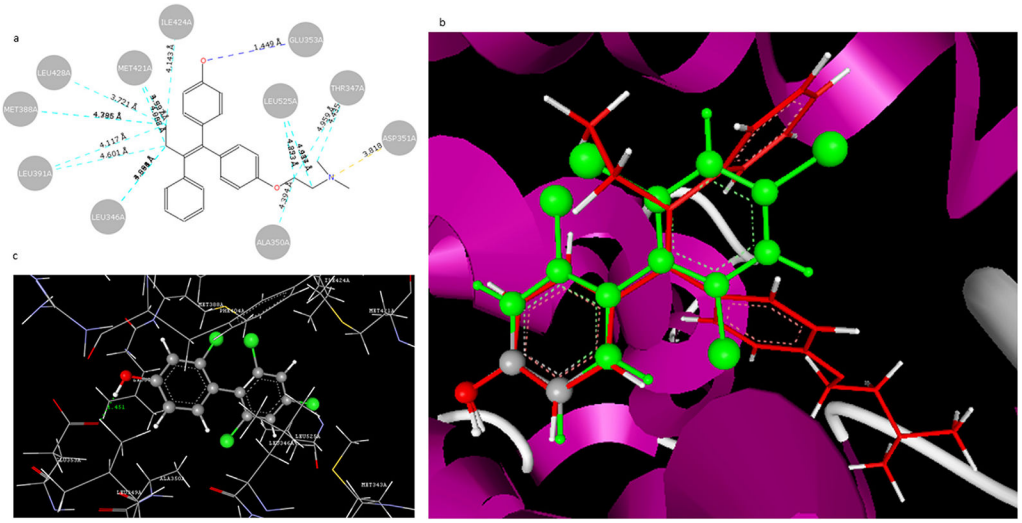

Fig. 1.

(a) The 2D representation of interaction between co-crystal tamoxifen and the estrogen receptor (ER). (b) Overlay of the co-crystal ligand (Red stick) and best docking pose (Green Ball and stick). (c) Key Aromatic Pi stacking Interactions of best docked compound with the receptor.

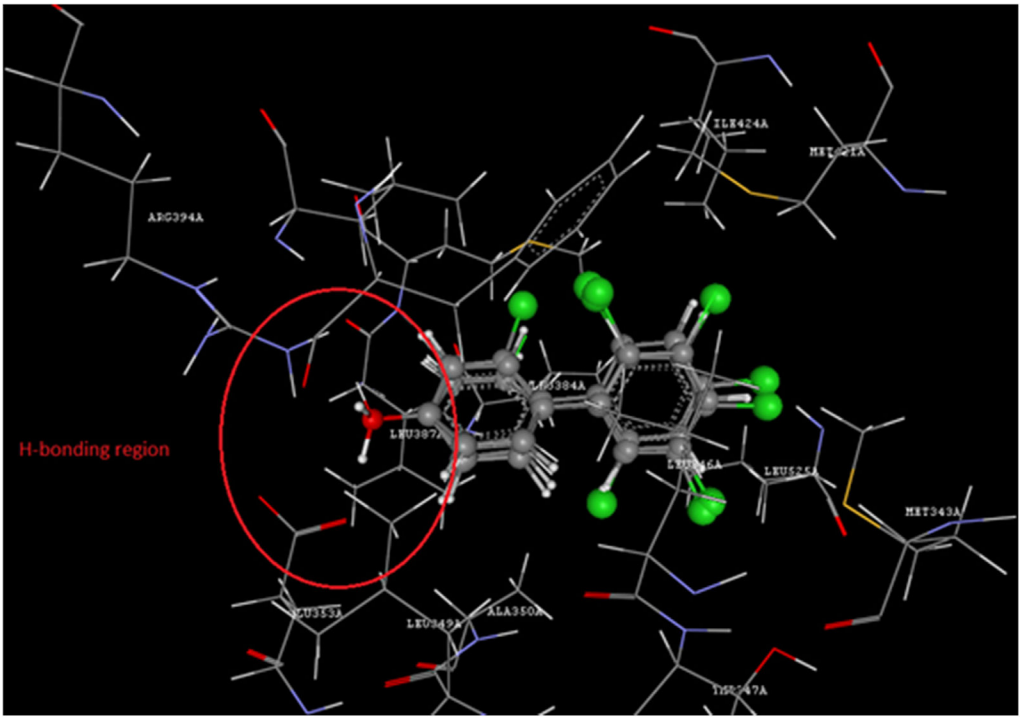

Fig. 2.
Overlay of best docking poses of active ligands 56, 58 and 59 highlighting H-bonding interactions with TRP72.

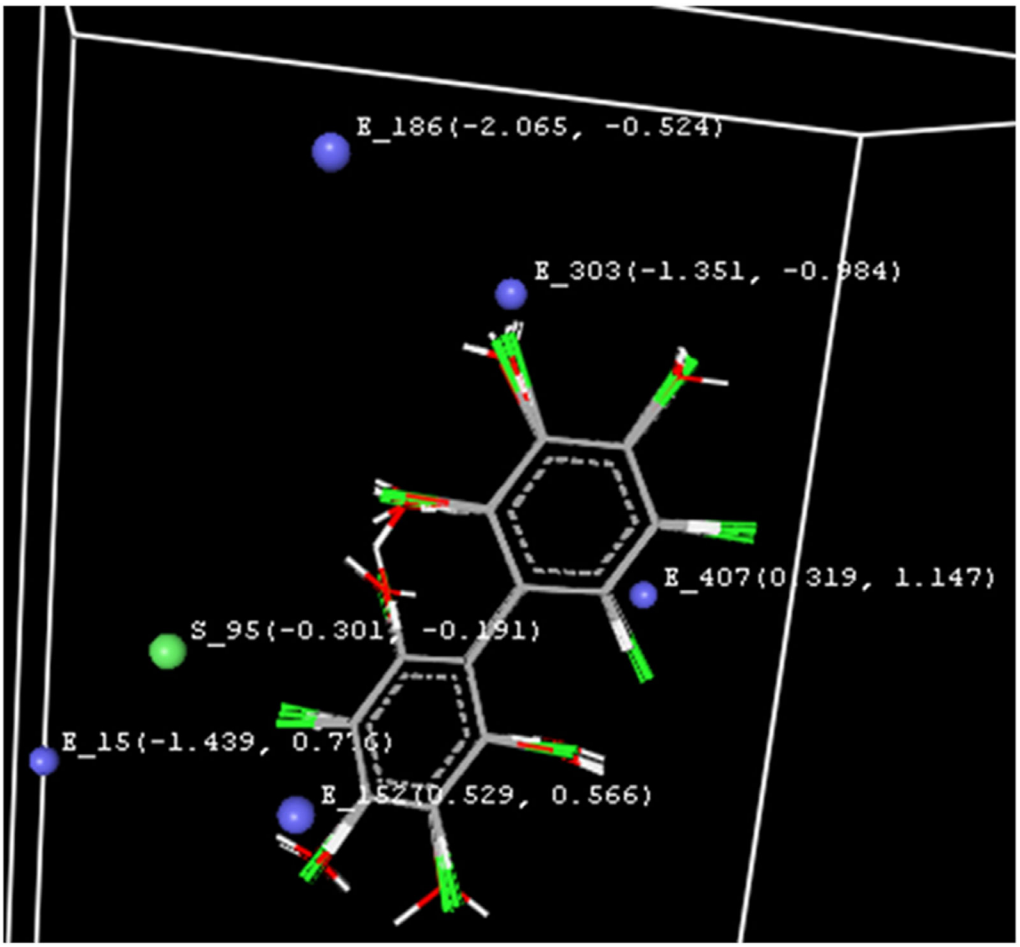

Fig. 3.
Selected hydrophobic and electrostatic Grid points in the 3D QSAR model.

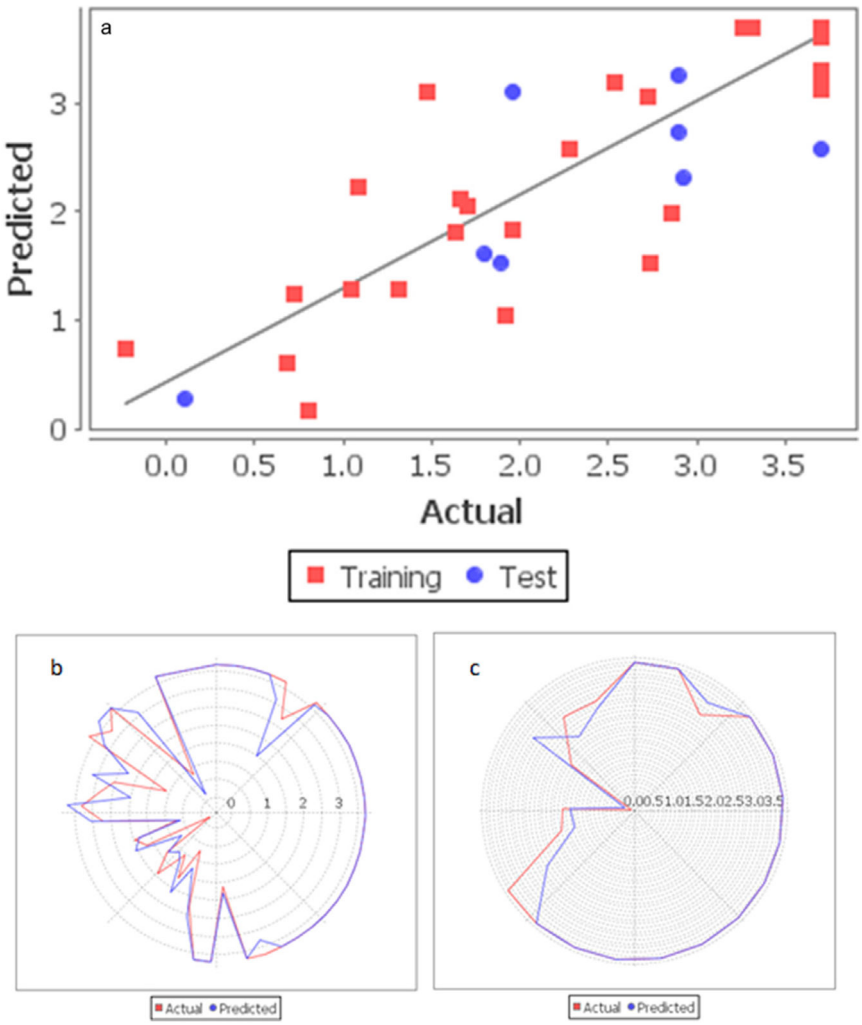

Fig. 4.

(a) Predicted versus actual estrogenicity activity as measured in the yeast two hybrid bioassays (Arulmozhiraja et al., 2005) of OH-PCBs using the 3D-QSAR model. (b) Radar plot showing fitness of predicted and actual activity values of training set. (c) Radar plot exploring fitness of predicted and actual activity values of test set.

Table 1

Chemical name and estrogenicity activity of hydroxylated polychlorinated biphenyls under the current study, and Piecewise Linear Pairwise Potential (PLP) score with key interacting residues.

Molecule number	IUPAC NAME	*ER_pIC$_{50}$	PLP score	Interactions	
				H-bond	Pi stacking
1	2′,3′,4′,5′-tetrachlorobiphenyl-2-ol	3.7	−45.56	–	TRP383
2	2′,3′,4′,5,5′-pentachlorobiphenyl-2-ol	3.7	−43.46	–	TRP383
3	2′,3′,5′,6′-tetrachlorobiphenyl-2-ol	3.7	−42.78	–	TRP383
4	2′,3′,5′,6′-tetrachlorobiphenyl-3-ol	3.7	−43.56	–	TRP383
5	2′,3′,5,5′,6′-pentachlorobiphenyl-2-ol	3.7	−37.65	–	TRP383
6	2′,3,3′,4′,5′-pentachlorobiphenyl-4-ol	3.7	−41.93	–	TRP383
7	2′,3,3′,4′,5,5′-hexachlorobiphenyl-4-ol	3.7	−36.45	–	TRP383
8	2′,3,3′,4-tetrachlorobiphenyl-2-ol	3.7	−34.8	–	TRP383
9	2′,3,3′,5′,6′-pentachlorobiphenyl-4-ol	3.7	−39.06	–	TRP383
10	2′,3,3′,5,5′,6′-hexachlorobiphenyl-4-ol	3.7	−38.52	–	TRP383
11	2′,3,4′,5,6′-pentachlorobiphenyl-ol	3.7	−38.83	–	TRP383
12	2′,3,4′,5-tetrachlorobiphenyl-2-ol	3.7	−40.12	–	TRP383
13	2′,3,4′-trichlorobiphenyl-2-ol	3.7	−35.8	–	–
14	2′,3,5,5′-tetrachlorobiphenyl-4-ol	3.7	−45.47	–	TRP383
15	2′,4′,5,5′-tetrachlorobiphenyl-2-ol	3.7	−41.34	–	TRP383
16	2′,3,3′-trichlorobiphenyl-2-ol	3.26	−39.46	MET522	TRP383
17	2′,4′,6′-trichlorobiphenyl-2-ol	3.31	−43.36	–	–
18	2′,3′-dichlorobiphenyl-2-ol	2.89	−40.47	MET343	–
19	2,2′,3′,5′,6′-pentachlorobiphenyl-4-ol	3.7	−35.4	–	TRP383
20	2,2′,3′,6-tetrachlorobiphenyl-3-ol	3.7	−39.26	–	TRP383
21	2,2′,4,4′-tetrachlorobiphenyl-3-ol	3.7	−44.67	–	–
22	2,2′,5′,6-tetrachlorobiphenyl-3-ol	3.7	−39.1	–	TRP383
23	2,3′,4′,5,6-pentachlorobiphenyl-3-ol	3.7	−43.68	–	TRP383
24	2,3′,4,4′,6-pentachlorobiphenyl-3-ol	3.7	−39.51	–	TRP383
25	2,3′,4,4′-tetrachlorobiphenyl-3-ol	3.7	−38.53	–	TRP383
26	2,3′,5,5′-tetrachlorobiphenyl-4-ol	3.7	−42.12	–	TRP383

Molecule number	IUPAC NAME	*ER_pIC$_{50}$	PLP score	Interactions	
				H-bond	Pi stacking
27	2′,4,4′,5-tetrachlorobiphenyl-3-ol	3.7	−43.48	–	TRP383
28	3′,4′,5-trichlorobiphenyl-2-ol	3.7	−49.06	–	TRP383
29	3′,4′-dichlorobiphenyl-2-ol	3.7	−42.96	–	TRP383
30	3′,4,4′,6-tetrachlorobiphenyl-3-ol	3.7	−44.44	–	TRP383
31	3′,4,5′,6-tetrachlorobiphenyl-3-ol	3.7	−44.73	–	TRP383
32	3′,4,5′-trichlorobiphenyl-2-ol	3.7	−39.87	–	TRP383
33	3′,4,6-trichlorobiphenyl-2-ol	3.7	−37.56	–	TRP383
34	3′,5,5′-trichlorobiphenyl-2-ol	3.7	−41.56	–	TRP383
35	2,2′,5′-trichlorobiphenyl-3-ol	1.47	−45.52	–	TRP383
36	2′,4′,6′-trichlorobiphenyl-3-ol	1.7	−41.23	–	TRP383
37	2′,3′,4′,5′-tetrachlorobiphenyl-3-ol	2.28	−34.79	–	TRP383
38	2′,5′-dichlorobiphenyl-3-ol	1.96	−40.43	–	TRP383
39	2,2′,3′,4,6-pentachlorobiphenyl-3-ol	2.54	−31.27	–	TRP383
40	3,3′,4′,5,6-pentachlorobiphenyl-2-ol	3.7	−43.28	–	TRP383
41	3,3′,4′,5-tetrachlorobiphenyl-2-ol	3.7	−44.48	–	TRP383
42	3,3′,4′,5-tetrachlorobiphenyl-4-ol	3.7	−42.72	–	TRP383
43	3,3′,4′,6-tetrachlorobiphenyl-2-ol	3.7	−45.94	–	TRP383
44	3,3′,4′-trichlorobiphenyl-2-ol	3.7	−44.08	–	TRP383
45	3,3′,5′-trichlorobiphenyl-ol	3.7	−39.48	–	TRP383
46	3,3′,5,5′-tetrachlorobiphenyl-2-ol	3.7	−43.16	–	TRP383
47	3,3′,5′-trichlorobiphenyl-2-ol	3.7	−38.6	–	TRP383
48	3,3′,6-trichlorobiphenyl-2-ol	3.7	−40.17	–	TRP383
49	3,4′,5-trichlorobiphenyl-2-ol	3.7	−44.87	–	TRP383
50	3,4′,6-trichlorobiphenyl-2-ol	3.7	−44.29	–	TRP383
51	3,5-dichlorobiphenyl-4-ol	3.7	−41.46	–	TRP383
52	2,2′,5′-trichlorobiphenyl-ol	0.8	−49.89	GLU353	TRP383
53	2′,3,3′-trichlorobiphenyl-ol	1.89	−38.12	MET343	–
54	2′,3,4′-trichlorobiphenyl-ol	1.92	−45.41	–	TRP383
55	2′,3,5′-trichlorobiphenyl-ol	1.8	−48.61	–	–
56	2′,4′,6′-trichlorobiphenyl-4-ol	0.1	−50.85	ARG394, GLU353	–

Molecule number	IUPAC NAME	*ER_pIC$_{50}$	PLP score	Interactions	
				H-bond	Pi stacking
57	2,2′,3′,5-tetrachlorobiphenyl-4-ol	2.85	−43.12	–	–
58	2,2′,4′,6′-tetrachlorobiphenyl-4-ol	−0.23	−54.64	GLU353	TRP383
59	2′,3′,4′,5′-tetrachlorobiphenyl-4-ol	0.72	−50.58	GLU353	TRP383
60	2,3,5,6-tetrachlorobiphenyl-4-ol	1.64	−32.33	–	TRP383
61	2′,3,4′,6′-tetrachlorobiphenyl-4-ol	1.31	−43.37	–	TRP383
62	2,2′,3′,4′,5′-pentachlorobiphenyl-4-ol	1.66	−42.59	–	TRP383
63	2′,5′-dichlorobiphenyl-4-ol	0.68	−49.05	ARG394, GLU353	TRP383
64	2-chlorobiphenyl-4-ol	1.09	−36.47	–	TRP383
65	3-chlorobiphenyl-4-ol	2.92	−37.09	–	TRP383
66	2,3′,5-trichlorobiphenyl-4-ol	2.89	−43.35	–	TRP383
67	2,4′,5-trichlorobiphenyl-4-ol	2.74	−35.9	–	TRP383
68	2,3′,4′-trichlorobiphenyl-ol	1.04	−49.05	–	TRP383
69	5-chlorobiphenyl-2-ol	3.7	−39.69	–	TRP383
70	3′,4′,6-trichlorobiphenyl-3-ol	1.96	−45.11	–	TRP383
71	3′,5′,6-trichlorobiphenyl-3-ol	2.72	−46.31	–	TRP383

*The estrogenic activity of active OH_PCBs was recorded as REC20 (the concentration of test compound showing 20% of the activity of 10^{-10} M estradiol (E2)) and percentage activity relative to estradiol ((calculated from REC20 E2/REC20 test compound)*100), the remaining OH-PCBs were listed as non-active (Arulmozhiraja et al., 2005). Non active chemicals were defined with values equivalent to the lowest REC20 activity plus two standard deviations, or the percentage activity relative to estradiol divided by two (score of 3.7). For modeling purposes, OH-PCBs activity was converted to negative logarithm values. Thus, the active OH-PCBs had scores <3.7, while the non-active had equal or > than 3.7. Residues and nomenclature: Alanine (ALA), Cysteine (CYS), Aspartic Acid (ASP), Glutamic Acid (GLU), Phenylalanine (PHE), Glycine (GLY), Histidine (HIS), Isoleucine (ILE), Lysine (LYS), Leucine (LEU), Methionine (MET), Asparagine (ASN), Proline (PRO), Glutamine (GLN), Arginine (ARG), Serine (SER), Threonine (THR), Valine (VAL), Tryptophan (TRP), Tyrosine (TYR).

Table 2

Model statistics for the generated 3D QSAR models.

Method	N_{tr}/N_{ts}*	kNN	q^2	Pred_r^2	q^2_SE	Pred_r^2_SE	Grid points	kNN range	
								Min	Max
kNN_1	50/21	2	0.77	0.58	0.61	0.63	E_152	−0.087	−0.056
							S_342	−0.097	−0.057
							S_172	0.462	0.684
							E_259	0.136	4.819
							E_141	−0.5267	2.408
kNN_2	49/21	2	0.82	0.83	0.47	0.40	E_15	−1.439	0.776
							E_186	−2.065	−0.524
							E_407	0.319	1.147
							E_303	−1.351	−0.984
							S_95	−0.301	−0.191
							E_152	0.529	0.566

*Ntr: number of molecules in training set and Nts: Number of molecules in test set, kNN: k-nearest neighbor, q2: Cross validated Correlation coefficient, Pred_r^2: Correlation Coefficient for test/predicted set, q2_SE: Standard error of estimate for q2, Pred_r^2: Standard error of estimate for Pred_r2, E:electrostatic molecular descriptor, S:steric molecular descriptor.

Table 3

Actual and predicted estrogenic activity along with residuals.

Molecule[a]	Actual activity	Model 1	Model 2	Residual model 1	Residual model 2
1	3.70	3.70	3.70	0.00	0.00
2£	3.70	3.70	3.70	0.00	0.00
3	3.70	3.70	3.70	0.00	0.00
4£	3.70	3.70	3.70	0.00	0.00
5£	3.70	3.70	3.70	0.00	0.00
6£	3.70	3.70	3.70	0.00	0.00
7	3.70	3.64	3.70	0.06	0.00
8	3.70	3.70	3.70	0.00	0.00
9£	3.70	3.70	3.70	0.00	0.00
10	3.70	3.66	3.70	0.04	0.00
11	3.70	3.41	3.70	0.29	0.00
12	3.70	3.70	3.70	0.00	0.00
13	3.70	3.12	3.21	0.58	0.49
14	3.70	3.65	3.70	0.05	0.00
15	3.70	3.70	3.70	0.00	0.00
16	3.26	3.09	3.70	0.17	−0.44
17	3.31	3.05	3.70	0.26	−0.39
18£	2.89	2.37	2.73	0.52	0.16
19	3.70	3.70	3.70	0.00	0.00
20	3.70	3.70	3.70	0.00	0.00
21£	3.70	2.80	3.70	0.90	0.00
22	3.70	3.62	3.70	0.08	0.00
23£	3.70	3.70	3.70	0.00	0.00
24	3.70	3.70	3.70	0.00	0.00
25	3.70	3.70	3.70	0.00	0.00
26	3.70	3.29	3.13	0.41	0.57
27	3.70	3.10	3.70	0.60	0.00
28	3.70	3.62	3.70	0.08	0.00

Molecule[a]	Actual activity	Model 1	Model 2	Residual model 1	Residual model 2
29	3.70	3.63	3.70	0.07	0.00
30£	3.70	3.27	3.70	0.43	0.00
31	3.70	3.67	3.70	0.03	0.00
32	3.70	3.61	3.70	0.09	0.00
33	3.70	3.57	3.70	0.13	0.00
34£	3.70	3.70	3.70	0.00	0.00
35	1.47	3.53	3.10	−2.06	−1.63
36	1.70	2.08	2.04	−0.38	−0.34
37	2.28	1.98	2.56	0.30	−0.28
38	1.96	1.97	1.82	−0.01	0.14
39	2.54	3.52	3.18	−0.98	−0.64
40	3.70	3.70	3.70	0.00	0.00
41£	3.70	3.56	3.70	0.14	0.00
42	3.70	3.70	3.61	0.00	0.09
43	3.70	3.66	3.70	0.04	0.00
44	3.70	3.70	3.29	0.00	0.41
45	3.70	3.70	3.70	0.00	0.00
46£	3.70	3.70	3.70	0.00	0.00
47£	3.70	3.37	2.57	0.33	1.13
48£	3.70	3.70	3.70	0.00	0.00
49	3.70	2.68	3.70	1.02	0.00
50£	3.70	1.89	3.70	1.81	0.00
51[¥]	3.70	2.68	–	1.02	–
52	0.80	1.08	0.17	−0.28	0.63
53£	1.89	1.48	1.53	0.41	0.36
54	1.92	1.05	1.05	0.87	0.87
55£	1.80	1.47	1.60	0.33	0.20
56£	0.10	0.91	0.27	−0.81	−0.17
57	2.85	1.59	1.98	1.26	0.87
58	−0.23	0.98	0.73	−1.21	−0.96
59	0.72	0.79	1.24	−0.07	−0.52

Molecule[a]	Actual activity	Model 1	Model 2	Residual model 1	Residual model 2
60	1.64	2.11	1.82	−0.47	−0.18
61	1.31	1.44	1.27	−0.13	0.04
62	1.66	1.02	2.12	0.64	−0.46
63	0.68	0.71	0.60	−0.03	0.08
64	1.09	2.13	2.23	−1.04	−1.14
65£	2.92	3.51	2.30	−0.59	0.62
66£	2.89	2.85	3.25	0.04	−0.36
67	2.74	1.77	1.51	0.97	1.23
68	1.04	0.93	1.28	0.11	−0.24
69	3.70	3.70	3.70	0.00	0.00
70£	1.96	3.44	3.10	−1.48	−1.14
71	2.72	3.70	3.06	−0.98	−0.34

¥ outlier eliminated from 3D QSAR Model 2.

[a] The molecules succeeded by "£" are the molecules in the test set.