Journal of Hazardous Materials 480 (2024) 136188

Contents lists available at ScienceDirect

Journal of Hazardous Materials

journal homepage: www.elsevier.com/locate/jhazmat

In silico models for the screening of human transthyretin disruptors

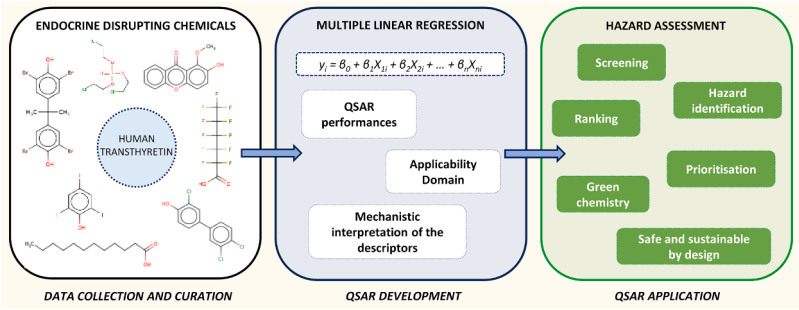
Marco Evangelista, Nicola Chirico, Ester Papa *

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Theoretical and Applied Sciences, University of Insubria, via J.H. Dunant 3, 21100 Varese, Italy

HIGHLIGHTS

- New datasets of T4-hTTR competing potencies of chemicals were generated.
- Three new robust QSAR models were developed based on simple molecular descriptors.
- The models were validated through internal and external validation procedures.
- Mechanistic interpretation highlighted relevant structural features.
- Models to support the Thyroid Hormone System-Disrupting Chemicals hazard assessment.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
QSAR
New approach methodologies
Thyroid hormone-system disruption
Hazard assessment
Structural features

ABSTRACT

The use of New Approach Methodologies (NAMs), such as Quantitative Structure-Activity Relationship (QSAR) models, is highly recommended by international regulations to speed up hazard and risk assessment of Endocrine Disruptors, which are known to be linked to a wide spectrum of severe diseases on humans and wildlife. A very sensitive target for these chemicals is the thyroid hormone system, which plays a key role in regulating metabolic and cognitive functions. Several chemicals have been demonstrated to compete with the thyroid hormone thyroxine (T4) for binding to human thyroid hormone distributor protein transthyretin (hTTR). In this work, we generated three new datasets composed by T4-hTTR competing potencies of more than 200 heterogeneous chemicals measured by three different *in vitro* assays. These datasets were used for the development of new regression QSAR models. The best models were thoroughly validated by internal and external validation procedures. The mechanistic interpretation of the selected molecular descriptors provided information on structural features which are relevant to characterise hTTR binders, such as the presence of hydroxylated and halogenated aromatic rings. PCA analysis was used to rank the studied chemicals according to their increasing T4-hTTR competing potency. Hydroxylated and halogenated bicyclic aromatic compounds are ranked as the strongest hTTR binders. The new QSARs are useful to screen potential Thyroid Hormone System-Disrupting Chemicals (THSDCs), and to support the identification of sustainable alternatives to hazardous chemicals.

* Correspondence to: Department of Theoretical and Applied Sciences, University of Insubria, via J.H. Dunant 3, 21100 Varese, Italy.
E-mail addresses: mevangelista@uninsubria.it (M. Evangelista), nicola.chirico@uninsubria.it (N. Chirico), ester.papa@uninsubria.it (E. Papa).

https://doi.org/10.1016/j.jhazmat.2024.136188
Received 21 July 2024; Received in revised form 28 September 2024; Accepted 15 October 2024
Available online 16 October 2024

1. Introduction

The concern about Endocrine Disrupting Chemicals (EDCs) is significantly increasing worldwide, owing to spreading evidences of endocrine-related disorders both in human and wildlife, such as reduction of fertility, cancers, metabolic and developmental dysfunctions [1,2]. According to the World Health Organization International Programme on Chemical Safety (WHO/IPCS), an endocrine disruptor is defined as "*an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations*" [3]. The risk assessment of EDCs is extremely challenging [4,5] due to their structural heterogeneity and large variety of physicochemical properties and environmental behaviours, their multiple emission sources, and the complexity of the endocrine system [3]. These features, on the one hand, lead to the ubiquitous presence of EDCs in the environmental compartments, increasing the number and complexity of the potential exposure pathways, while, on the other hand, determine several modes of action and toxicity pathways by which those chemicals exert their adverse effects on living organisms [6]. In Europe, several political and scientific efforts followed the Community Strategy for Endocrine Disruptors [7], aiming to fill knowledge gaps, to set validated test methods and to harmonize the current chemical legislations [8]. A priority of the recent European Green Deal [9] is to improve the protection of human health and of the environment against harmful chemicals, moving towards a toxic-free environment through a Chemical Strategy for Sustainability (CSS) [10], by strengthening and harmonizing the European regulatory system, with a significant focus on endocrine disruptors such as in the recent Classification, Labelling and Packaging (CLP) Regulation update [11].

Among EDCs, a variety of industrial chemicals, pesticides, plasticizers, surfactants, pharmaceuticals and personal care products [12-15] have been recognised to interfere with multiple molecular targets involved in the Hypothalamic-Pituitary-Thyroid (HPT) axis through several mechanisms, altering the Thyroid Hormones (THs) homeostasis [16,17]. Alterations of the TH homeostasis may induce severe adverse diseases to nervous system development, metabolism, immune and cardiovascular systems [18,19]. Substances able to exert this type of activity are known as Thyroid Hormone System-Disrupting Chemicals (THSDCs) [20,21]. Triiodothyronine (T3) and thyroxine (T4) represent the two main hormones produced by the thyroid gland, from where THs are transported into the bloodstream by the activity of THs transporters (e.g., monocarboxylate transporter 8 or MCT8) [20,22]. In mammals, T4 is the predominant TH form secreted by the thyroid, while T3 is the more biologically active one, due to its higher binding affinity with Thyroid Hormone Nuclear Receptors (TRs). T4 is often referred to as a "prohormone", as it is converted to T3 in peripheral tissues through deiodination catalysed by deiodinase enzymes [23,24]. Nearly all the amount of THs in human blood is bound to distributor proteins, such as transthyretin (TTR), thyroxine-binding globulin (TBG), and albumin (ALB) [25]. Differently from TBG and ALB, which are primarily synthetised by the liver, TTR is also synthetised by the choroid plexus, by the retinal pigment epithelium of the eye, and by other tissues and/or organs, implying that TTR is involved in multiple biological processes [26,27]. Among these different functions, TTR has the role to buffer abnormal alterations of free THs physiological concentrations. Free THs are devoted to reach the target tissues to guarantee the proper functioning of the TH system [28]. Any alteration in the THs binding to distributor proteins can modulate the free THs physiological concentrations in blood, leading to TH system dysfunctions. Particular attention is posed in studying the ability of THSDCs to displace the thyroid hormone T4 from the TTR. Differently from the other TH distributor proteins, TTR has a key role in delivering T4 in cerebrospinal fluid [18] across the blood-brain barrier and placenta during foetal development: exogenous chemicals able to bind to TTR could be transferred to the foetus, leading to a T4 deficiency and subsequent severe, irreversible effects, primarily cognitive dysfunctions [19,29]. For these reasons, the identification of chemicals with this behaviour is considered a priority [20,22].

In the literature, several chemical classes of compounds have been experimentally demonstrated to compete with T4 for binding to the human TTR (hTTR), such as per- and polyfluoroalkyl substances (PFAS) [30], bisphenols [31], halogenated phenols [32], polychlorinated biphenyls (PCBs) [33], polybrominated diphenyl ethers (PBDEs) [34] and both PCBs and PBDEs hydroxylated metabolites [35,36]. The traditional approach based on *in vivo* experiments became inadequate and unsustainable for testing the potential endocrine properties of all the existing substances, and so for the identification of potential THSDCs [20,37]. In particular, the development of New Approach Methodologies (NAMs), such as *in vitro* test (e.g. *in vitro* assays to detect estrogen [38] and androgen [39] receptors agonists and antagonists), and *in silico* methods (e.g. Quantitative Structure-Activity Relationships (QSARs)), is highly recommended in the literature, and by several authorities, to support the identification of EDCs, as well as of THSDCs [8,20,37,40–42]. Worldwide, several efforts and projects aimed for an integration of the results in the Weight of Evidence (WoE) approach [43-46]. Focusing on the prediction of the hTTR binding potencies of chemicals, a limited number of regression QSARs have been developed so far using techniques such as Multiple Linear Regression (MLR) [32,47–50], k-Nearest Neighbors (k-NN) [50], and 3D QSAR approaches [51,52]. Most of these QSARs are based on small datasets (between 15 and 32 chemicals) mainly representative for specific chemical classes (such as PBDEs, perfluorinated compounds (PFCs), halogenated phenols and thiophenols). Among the aforementioned models, only those developed by Yang and coworkers [50] have a wide applicability domain in terms of structural heterogeneity represented in their training sets. Statistical performances of the existing regression models range, in terms of fitting (R^2), between 0.810 and 0.960. However, with the exclusion of the models by Yang and coworkers [50], the structural and response domains of these QSARs is very narrow due to the limited size of the respective training sets. Furthermore, the use of commercial software for the calculation of the molecular descriptors, such as the Dragon software, may limit their application.

This study aims to develop new regression QSAR models, following the OECD guidelines for QSAR development [53], for the identification of potential hTTR binders. This work takes into account current research shortcomings [41,54] and in particular the need for new QSAR models for endocrine disruption modalities different than androgen and estrogen (e.g. here thyroid hormone system disruption), based on curated datasets and with broad applicability domains. To this end, the QSARs proposed in this work were developed using three new and curated datasets based on data collected from the literature, which include values of T4-hTTR competing potencies for more than 200 different chemicals, expressed as Relative competitive Potency (RP), and calculated from experimental binding affinities measured using three *in vitro* assays (the 8-anilino-1-naphthalenesulfonic acid (ANSA) based binding assay [20], the fluorescence conjugate isothiocyanate (FITC)-T4 based binding assay [20], and the radiolabeled [125I]-T4 binding assay (RLBA) [34]). The development of models, based on these datasets, extends the experimental and chemical information contained in previous QSAR studies for the same endpoint. Furthermore, since the three datasets cover a wide range of different chemicals, each QSAR is applicable to fill the experimental data gaps for ANSA, FITC-T4, and RLBA assays. The experimental and predicted RP values for more than 200 chemicals were used to screen potential hTTR binders using the QSAR models developed for the different *in vitro* assays, to evaluate their agreement. A final important aspect that was considered in this study was the need to ease the reproducibility and the application of the new QSARs by implementing them in a non-commercial software.

2. Material and methods

2.1. Datasets

The raw dataset reported in the Supplementary Material S1 lists binding affinity (commonly quantified in terms of IC50, i.e. half-maximal inhibitory concentration) or RP values measured for 240 different chemicals, mostly aromatic and halogenated. This collection includes, to the best of our knowledge, all the binding affinity or RP values currently available in the literature [29,30,32–36,55–86]. RP is defined as the ratio between the binding affinity of T4 and the binding affinity of a chemical with hTTR. When only binding affinity values were available in the original literature, their RP values were calculated and included in the raw dataset (Supplementary Material S1). The use of RP values to define the hTTR binding affinity reduces the data bias caused by minor modifications of the procedures within the same *in vitro* assays, and by their measurements in different laboratories [84]. Data for mixtures, ambiguous molecular structures, and unclear chemical identifiers, were not included in the raw dataset. Moreover, in order to avoid further reduction of the already limited structural and experimental information, two chiral forms of the same chemical (hexabromocyclodecane alpha and beta) were converted in the non-chiral form, while structures of salts and anions (11 chemicals) were converted to their respective acid, as in a previous work [87]. Geometric mean of RP values was calculated when multiple experimental data were available for one chemical [88].

Three datasets (i.e. ANSA, FITC-T4, and RLBA), reported in Supplementary Material S2, were extracted from the raw dataset to develop QSAR models for the prediction of T4-hTTR competing potencies, according to different *in vitro* assays explained as follows. The radiolabeled [125I]-T4 binding assay (RLBA), firstly described by Somack et al. [89], and further elaborated by minor modifications [34,35], uses radioactive iodine-125 (125I) to label T4 in order to make T4 detectable. A gamma counter is used to measure the radioactivity of the radiolabeled-T4 bound with TTR. Variation in the measured signal in the presence of a competitor is used to quantify the relative binding [34,35,89]. The 8-anilino-1-naphthalenesulfonic acid (ANSA) based binding assay (firstly described by Nilsson et al. [90], and further elaborated by minor modifications [55,71]), and the fluorescence conjugate isothiocyanate (FITC)-T4 based binding assay (firstly described by Smith et al. [91], and further elaborated by minor modifications [36,72]) are two competitive fluorescence displacement assays: based on different principles, the TTR binding of chemicals is quantified by measuring the degree of reduction of the fluorescence signal [20]. Particularly, competitive fluorescence displacement assays have been highlighted as powerful methods to detect chemicals able to interfere with TH distributor proteins TTR and TBG [84], and have been recently used in the Joint Research Centre (JRC)'s EURL ECVAM thyroid validation study for the validation of methods for THSDCs identification [20].

The raw dataset also includes 35 binding affinities measured with Surface Plasmon Resonance-based Bioassay (SPRB) [69] and Isothermal Titration Calorimetry (ITC) [85], reported in Supplementary Materials S1. However, these data were too limited for modelling purposes and were excluded from further analysis.

Dataset ANSA includes 79 compounds, such as hydroxylated PBDEs, hydroxylated PCBs, sulfated PCBs, phosphates, fatty acids, halogenated phenols and thiophenols, and halogenated and/or hydroxylated benzoic acids. Dataset FITC-T4 is composed by 50 compounds, such as PFAS, bisphenols, hydroxylated PBDEs, PCBs, hydroxylated PCBs, sulfated PCBs, parabens, phthalates, and halogenated phenols. Dataset RLBA includes 137 compounds, such as PBDEs, hydroxylated PBDEs, PCBs, hydroxylated PCBs, PFAS, halogenated phenols, bisphenol A derivatives, and xanthones. RP values were log transformed prior to QSAR modelling. Finally, all the experimental LogRP values from the ANSA, FITC-T4, and RLBA datasets, were combined into Dataset I, which includes 223 chemicals. Missing experimental LogRPs were filled by applying the respective assay-specific QSAR model developed in this work. This procedure led to the development of Dataset II, which includes 3 assay-specific columns of experimental and, where not available, predicted LogRP values. Finally, only predicted LogRPs within the range of the experimental response of the respective training set, with a prediction uncertainty below the maximum uncertainty predicted for the respective training set, and calculated for chemicals within the structural applicability domain of each model, were included into Dataset III (150 unique chemicals). Principal Component Analysis (PCA) [92] was applied to evaluate the relationships between the LogRP values within Dataset III. Dataset I, II, and III, are reported in Supplementary Material S2.

2.2. Molecular descriptors

The chemical structures were coded as SMILES (Simplified Molecular Input Line Entry System), which were downloaded from the PubChem website (available at https://pubchem.ncbi.nlm.nih.gov/) or calculated by drawing the chemical structures in the ACD/ChemSketch software v. 2021.1.1 (available at https://www.acdlabs.com/, ACDLabs, Toronto, ON, Canada). SMILES were harmonized using the Open Babel software v. 2.4.1 (available at https://openbabel.org/) [93] to generate unique strings for the same molecular structure, and were used as input by the PaDEL-Descriptor software v. 2.21 [94] for the calculation of 7185 mono- /bi- dimensional theoretical molecular descriptors and fingerprints, which encode for the structural features of the compounds. Molecular descriptors were filtered using the QSARINS software [95] prior to modelling, in order to reduce useless and redundant information. Descriptors with at least one missing value, as well as those characterized by low variance (i.e., the same value for more than 80 % of the molecules) or a pairwise correlation exceeding 95 %, were excluded from the further analysis. At the end of this procedure, 546, 483 and 503 molecular descriptors were selected for the datasets i.e., ANSA, FITC-T4 and RLBA, respectively.

2.3. Splitting

To estimate the ability of the QSAR models to make reliable predictions, each dataset was split in a training set for the model development, and in a test set for the external validation. Chemicals were sorted by response, then one every three chemicals were assigned to the test set, and the remaining chemicals were assigned to the training set. The first and the last chemicals were assigned to the training set, in order to limit the endpoint experimental space of the model and to reduce extrapolation. Finally, each training set was further pre-filtered using QSARINS by removing low variance and highly correlated descriptors, as described in Section 2.2.

2.4. Modelling procedure

QSARs were developed using MLR by means of Ordinary Least Squares (OLS), and several statistical metrics were calculated to verify their goodness-of-fit (Coefficient of Determination, R^2, and Root Mean Square Error of the training set, $RMSE_{TR}$), internal robustness (leave-one-out, Q^2_{LOO}, and leave-more-out, Q^2_{LMO}, which are cross validation metrics), and external predictive ability ($RMSE_{PR}$, Q^2_{F3}). Details about the calculation of these metrics are reported in Supplementary Material S3.

The descriptors selection procedure applied to the filtered descriptors (see Section 2.3) explored, in a first step, all their combinations (all-subset) up to two descriptors; in a second step, a Genetic Algorithm (GA) was applied to explore the most promising combinations of three or more descriptors [95]. Q^2_{LOO} was used as the fitness function both for all-subset and GA.

Moreover, to further check for the QSAR models robustness, both the QUIK rule [96] and the Y-scrambling procedures [97] were applied. The

QUIK rule was performed to ensure that the correlations among the modelling descriptors is lower than their correlation with the response (in this work, models were filtered according to $K_{XY} - K_X > 0.015$ [96]). To check for chance correlation between the molecular descriptors and the response of the developed QSARs, 2000 QSARs were developed using training sets with random shuffled responses (Y-scrambling). The average of R^2 and Q^2_{LOO} (called R^2_{YS} and Q^2_{YS}, respectively) of the 2000 QSARs should be much lower than the ones of the QSARs under scrutiny, since the structure-activity relationship of the scrambled QSARs should be negligible.

Calculations concerning the QSARs and the corresponding performances were performed using the QSARINS software [95].

2.5. Applicability domain

The definition of the Applicability Domain (AD) of a QSAR model is required to evaluate the reliability of predictions and/or the degree of extrapolation. Each QSAR model was developed using the structural and the experimental information included in the respective training set, which defines the space of the AD. Concerning the chemical structures, the AD of each model was defined by the leverage approach, with the graphical support of the Williams plot. In particular, this charts plots the standardized residuals of predictions (a measure of the response AD) on the y-axis, allowing for the identification of response outliers (i.e., chemicals with standardized residuals that fall outside the range defined by ±2.5 standard deviation units), and the leverage values (HAT matrix diagonal elements) of the chemicals on the x-axis, allowing for the identification of structural outliers (i.e., chemicals with a leverage value greater than h^*, defined as $3\,^*(p+1)/n$, where p is the number of model descriptors and n is the number of the chemicals in the training set). The HAT matrix, also known as leverage or influence matrix, is calculated as

$$HAT = X(X^TX)^{-1}X^T \quad (1)$$

where X is the data matrix consisting of n rows and p columns.

The leverage value of each chemical is a measure of its distance from the centroid of the model. It quantifies the influence of each chemical on the model and the reliability of the predictions. Predictions associated to hat values larger than h^* (defined above) are considered outside the structural AD of the model and, therefore, are less reliable than those falling within the h^* cut off value. The inclusion of predictions for new chemicals within the AD of the here proposed models can be verified in the QSAR-ME Profiler beta version 1.02 software available at https://dunant.dista.uninsubria.it/qsar/.

To further evaluate the QSAR models, the normality of the distribution of the residuals was graphically inspected through the QQ plots (quantile–quantile plots).

2.6. Prediction of missing LogRPs in Dataset I

The prediction of missing experimental LogRP values of chemicals included in Dataset I was performed by applying the QSAR models reported in Eqs. 3, 5 and 7, as described in the Section 2.1, using the QSAR-ME Profiler beta version 1.02 software (available at https://dunant.dista.uninsubria.it/qsar/) leading to Dataset II. This software allowed for the evaluation of the reliability of predictions by identifying the chemicals within the applicability domains, and by comparing the uncertainty of predictions with the range of the uncertainty in the training sets. This procedure led to the development of Dataset III, as described in Section 2.1. Predicted LogRPs, hat values, and uncertainties values, are reported in Supplementary Material S2. Additional details on the calculation of uncertainties in the training and in the test set are reported in Supplementary Material S3.

3. Results and discussion

The ANSA, FITC-T4, and RLBA datasets, were modelled following the principles proposed by the OECD for the regulatory acceptability of QSARs [53]. A population of about 100 QSAR models was generated for each dataset, as the result of the GA variables subset selection procedure applied in QSARINS. The maximum number of descriptors in each population was kept as low as possible, to reduce the possibility of overfitting, and in agreement to the parsimony principle. The ratio "number of chemicals/number of molecular descriptors", calculated for the best models developed from the three datasets, was about 15, which is largely above the suggested regulatory threshold of five [53]. The best model from each population was selected on the basis of the best balance between fitting and measures of internal cross validation, taking into account both the applicability domain and the residuals distribution of the predicted endpoints. The plot of experimental versus predicted LogRP values from the ANSA, FITC-T4, and RLBA models, are reported in Fig. S1, Fig. S4, and Fig. S7, respectively. All the points are regularly distributed along the diagonal, and no relevant anomalies can be highlighted from these plots.

3.1. QSAR model for the ANSA dataset

The best QSAR model developed for the ANSA dataset is based on four molecular descriptors and is reported in Eq. 2:

$$\begin{aligned} \boldsymbol{LogRP_{ANSA}} = {} & -1.2\ (\pm 0.30)\ +\ 1.6 \times 10^2\ (\pm 23)\ \bullet \boldsymbol{AATSC1c} \\ & +\ 1.7\ (\pm 0.24) \bullet \boldsymbol{PubchemFP381}\ +\ 1.5 \times 10^{-2}\ (\pm 4.8 \\ & \times 10^{-3}) \bullet \boldsymbol{ATSC2s}\ +\ 0.28\ (\pm 7.6 \times 10^{-2}) \bullet \boldsymbol{nX} \end{aligned} \quad (2)$$

(n° Training set = 59; n° Test set = 20; $R^2 = 0.89$; $RMSE_{TR} = 0.38$; $Q^2_{loo} = 0.86$; $Q^2_{lmo} = 0.86$; $\Delta K = 0.088$; $Q^2_{F3} = 0.88$; $RMSE_{TEST} = 0.39$; $R^2_{YS} = 0.070$; $Q^2_{YS} = -0.11$)

The model fits well ($R^2 = 0.89$; $RMSE_{TR} = 0.38$), is internally robust ($Q^2_{loo} = 0.86$; $Q^2_{lmo} = 0.86$), and externally predictive ($Q^2_{F3} = 0.88$ $RMSE_{TEST} = 0.39$). Furthermore, $R^2_{YS} = 0.070$ and $Q^2_{YS} = -0.11$ exclude chance correlation between the selected descriptors and the response, while positive value of ΔK confirms the absence of multicollinearity. This QSAR covers a wider structural domain compared to literature ones with similar performances and complexity [50].

Fig. S1 shows the plot of experimental versus predicted LogRP values by the ANSA model (Eq. 2). As was mentioned above, the analysis of the AD of the model reported in Fig. S2 does not highlight problematic chemicals. Only 2,3,4,5,6-Pentafluorobenzoic acid (ID = 54) has a leverage value ($h = 0.26$) slightly larger than the threshold ($h^* = 0.25$). This is probably due to the fact that only ID 54 and ID 73 (3,5-Dibromo-4-hydroxybenzotrifluoride) contain fluorine atoms in Dataset ANSA. Indeed, not surprisingly, ID 73 has a leverage value ($h = 0.24$) lower but very close to the threshold. However, LogRP is correctly predicted for both chemicals with standardised residuals smaller than ± 2.5 standard deviation units. The QQ plot (Fig. S3) indicates reasonably normally distributed residuals.

The four molecular descriptors selected in Eq. 2 are listed and defined in Table S1. These descriptors have a positive relationship with LogRP (i.e., positive value of the regression coefficient in Eq. 2), meaning that an increase of their values promotes the binding with hTTR of chemicals. AATSC1c and PubchemFP381 are the most influential descriptors of Eq. 2, which are also the two most frequently selected descriptors across the whole population of developed models. AATSC1c is an autocorrelation descriptor which takes into account the spatial distribution of charges along the molecular structure at lag 1, where the lag is the topological distance between pairs of atoms. Previous literature work [32] demonstrated that the dominant binding interactions between halogenated thiophenols and hTTR mainly involve

noncovalent interactions, where the spatial distribution of charges can play an important role. Since halogenated thiophenols are those with the highest values of AATSC1c, among chemicals included in the training set, this descriptor could encode information about the tendency of chemicals in forming intermolecular bonds with hTTR. PubchemFP381 is a fragment that, concerning our training set, distinguishes chemicals with a phenoxy group (i.e., PubchemFP381 = 1 when at least one oxygen atom in the structure is bound to an aromatic ring) from those missing this feature (in this case, PubchemFP381 = 0). In the training set, this descriptor allows for the identification of phenols, hydroxylated PBDEs, and hydroxylated PCBs. The selection of this descriptor is in line with literature findings [19,82], showing that the binding with hTTR is favoured by the presence of hydroxyl groups, aromatic rings, and halogen atoms, which recall the main structural features of THs. This is supported by the fact that most of the training set chemicals with PubchemFP381 = 1 are those with the largest values of experimental RPs. The ATSC2s descriptor reflects how electronic properties, such as polarizability and electronegativity, are distributed along the molecular structure at a lag 2 distance [98]. As seen for AATSC1c, also ATSC2s encodes useful information about the formation of intermolecular bonds with hTTR, such as hydrogen bonds where electronegativity is the relevant factor [99]. Both AATSC1c and ATSC2s are calculated from the Moreau-Broto's autocorrelation coefficient [100]. The nX descriptor is a counter of the number of halogen atoms. Previous studies demonstrated that halogen atoms are relevant to the binding of chemicals with the hTTR [32,35], recalling the presence of iodine atoms in the THs molecular structures. Globally, in the training set, chemicals with the highest experimental response values include from three to six halogen atoms. Chemicals with the lowest LogRP values include a lower number of halogen atoms, ranging from zero to three. After demonstrating the model's predictive potential through external validation, the model was re-developed by combining both training and test sets, in order to use all the available information which is expected to improve the model reliability. The equation of the model is reported below:

$$\begin{aligned} \boldsymbol{LogRP_{ANSA}} = &-1.1\ (\pm 0.26) + 1.7 \times 10^{2}\ (\pm 20) \bullet \boldsymbol{AATSC1c} \\ &+ 1.7\ (\pm 0.20) \bullet \boldsymbol{PubchemFP381} + 0.29\ (6.7 \times 10^{-2}) \\ &\bullet \boldsymbol{nX} + 1.5 \times 10^{-2}\ (\pm 4.1 \times 10^{-3}) \bullet \boldsymbol{ATSC2s} \end{aligned} \quad (3)$$

(n° Training set = 79; R^2 = 0.89; $RMSE_{TR}$ = 0.38; Q^2_{loo} = 0.87; Q^2_{lmo} = 0.87; ΔK = 0.14; R^2_{YS} = 0.052; Q^2_{YS} = −0.082)

As expected, taking into account the good results of the external validation of Eq. 2, as well as the distribution of the training and test sets in Fig. S1 and Fig. S2, the values of the intercept and of the coefficients in Eq. 2 and in Eq. 3 are very similar. This further confirms the robustness of the selected descriptors, and Eq. 3 can be suggested as a reliable QSAR for the prediction of LogRP ANSA of new chemicals.

3.2. QSAR model for the FITC-T4 Dataset

The equation of the best QSAR model developed using the information included in the Dataset FITC-T4 is:

$$\begin{aligned} \boldsymbol{LogRP_{FITC-T4}} = &-0.84\ (\pm 0.35) + 0.32\ (\pm 4.8 \times 10^{-2}) \bullet \boldsymbol{naasC} \\ &- 1.6\ (\pm 0.48) \bullet \boldsymbol{SpMin4\_Bhs} + 1.0 \times 10^{-2}\ (\pm 6.2 \\ &\times 10^{-3}) \bullet \boldsymbol{VE3\_Dzs} \end{aligned} \quad (4)$$

(n° Training set = 38; n° Test set = 12; R^2 = 0.85; $RMSE_{TR}$ = 0.38; Q^2_{loo} = 0.81; Q^2_{lmo} = 0.80; ΔK = 0.030; Q^2_{F3} = 0.83; $RMSE_{TEST}$ = 0.40; R^2_{YS} = 0.082; Q^2_{YS} = −0.5)

This model has good fitting (R^2 = 0.85; $RMSE_{TR}$ = 0.38), is internally robust (Q^2_{loo} = 0.81; Q^2_{lmo} = 0.80), and has a good external predictive potential (Q^2_{F3} = 0.83; $RMSE_{TEST}$ = 0.40). The absence of chance correlation between the descriptors and the response is confirmed by the Y-scrambling procedure (R^2_{YS} = 0.082; Q^2_{YS} = −0.15), while the absence of multicollinearity is confirmed by the positive value of ΔK.

The plot of experimental versus predicted LogRP values by the FITC-T4 model is reported in Fig. S4.

The Williams plot, reported in Fig. S5, shows that all the chemicals fall into the structural applicability domain of the model. Only tetrabromobisphenol A-mono (allyl ether) (ID = 21) has a standardised residual in prediction of –2.6, slightly exceeding the reference range of ± 2.5 standard deviation units. It is interesting to note that this chemical is the only bisphenol derivative in the test set, having also the largest value of experimental RP compared to all the bisphenol derivatives included in the dataset. This may be the reason of the large residual in prediction. The QQ plot (Fig. S6) shows reasonably normally distributed residuals. The three molecular descriptors selected in Eq. 4 are listed in Table S2.

The most important molecular descriptor is naasC, positively correlated with the response. This suggests that larger values of this descriptor contribute to a greater binding ability with hTTR. NaasC is an electrotopological state index [101], encoding for electronic and topological information, such as electronegativity and polarizability, as seen for ATSC2s in Eq. 2; it counts the number of bonds that involve aromatic carbons, excluding those with hydrogen and carbon atoms in the same aromatic ring. This descriptor, in the training set of the FITC-T4 model, discriminates between more and less substituted diphenyls (mainly in terms of hydroxyl groups, halogen atoms, ether bonds), phenols, and not aromatic chemicals (such as PFAS), with naasC values decreasing in this order. Globally, chemicals with the highest values of experimental RPs are those with the highest values of naasC. Overall, it seems that the presence of hydroxyl groups, aromatic rings, and halogen atoms, is important in promoting the binding affinity with the hTTR. These results are in agreement with those highlighted for the ANSA model, despite the use of different training sets for the development of the two models.

The second most important molecular descriptor is SpMin4_Bhs, which takes into account the topology of the chemicals and has negative sign in Eq. 4. Indeed, in this dataset, the smallest values of this descriptor are calculated for PFAS, which are known to bind with hTTR [30,47,62,74]. The correlation between SpMin4_Bhs and LogRP values for the 14 PFAS in the training set is −0.88. These results are supported by the dependency of SpMin4_Bhs with the carbon chain length of PFAS, which specularly reflects the dependency of the binding capacity of PFAS with their own carbon chain length. In the literature, it is demonstrated that PFAS with a terminal carboxylic acid group, and with a carbon chain length between six and ten, are more potent hTTR binders than the equivalents with a longer or shorter carbon chain length [47]. For sulfonic PFAS, a carbon chain length equal to eight optimises the binding with hTTR [47,74]. In the studied dataset, SpMin4_Bhs values are lower for PFAS with a terminal carboxylic acid group and with a carbon chain length between six and ten, while are higher for the equivalents with a longer or shorter carbon chain length. Similarly, SpMin4_Bhs values exhibited a decreasing trend with increasing length of the carbon chain in sulfonic PFAS, reaching a minimum value for the eight-carbon sulfonic PFAS (sulfonic PFAS with a carbon chain length larger than eight are missing in the studied dataset). Therefore, the selection of this descriptor in Eq. 4, seems to be particularly related to PFAS disruptive effect on hTTR functions, considering its negative sign in Eq. 4. It should be noted that naasC and SpMin4_Bhs are two of the most frequently selected descriptors across the population of developed models, and the two most frequent ones in the population of models developed with three variables. VE3_Dzs is the less influential descriptor, which is calculated from the Barysz distance matrix that accounts simultaneously for heteroatoms and multiple bonds. In Eq. 4, it has a positive correlation with the response, suggesting that heteroatoms and multiple bonds together have a role in increasing the binding affinity with hTTR. In particular, the value of this descriptor increases with the response for chemicals that share the same value of naasC. Therefore, VE3_Dzs

provides additional information to discriminate different molecular and atomic aspects for chemicals having the same value of naasC.

Finally, FITC-T4 model was re-developed by combining both training and test sets in order to use all the available information. The equation of the full model is reported as follows (Eq. 5):

$$\boldsymbol{LogRP}_{\boldsymbol{FITC-T4}} = -0.97\ (\pm 0.30)\ +\ 0.33\ (\pm 4.3 \times 10^{-2}) \bullet \boldsymbol{naasC} - 1.5\ (\pm 0.42) \bullet \boldsymbol{SpMin4\_Bhs}\ +\ 1.1 \times 10^{-2}\ (\pm 5.8 \times 10^{-3}) \bullet \boldsymbol{VE3\_Dzs} \quad (5)$$

(n° Training set = 50; $R^2 = 0.84$; $RMSE_{TR} = 0.38$; $Q^2_{loo} = 0.81$; $Q^2_{lmo} = 0.80$; $\Delta K = 0.058$; $R^2_{YS} = 0.061$; $Q^2_{YS} = -0.11$)

As was described for Eq. 2 and Eq. 3, the values of the intercept and of the coefficients in the Eq. 4 and Eq. 5 are very similar. Eq. 5 is the final QSAR for the prediction of LogRP FITC-T4 for new chemicals.

3.3. QSAR model for RLBA Dataset

The first modelling attempt performed on Dataset RLBA highlighted four chemicals as frequent outliers in the population of models, which may negatively affect their performances: 3,3',5,5'-tetrachlorobiphenyl (ID = 46), 3,3',4,4',5-pentachlorobiphenyl (ID = 93), 3,3',4,5,5'-pentachlorobiphenyl (ID = 75), and perfluorobutanesulfonic acid (ID = 109). The LogRP value (0.85) of 3,3',5,5'-tetrachlorobiphenyl is 2 log units larger than LogRP reported for the other four tetrachlorobiphenyls, which ranges from −1.1 to −1.3. By excluding the two pentachlorobiphenyls listed above from the dataset, LogRP ranges from −1.8 to 0.44 for the other nine pentachlorobiphenyls, while 3,3',4,4',5-pentachlorobiphenyl and 3,3',4,5,5'-pentachlorobiphenyl LogRP values are respectively −3.4 and 0.91. Perfluorobutanesulfonic acid is also an outlier, whose LogRP value (−2.5) is smaller compared to the other three perfluorinated sulfonic acids included in the dataset (LogRP range between −1.2 and −0.77). It is important to highlight that the LogRP values of these outliers are not averages of experimental values taken from different sources, which internal variability may have affected the quality of the model.

By removing the aforementioned four chemicals from the dataset, a population of models was developed, and the following QSAR was selected:

$$\boldsymbol{logRP}_{\boldsymbol{RLBA}} = -32\ (\pm 7.0)\ +\ 1.8\ (\pm 0.29) \bullet \boldsymbol{PubchemFP590} + 7.7\ (\pm 1.8) \bullet \boldsymbol{SpMax1\_Bhe}\ -\ 1.3\ (\pm 0.43) \bullet \boldsymbol{PubchemFP18}\ -1.2\ (\pm 0.39) \bullet \boldsymbol{GATS5c}\ -\ 20\ (\pm 7.9) \bullet \boldsymbol{AATSC1e}\ +\ 4.9 \times 10^{-3}\ (\pm 2.0 \times 10^{-3}) \bullet \boldsymbol{AATS4v} \quad (6)$$

(n° Training set = 100; n° Test set = 33; $R^2 = 0.81$; $RMSE_{TR} = 0.52$; $Q^2_{loo} = 0.77$; $Q^2_{lmo} = 0.77$; $\Delta K = 0.073$; $Q^2_{F3} = 0.69$; $RMSE_{TEST} = 0.66$; $R^2_{YS} = 0.061$; $Q^2_{YS} = -0.088$)

The RLBA model has good validation metrics values, slightly worse compared to the ANSA and FITC-T4 models. This could be due to the higher structural heterogeneity and size of the RLBA dataset, compared to the other datasets. However, the model has good fitting according to R^2 and $RMSE_{TR}$ values, Q^2_{loo} and Q^2_{lmo} values support the model internal robustness, while Q^2_{F3} and $RMSE_{TEST}$ support its external predictivity when applied to new chemicals. The R^2_{YS} and Q^2_{YS} values exclude chance correlation, while ΔK value greater than zero excludes multicollinearity. The removal of the aforementioned outliers led to an improvement of the model performances in terms of fitting and internal robustness (i.e. former QSARs with six descriptors had R^2 from 0.72 to 0.75, $RMSE_{TR}$ from 0.61 to 0.64, while Q^2_{loo} from 0.68 to 0.71).

Experimental versus predicted LogRP values of RLBA model can be found in Fig. S7.

The Williams plot, reported in Fig. S8, shows that most chemicals are within the applicability domain of the model, confirming the absence of both response and structural outliers, with an exception for the chemical perfluoro-n-pentanoic acid (ID = 120). Despite this chemical has a leverage value ($h = 0.24$) greater than the defined threshold ($h^* = 0.21$), the model predicts reasonably LogRP (standardised residual smaller than ±2.5 standard deviation units). Residuals are reasonably normally distributed according to the QQ plot (Fig. S9). The selected molecular descriptors are listed in Table S3.

The PubchemFP590 descriptor encodes for the presence of hydroxylated aromatic rings in the molecular structure (value 1 of the descriptor). It is the most relevant variable in Eq. 6, and its positive sign indicates that the presence of the molecular fragment leads to an increase of the binding affinity with hTTR. In the training set, this descriptor identifies chemicals such as hydroxylated PCBs and PBDEs, phenols, parabens, hydroxylated dioxins and furans, hydroxylated xanthones, and bisphenol A derivatives. Most of the chemicals with PubchemFP590 = 1, included in the training set, are those with the greatest values of experimental RPs. This is in agreement with the structural information selected in ANSA and FITC-T4 models, despite using different training sets. Furthermore, PubchemFP590 highlights the same critical structural feature that promotes the binding with hTTR, previously highlighted.

The second most relevant descriptor in Eq. 6 is SpMax1_Bhe, which is calculated from the Burden matrix and is weighted by relative Sanderson electronegativities. SpMax1_Bhe belongs to the same family of descriptors as SpMin4_Bhs, selected in the FITC-T4 model. In Eq. 6, this descriptor has a positive sign, indicating its correlation with the response. Chemicals with the greatest values of this descriptor are mainly PCBs, hydroxylated PCBs, hydroxylated xanthones, then PBDEs, hydroxylated PBDEs, PFAS, bisphenol A derivatives, and finally phenols. The role of this descriptor may be to take into account the binding ability of non-hydroxylated biphenyls (i.e. PCBs, PBDEs) and PFAS, which is not considered by PubchemFP590. Phenols are the compounds with the smallest SpMax1_Bhe value and the lowest binding affinity, compared to chemicals with two aromatic rings. This is not surprising, since the binding activity of a chemical is affected by the structural similarity to THs (which is high for halogenated and hydroxylated biphenyls and biphenyls ethers).

The third most relevant molecular descriptor is PubchemFP18, which has an absolute standardized coefficient comparable to SpMax1_Bhe in Eq. 6 (–0.45 and 0.46, respectively). In this training set, PubchemFP18 identifies chemicals with one or more oxygen atoms (value 1 of the descriptor) but no hydroxyl groups bound to the aromatic ring. Therefore, the combination of PubchemFP18 and PubchemFP590 allows for the correct modelling and the discrimination across the activity of hydroxylated aromatics, mentioned above, from non-hydroxylated aromatics (mainly PBDEs) and not aromatic but hydroxylated compounds (mainly carboxylic and sulfonic PFAS).

The three less important molecular descriptors selected in Eq. 6 are GATS5c, AATSC1e, and AATS4v, which belong to the 2D autocorrelation descriptors, which are known for their usefulness and applicability [99]. In this context, these descriptors capture the topological distribution of several properties of the chemical structures. Differently from the other autocorrelation descriptors, which are calculated from the Moreau-Broto's autocorrelation coefficient, GATS5c is calculated from the Geary's coefficient and is related to the distribution of charges across the molecules.

Once the model's predictive ability was demonstrated through external validation, training and test sets were pooled and the model was re-developed, resulting in the following equation:

$$\begin{aligned} \boldsymbol{LogRP_{RLBA}} = & -30\ (\pm 6.4) + 1.9\ (\pm 0.28) \bullet \boldsymbol{PubchemFP590} \\ & - 1.5\ (\pm 0.42) \bullet \boldsymbol{PubchemFP18} + 7.4\ (\pm 1.6) \\ & \bullet \boldsymbol{SpMax1_{Bhe}} - 22\ (\pm 7.1) \bullet \boldsymbol{AATSC1e} - 1.3\ (\pm 0.38) \\ & \bullet \boldsymbol{GATS5c} + 4.9 \times 10^{-3}\ (\pm 1.8 \times 10^{-3}) \bullet \boldsymbol{AATS4v} \end{aligned} \quad (7)$$

(n° Training set = 133; $R^2 = 0.78$; $Q^2_{loo} = 0.75$; $Q^2_{lmo} = 0.75$; $RMSE_{TR} = 0.55$; $\Delta K = 0.067$; $R^2_{YS} = 0.045$; $Q^2_{YS} = -0.065$)

Consistently with former models, considering the similarities between the Eq. 6 and Eq. 7, Eq. 7 is suggested for the prediction of LogRP RLBA of new chemicals.

Finally, it is interesting to highlight that the main structural features selected in Eq. 2, Eq. 4, and Eq. 6, are similar. Although the structural information included in the training set of each model led to the selection of different sets of molecular descriptors, the mechanistic interpretation of the overall set of the selected descriptors mainly converges towards the same key molecular structures that affect binding capacity of compounds, as was previously described.

PubchemFP381, selected in the ANSA model, distinguishes chemicals with a phenoxy group (mainly hydroxylated aromatic compounds). Similar structural information (i.e. the discrimination of non-hydroxylated aromatic compounds from hydroxylated aromatic compounds (e.g., PBDEs and PCBs from their hydroxylated metabolites)) is encoded by the combination of PubchemFP590 and PubchemFP18 selected in the RLBA model. Furthermore, the relevance of substituted aromatic rings (in terms of hydroxyl groups and halogen atoms, based on our training set) is encoded by naasC descriptor selected in the FITC-T4 model.

The descriptor nX selected in the ANSA model counts the number of halogen atoms, as well as VE3_Dzs, selected in the FITC-T4 model, is sensitive to this structural feature.

Autocorrelation descriptors, selected in the ANSA and RLBA models, encode for the topological distribution of different properties along the chemical structure, which can influence the tendency of compounds in forming chemical bonds with hTTR. In particular, the same descriptor AATSC1, weighted on charges (i.e., AATSC1c) and on electronegativity (i.e., AATSC1e), is selected in the ANSA and RLBA models, respectively. The influence of charges is also encoded by the descriptor GATS5c, which is selected in the RLBA model.

3.4. Models application to Dataset II and PCA analysis of Dataset III

Equations calibrated on the full datasets ANSA, FITC-T4, and RLBA, respectively Eq. 3, Eq. 5 and Eq. 7, were applied to predict missing experimental values of T4-hTTR competing potencies of 223 chemicals included in Dataset I, leading to Dataset II. Among these chemicals, 73 belonging to different chemical classes were discarded from further analysis because of unreliable predictions provided by at least one QSAR (see Section 2.1), mainly due to poor representation of these chemicals in the QSARs training sets. In particular, predictions were excluded for PFAS and non-hydroxylated PBDEs (ANSA model), as well as for phosphates and fatty acids (FITC-T4 and RLBA models). Experimental and reliable LogRP predictions were available for 150 chemicals and included in Dataset III, which was explored by PCA. Fig. 1 shows the PCA biplot (scores and loadings) of the first two components, where the first component explains more than 65 % of the total variance (cumulative explained variance for PC1 and PC2 is about 85 % of the total variance). Scores and loadings values calculated for the first two principal components are reported in Supplementary Material S2.

The loadings are positively correlated and equally important along PC1 (internal correlations: LogRP ANSA – LogRP FITC-T4 = 0.44; LogRP ANSA – LogRP RLBA = 0.52; LogRP FITC-T4 – LogRP RLBA = 0.49), while PC2 highlights intra methods differences between LogRP FITC-T4 and LogRP ANSA (loadings are opposite in sign).

Considering the values of the loadings and their agreement in the biplot, PC1 ranks chemicals from right to left according to their increasing T4-hTTR competing potency. Chemicals on the left side of the graph are strong hTTR binders with large LogRP values, either experimental or predicted by the three QSARs. On the contrary, chemicals on the right side of the graph have small values of LogRP, and consequently are weak hTTR binders. As expected, the strongest hTTR binders have two hydroxylated or non-hydroxylated aromatic rings with a high number of halogen atoms (at least three), such as PCBs, PBDEs, and their hydroxylated metabolites, halogenated bisphenol A derivatives (e.g. tetrabromobisphenol A), halogenated dioxins (e.g. 2-hydroxy-1,3,7,8-tetrachlorodibenzo-4-dioxin), highly halogenated phenols (e.g. pentabromophenol), and thiophenols (e.g. pentachlorothiophenol). Chemicals with less or no substituents as hydroxyl groups and halogen atoms, with one or two aromatic rings, and with heteroatoms such as nitrogen and sulphur in their structure, are less powerful competitors with T4 in hTTR

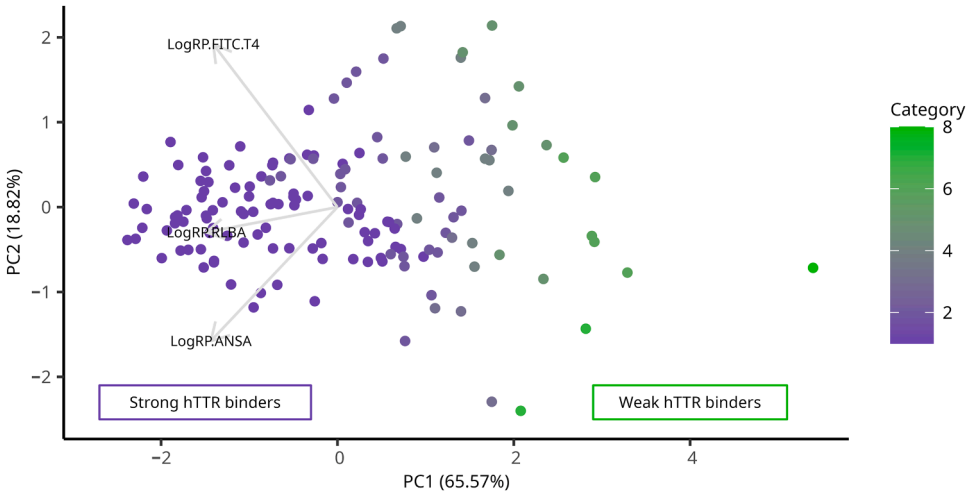


Fig. 1. Biplot resulting from PCA analysis carried out on Dataset III to compare experimental or predicted LogRPs for the three *in vitro* assays ANSA, FITC−T4, and RLBA. For graphical visualization, all the compounds in Dataset III have been categorised on the basis of their respective experimental or predicted LogRPs, according to the classes defined by Yang et al. (strong hTTR binder: LogRP ≥ -1.26; moderate hTTR binder: $-2.26 \leq$ LogRP < -1.26; weak hTTR binder: LogRP < -2.26) [50]. Different combinations of the aforementioned classes in experimental or predicted data define categories from 1 to 8 (e.g., a chemical falls into category 1 if it is a strong hTTR binder according to each experimental or predicted LogRP; a chemical falls into category 8 if it is a weak hTTR binder according to each experimental or predicted LogRP). Additional details on the definition of the eight categories are reported in Supplementary Material S2.

binding. Examples of weaker hTTR binders are non-hydroxylated and low halogenated PCBs and PBDEs, and related sulfated metabolites, one or two halogen-substituted thiophenols (e.g. 2-bromothiophenol) and phenols (e.g. 3-chlorophenol), xanthone derivatives (e.g. 2-hydroxy-1-methoxyxanthone), and parabens (e.g. ethyl- and propylparaben). Tris (2-chloroethyl) phosphate has the lowest LogRP values experimentally measured with the ANSA method and predicted by FITC-T4 and RLBA models (all LogRPs less than –3). This chemical does not have any aromatic ring.

Finally, PC2, which explains about 20 % of the residual variance, highlights that intra methods differences are more evident between LogRP-FITC and LogRP-ANSA values, in particular for moderate/weak binders. These differences may be explained by a prevalence of predicted LogRP values by Eq. 3 and/or Eq. 5 for substances with PC2 scores smaller than −1 or PC2 scores greater than 1. Therefore, the separation of the loadings LogRP FITC-T4 and LogRP-ANSA along PC2, is probably due to the uncertainty in prediction and to the intra methods experimental uncertainty.

4. Conclusions

On the basis of the increasing concern against EDCs and the need for their early identification using *in silico* approaches, this work was primarily focused on the development of three new MLR QSARs, to detect THSDCs by predicting the competition with the T4 for the binding with hTTR. These QSARs were developed using new and curated datasets which included, to the best of our knowledge, all the currently available quantitative experimental binding affinity values measured using three *vitro* assays (i.e., ANSA, FITC-T4, and RLBA). QSARs developed to predict LogRP values for the ANSA and RLBA datasets, which are based on heterogeneous training sets characterised by broad response ranges (from weak to strong hTTR binders), have wider applicability domains compared to existing similar QSARs. To our knowledge, no QSARs have been developed, so far, to predict LogRP measured with the FITC-T4 methodology, other than the here proposed model.

The mechanistic interpretation of the descriptors selected in Eq. 2, Eq. 4 and Eq. 6, is coherent with the most important structural features which are known to be involved in the binding of chemicals to hTTR, such as hydroxylated and halogenated aromatic rings, as well as descriptors encoding for specific electronic environments. The binding of chemicals to hTTR was further investigated running PCA analysis on experimental and reliable LogRPs predictions for the three datasets. PCA ranked the studied chemicals along PC1, according to their relative competitive potency and confirmed that the structural features mentioned above are relevant to distinguish between strong and weak binding activity.

A limitation of this study is related to the fact that some specific chemical classes (i.e. PFAS, non-hydroxylated PBDEs, phosphates, and fatty acids) were not included in all the training sets, and therefore they were predicted outside the applicability domain of one or more of the here proposed QSARs. This led to the exclusion of these chemicals from the PCA analysis, which underlines the need for additional *in vitro* tests in these specific areas of the structural space. These new data would be useful to refine the existing QSARs, and to generate new models covering wider structural domains. Furthermore, as the proposed models target only hTTR binders, they may overestimate the binding affinity of inactive compounds.

However, the here proposed QSARs, which are characterised by low complexity and are statistically robust and predictive, can be applied for the early screening of THSDCs on the basis of the molecular structure. To ease their application and support hazard assessment procedures, they have been implemented in the QSAR-ME Profiler beta version 1.02 software (freely available at https://dunant.dista.uninsubria.it/qsar/).

Environmental implication

Thyroid Hormone System-Disrupting Chemicals (THSDCs) are endocrine disruptors able to induce severe adverse effects in living organisms. International regulations call for the identification, regulation, and substitution of these hazardous substances. This study provides *in silico* models to identify potential THSDCs from the molecular structure. These models are based on a new curated collection of data for heterogeneous organics and cover large applicability domains. The models are valid, time and cost-effective alternatives to experimental tests to speed up hazard and risk assessment procedures, and to provide suggestions and mechanistic insight for the identification of THSDCs.

CRediT authorship contribution statement

Nicola Chirico: Writing – original draft, Software, Methodology, Investigation. **Ester Papa:** Writing – original draft, Supervision, Project administration, Methodology, Investigation, Conceptualization. **Marco Evangelista:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the PhD Program in Chemical and Environmental Sciences (DiSCA, University of Insubria) (PhD scholarship to Marco Evangelista).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jhazmat.2024.136188.

Data availability

All data used in this study are reported as Supplementary Material.

References

[1] De Coster, S., van Larebeke, N., 2012. Endocrine-disrupting chemicals: associated disorders and mechanisms of action. J Environ Public Health 2012, 713696. https://doi.org/10.1155/2012/713696.
[2] Kahn, L.G., Philippat, C., Nakayama, S.F., Slama, R., Trasande, L., 2020. Endocrine-disrupting chemicals: implications for human health. Lancet Diabetes Endocrinol 8, 703–718. https://doi.org/10.1016/S2213-8587(20)30129-7.
[3] World Health Organization, International Programme on Chemical Safety (WHO-IPCS), 2002.Global assessment on the state of the science of endocrine disruptors.
[4] Khan, S., Naushad, Mu, Govarthanan, M., Iqbal, J., Alfadul, S.M., 2022. Emerging contaminants of high concern for the environment: current trends and future research. Environ Res 207, 112609. https://doi.org/10.1016/j.envres.2021.112609.
[5] Zoeller, R.T., Brown, T.R., Doan, L.L., Gore, A.C., Skakkebaek, N.E., Soto, A.M., Woodruff, T.J., Vom Saal, F.S., 2012. Endocrine-disrupting chemicals and public health protection: a statement of principles from the endocrine society. Endocrinology 153, 4097–4110. https://doi.org/10.1210/en.2012-1422.
[6] Diamanti-Kandarakis, E., Bourguignon, J.-P., Giudice, L.C., Hauser, R., Prins, G. S., Soto, A.M., Zoeller, R.T., Gore, A.C., 2009. Endocrine-disrupting chemicals: an endocrine society scientific statement. Endocr Rev 30, 293–342. https://doi.org/10.1210/er.2009-0002.
[7] European Commission, 1999. Community Strategy for Endocrine Disrupters - a range of substances suspected of interfering with the hormone systems of humans and wildlife.
[8] Andersson, N., Arena, M., Auteri, D., Barmaz, S., Grignard, E., Kienzler, A., Lepper, P., Lostia, A.M., Munn, S., Parra Morte, J.M., Pellizzato, F., Tarazona, J., Terron, A., Van der Linden, S., 2018. Guidance for the identification of endocrine

disruptors in the context of regulations (EU) No 528/2012 and (EC) No 1107/2009. EFSA J 16, e05311. https://doi.org/10.2903/j.efsa.2018.5311.
[9] European Commission, 2019. The European Green Deal.
[10] European Commission, 2020. Chemicals Strategy for Sustainability towards a Toxic-free Environment.
[11] European Commission, 2022. Amending Regulation (EC) No 1272/2008 as regards hazard classes and criteria for the classification, labelling and packaging of substances and mixtures.
[12] Calsolaro, V., Pasqualetti, G., Niccolai, F., Caraccio, N., Monzani, F., 2017. Thyroid disrupting chemicals. Int J Mol Sci 18, 2583. https://doi.org/10.3390/ijms18122583.
[13] Crofton, K.M., 2008. Thyroid disrupting chemicals: mechanisms and mixtures. Int J Androl 31, 209–223. https://doi.org/10.1111/j.1365-2605.2007.00857.x.
[14] Köhrle, J., Frädrich, C., 2021. Thyroid hormone system disrupting chemicals. Best Pract Res Clin Endocrinol Metab 35, 101562. https://doi.org/10.1016/j.beem.2021.101562.
[15] Sokal, A., Jarmakiewicz-Czaja, S., Tabarkiewicz, J., Filip, R., 2021. Dietary intake of endocrine disrupting substances presents in environment and their impact on thyroid function. Nutrients 13, 867. https://doi.org/10.3390/nu13030867.
[16] Miller, M.D., Crofton, K.M., Rice, D.C., Zoeller, R.T., 2009. Thyroid-disrupting chemicals: interpreting upstream biomarkers of adverse outcomes. Environ Health Perspect 117, 1033–1041. https://doi.org/10.1289/ehp.0800247.
[17] Noyes, P.D., Friedman, K.P., Browne, P., Haselman, J.T., Gilbert, M.E., Hornung, M.W., Barone, S., Crofton, K.M., Laws, S.C., Stoker, T.E., Simmons, S. O., Tietge, J.E., Degitz, S.J., 2019. Evaluating chemicals for thyroid disruption: opportunities and challenges with in vitro testing and adverse outcome pathway approaches. Environ Health Perspect 127, 095001. https://doi.org/10.1289/EHP5297.
[18] Boas, M., Feldt-Rasmussen, U., Main, K.M., 2012. Thyroid effects of endocrine disrupting chemicals. Mol Cell Endocrinol, Health Impacts Endocr Disrupters 355, 240–248. https://doi.org/10.1016/j.mce.2011.09.005.
[19] Brucker-Davis, F., 1998. Effects of environmental synthetic chemicals on thyroid function. Thyroid 8, 827–856. https://doi.org/10.1089/thy.1998.8.827.
[20] Bernasconi, C., Langezaal, I., Bartnicka, J., Asturiol, D., Bowe, G., Coecke, S., Kienzler, A., Liska, R., Milcamps, A., Munoz-Pineiro, M.A., Pistollato, F., Whelan, M., 2023. Validation of a battery of mechanistic methods relevant for the detection of chemicals that can disrupt the thyroid hormone system. https://doi.org/10.2760/862948.
[21] Kortenkamp, A., Axelstad, M., Baig, A.H., Bergman, Å., Bornehag, C.-G., Cenijn, P., Christiansen, S., Demeneix, B., Derakhshan, A., Fini, J.-B., Frädrich, C., Hamers, T., Hellwig, L., Köhrle, J., Korevaar, T.I.M., Lindberg, J., Martin, O., Meima, M.E., Mergenthaler, P., Nikolov, N., Du Pasquier, D., Peeters, R.P., Platzack, B., Ramhøj, L., Remaud, S., Renko, K., Scholze, M., Stachelscheid, H., Svingen, T., Wagenaars, F., Wedebye, E.B., Zoeller, R.T., 2020. Removing critical gaps in chemical test methods by developing new assays for the identification of thyroid hormone system-disrupting chemicals—the ATHENA project. Int J Mol Sci 21, 3123. https://doi.org/10.3390/ijms21093123.
[22] OECD, 2014a. New Scoping Document on in vitro and ex vivo Assays for the Identification of Modulators of Thyroid Hormone Signalling. Organisation for Economic Co-operation and Development, Paris.
[23] Giammanco, M., Di Liegro, C.M., Schiera, G., Di Liegro, I., 2020. Genomic and non-genomic mechanisms of action of thyroid hormones and their catabolite 3,5-Diiodo-L-thyronine in mammals. Int J Mol Sci 21, 4140. https://doi.org/10.3390/ijms21114140.
[24] Schroeder, A.C., Privalsky, M.L., 2014. Thyroid hormones, T3 and T4, in the brain. Front Endocrinol 5. https://doi.org/10.3389/fendo.2014.00040.
[25] Richardson, S.J., 2008. Marsupial models for understanding evolution of thyroid hormone distributor proteins. Mol Cell Endocrinol 293, 32–42. https://doi.org/10.1016/j.mce.2008.04.002.
[26] Alshehri, B., D'Souza, D.G., Lee, J.Y., Petratos, S., Richardson, S.J., 2015. The diversity of mechanisms influenced by transthyretin in neurobiology: development, disease and endocrine disruption. J Neuroendocrinol 27, 303–323. https://doi.org/10.1111/jne.12271.
[27] Gião, T., Saavedra, J., Cotrina, E., Quintana, J., Llop, J., Arsequell, G., Cardoso, I., 2020. Undiscovered roles for transthyretin: from a transporter protein to a new therapeutic target for alzheimer's disease. Int J Mol Sci 21, 2075. https://doi.org/10.3390/ijms21062075.
[28] Schussler, G.C., 2000. The thyroxine-binding proteins. Thyroid 10, 141–149. https://doi.org/10.1089/thy.2000.10.141.
[29] Meerts, I.A.T.M., Assink, Y., Cenijn, P.H., van den Berg, J.H.J., Weijers, B.M., Bergman, Å., Koeman, J.H., Brouwer, A., 2002. Placental transfer of a hydroxylated polychlorinated biphenyl and effects on fetal and maternal thyroid hormone homeostasis in the rat. Toxicol Sci 68, 361–371. https://doi.org/10.1093/toxsci/68.2.361.
[30] Weiss, J.M., Andersson, P.L., Lamoree, M.H., Leonards, P.E.G., van Leeuwen, S.P. J., Hamers, T., 2009. Competitive binding of poly- and perfluorinated compounds to the thyroid hormone transport protein transthyretin. Toxicol Sci 109, 206–216. https://doi.org/10.1093/toxsci/kfp055.
[31] Šauer, P., Švecová, H., Grabicová, K., Gönül Aydın, F., Mackuľak, T., Kodeš, V., Blytt, L.D., Henninge, L.B., Grabic, R., Kocour Kroupová, H., 2021. Bisphenols emerging in Norwegian and Czech aquatic environments show transthyretin binding potency and other less-studied endocrine-disrupting activities. Sci Total Environ 751, 141801. https://doi.org/10.1016/j.scitotenv.2020.141801.
[32] Yang, X., Ou, W., Zhao, S., Wang, L., Chen, J., Kusko, R., Hong, H., Liu, H., 2021. Human transthyretin binding affinity of halogenated thiophenols and halogenated phenols: An in vitro and in silico study. Chemosphere 280, 130627. https://doi.org/10.1016/j.chemosphere.2021.130627.
[33] Hamers, T., Kamstra, J.H., Cenijn, P.H., Pencikova, K., Palkova, L., Simeckova, P., Vondracek, J., Andersson, P.L., Stenberg, M., Machala, M., 2011. In vitro toxicity profiling of ultrapure non–dioxin-like polychlorinated biphenyl congeners and their relative toxic contribution to PCB mixtures in humans. Toxicol Sci 121, 88–100. https://doi.org/10.1093/toxsci/kfr043.
[34] Hamers, T., Kamstra, J.H., Sonneveld, E., Murk, A.J., Kester, M.H.A., Andersson, P.L., Legler, J., Brouwer, A., 2006. In vitro profiling of the endocrine-disrupting potency of brominated flame retardants. Toxicol Sci 92, 157–173. https://doi.org/10.1093/toxsci/kfj187.
[35] Lans, M.C., Klasson-Wehler, E., Willemsen, M., Meussen, E., Safe, S., Brouwer, A., 1993. Structure-dependent, competitive interaction of hydroxy-polychlorobiphenyls, -dibenzo-p-dioxins and -dibenzofurans with human transthyretin. Chem Biol Interact 88, 7–21. https://doi.org/10.1016/0009-2797(93)90081-9.
[36] Ren, X.M., Guo, L.-H., 2012. Assessment of the binding of hydroxylated polybrominated diphenyl ethers to thyroid hormone transport proteins using a site-specific fluorescence probe. Environ Sci Technol 46, 4633–4640. https://doi.org/10.1021/es2046074.
[37] Ramhøj, L., Axelstad, M., Baert, Y., Cañas-Portilla, A.I., Chalmel, F., Dahmen, L., De La Vieja, A., Evrard, B., Haigis, A.-C., Hamers, T., Heikamp, K., Holbech, H., Iglesias-Hernandez, P., Knapen, D., Marchandise, L., Morthorst, J.E., Nikolov, N. G., Nissen, A.C.V.E., Oelgeschlaeger, M., Renko, K., Rogiers, V., Schüürmann, G., Stinckens, E., Stub, M.H., Torres-Ruiz, M., Van Duursen, M., Vanhaecke, T., Vergauwen, L., Wedebye, E.B., Svingen, T., 2023. New approach methods to improve human health risk assessment of thyroid hormone system disruption–a PARC project. Front Toxicol 5, 1189303. https://doi.org/10.3389/ftox.2023.1189303.
[38] OECD, 2009. Test No. 455: The Stably Transfected Human Estrogen Receptor-alpha Transcriptional Activation Assay for Detection of Estrogenic Agonist-Activity of Chemicals. Organisation for Economic Co-operation and Development, Paris.
[39] OECD, 2023. Test No. 458: Stably Transfected Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals. Organisation for Economic Co-operation and Development, Paris.
[40] Eytcheson, S.A., Olker, J.H., Friedman, K.P., Hornung, M.W., Degitz, S.J., 2023. Assessing utility of thyroid in vitro screening assays through comparisons to observed impacts in vivo. Regul Toxicol Pharmacol 144, 105491. https://doi.org/10.1016/j.yrtph.2023.105491.
[41] Mitchell, C.A., Burden, N., Bonnell, M., Hecker, M., Hutchinson, T.H., Jagla, M., LaLone, C.A., Lagadic, L., Lynn, S.G., Shore, B., Song, Y., Vliet, S.M., Wheeler, J. R., Embry, M.R., 2023. New approach methodologies for the endocrine activity toolbox: environmental assessment for fish and amphibians. Environ Toxicol Chem 42, 757–777. https://doi.org/10.1002/etc.5584.
[42] Schneider, M., Pons, J.-L., Labesse, G., Bourguet, W., 2019. In silico predictions of endocrine disruptors properties. Endocrinology 160, 2709–2716. https://doi.org/10.1210/en.2019-00382.
[43] Mansouri, K., Abdelaziz, A., Rybacka, A., Roncaglioni, A., Tropsha, A., Varnek, A., Zakharov, A., Worth, A., Richard, A.M., Grulke, C.M., Trisciuzzi, D., Fourches, D., Horvath, D., Benfenati, E., Muratov, E., Wedebye, E.B., Grisoni, F., Mangiatordi, G.F., Incisivo, G.M., Hong, H., Ng, H.W., Tetko, I.V., Balabin, I., Kancherla, J., Shen, J., Burton, J., Nicklaus, M., Cassotti, M., Nikolov, N.G., Nicolotti, O., Andersson, P.L., Zang, Q., Politi, R., Beger, R.D., Todeschini, R., Huang, R., Farag, S., Rosenberg, S.A., Slavov, S., Hu, X., Judson, R.S., 2016. CERAPP: collaborative estrogen receptor activity prediction project. Environ Health Perspect 124, 1023–1033. https://doi.org/10.1289/ehp.1510267.
[44] Mansouri, K., Kleinstreuer, N., Abdelaziz, A.M., Alberga, D., Alves, V.M., Andersson, P.L., Andrade, C.H., Bai, F., Balabin, I., Ballabio, D., Benfenati, E., Bhhatarai, B., Boyer, S., Chen, J., Consonni, V., Farag, S., Fourches, D., García-Sosa, A.T., Gramatica, P., Grisoni, F., Grulke, C.M., Hong, H., Horvath, D., Hu, X., Huang, R., Jeliazkova, N., Li, J., Li, X., Liu, H., Manganelli, S., Mangiatordi, G.F., Maran, U., Marcou, G., Martin, T., Muratov, E., Nguyen, D.-T., Nicolotti, O., Nikolov, N.G., Norinder, U., Papa, E., Petitjean, M., Piir, G., Pogodin, P., Poroikov, V., Qiao, X., Richard, A.M., Roncaglioni, A., Ruiz, P., Rupakheti, C., Sakkiah, S., Sangion, A., Schramm, K.-W., Selvaraj, C., Shah, I., Sild, S., Sun, L., Taboureau, O., Tang, Y., Tetko, I.V., Todeschini, R., Tong, W., Trisciuzzi, D., Tropsha, A., Van Den Driessche, G., Varnek, A., Wang, Z., Wedebye, E.B., Williams, A.J., Xie, H., Zakharov, A.V., Zheng, Z., Judson, R.S., 2020. CoMPARA: collaborative modeling project for androgen receptor activity. Environ Health Perspect 128, 027002. https://doi.org/10.1289/EHP5580.
[45] Richard, A.M., Huang, R., Waidyanatha, S., Shinn, P., Collins, B.J., Thillainadarajah, I., Grulke, C.M., Williams, A.J., Lougee, R.R., Judson, R.S., Houck, K.A., Shobair, M., Yang, C., Rathman, J.F., Yasgar, A., Fitzpatrick, S.C., Simeonov, A., Thomas, R.S., Crofton, K.M., Paules, R.S., Bucher, J.R., Austin, C.P., Kavlock, R.J., Tice, R.R., 2021. The Tox21 10K compound library: collaborative chemistry advancing toxicology. Chem Res Toxicol 34, 189–216. https://doi.org/10.1021/acs.chemrestox.0c00264.
[46] Rotroff, D.M., Dix, D.J., Houck, K.A., Knudsen, T.B., Martin, M.T., McLaurin, K. W., Reif, D.M., Crofton, K.M., Singh, A.V., Xia, M., Huang, R., Judson, R.S., 2013. Using in vitro high throughput screening assays to identify potential endocrine-disrupting chemicals. Environ Health Perspect 121, 7–14. https://doi.org/10.1289/ehp.1205065.
[47] Kar, S., Sepúlveda, M.S., Roy, K., Leszczynski, J., 2017. Endocrine-disrupting activity of per- and polyfluoroalkyl substances: Exploring combined approaches

of ligand and structure based modeling. Chemosphere 184, 514–523. https://doi.org/10.1016/j.chemosphere.2017.06.024.
[48] Papa, E., Kovarich, S., Gramatica, P., 2013. QSAR prediction of the competitive interaction of emerging halogenated pollutants with human transthyretin. SAR QSAR Environ Res 24, 333–349. https://doi.org/10.1080/1062936X.2013.773374.
[49] Papa, E., Kovarich, S., Gramatica, P., 2010. QSAR modeling and prediction of the endocrine-disrupting potencies of brominated flame retardants. Chem Res Toxicol 23, 946–954. https://doi.org/10.1021/tx1000392.
[50] Yang, X., Ou, W., Zhao, S., Xi, Y., Wang, L., Liu, H., 2021. Rapid screening of human transthyretin disruptors through a tiered in silico approach. ACS Sustain Chem Eng 9, 5661–5672. https://doi.org/10.1021/acssuschemeng.1c00680.
[51] Natesan, S., Wang, T., Lukacova, V., Bartus, V., Khandelwal, A., Balaz, S., 2011. Rigorous treatment of multispecies multimode ligand−receptor interactions in 3D-QSAR: CoMFA analysis of thyroxine analogs binding to transthyretin. J Chem Inf Model 51, 1132–1150. https://doi.org/10.1021/ci200055s.
[52] Yang, W., Shen, S., Mu, L., Yu, H., 2011. Structure–activity relationship study on the binding of PBDEs with thyroxine transport proteins. Environ Toxicol Chem 30, 2431–2439. https://doi.org/10.1002/etc.645.
[53] OECD, 2014b. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. Organisation for Economic Co-operation and Development, Paris.
[54] Street, M.E., Audouze, K., Legler, J., Sone, H., Palanza, P., 2021. Endocrine disrupting chemicals: current understanding, new testing strategies and future research needs. Int J Mol Sci 22, 933. https://doi.org/10.3390/ijms22020933.
[55] Cao, J., Lin, Y., Guo, L.-H., Zhang, A.-Q., Wei, Y., Yang, Y., 2010. Structure-based investigation on the binding interaction of hydroxylated polybrominated diphenyl ethers with thyroxine transport proteins. Toxicology 277, 20–28. https://doi.org/10.1016/j.tox.2010.08.012.
[56] Chauhan, K.R., Kodavanti, P.R.S., McKinney, J.D., 2000. Assessing the role of ortho-substitution on polychlorinated biphenyl binding to transthyretin, a thyroxine transport protein. Toxicol Appl Pharmacol 162, 10–21. https://doi.org/10.1006/taap.1999.8826.
[57] Cheek, A.O., Kow, K., Chen, J., McLachlan, J.A., 1999. Potential mechanisms of thyroid disruption in humans: interaction of organochlorine compounds with thyroid receptor, transthyretin, and thyroid-binding globulin. Environ Health Perspect 107, 273–278. https://doi.org/10.1289/ehp.99107273.
[58] den Besten, C., Vet, J.J.R.M., Besselink, H.T., Kiel, G.S., van Berkel, B.J.M., Beems, R., van Bladeren, P.J., 1991. The liver, kidney, and thyroid toxicity of chlorinated benzenes. Toxicol Appl Pharmacol 111, 69–81. https://doi.org/10.1016/0041-008X(91)90135-2.
[59] Gales, L., Almeida, M.R., Arsequell, G., Valencia, G., Saraiva, M.J., Damas, A.M., 2008. Iodination of salicylic acid improves its binding to transthyretin. Biochim Biophis Acta 1784, 512–517. https://doi.org/10.1016/j.bbapap.2007.11.014.
[60] Grimm, F.A., Lehmler, H.-J., He, X., Robertson, L.W., Duffel, M.W., 2013. Sulfated metabolites of polychlorinated biphenyls are high-affinity ligands for the thyroid hormone transport protein transthyretin. Environ Health Perspect 121, 657–662. https://doi.org/10.1289/ehp.1206198.
[61] Hamers, T., Kamstra, J.H., Sonneveld, E., Murk, A.J., Visser, T.J., Van Velzen, M. J.M., Brouwer, A., Bergman, Å., 2008. Biotransformation of brominated flame retardants into potentially endocrine-disrupting metabolites, with special attention to 2,2′,4,4′-tetrabromodiphenyl ether (BDE-47). Mol Nutr Food Res 52, 284–298. https://doi.org/10.1002/mnfr.200700104.
[62] Hamers, T., Kortenkamp, A., Scholze, M., Molenaar, D., Cenijn, P.H., Weiss, J.M., 2020. Transthyretin-binding activity of complex mixtures representing the composition of thyroid-hormone disrupting contaminants in house dust and human serum. Environ Health Perspect 128, 017015. https://doi.org/10.1289/EHP5911.
[63] Harju, M., Hamers, T., Kamstra, J.H., Sonneveld, E., Boon, J.P., Tysklind, M., Andersson, P.L., 2007. Quantitative structure-activity relationship modeling on in vitro endocrine effects and metabolic stability involving 26 selected brominated flame retardants. Environ Toxicol Chem 26, 816–826. https://doi.org/10.1897/06-308R.1.
[64] Hill, K.L., Mortensen, Å.-K., Teclechiel, D., Willmore, W.G., Sylte, I., Jenssen, B. M., Letcher, R.J., 2018. In vitro and in silico competitive binding of brominated polyphenyl ether contaminants with human and gull thyroid hormone transport proteins. Environ Sci Technol 52, 1533–1541. https://doi.org/10.1021/acs.est.7b04617.
[65] Huang, K., Wang, X., Zhang, H., Zeng, L., Zhang, X., Wang, B., Zhou, Y., Jing, T., 2020. Structure-directed screening and analysis of thyroid-disrupting chemicals targeting transthyretin based on molecular recognition and chromatographic separation. Environ Sci Technol 54, 5437–5445. https://doi.org/10.1021/acs.est.9b05761.
[66] Legler, J., Cenijn, P.H., Malmberg, T., Bergman, A., Brouwer, A., 2002. Determination of the endocrine disrupting potency of hydroxylated PCB's and flame retardants with in vitro bioassays. Organo Compd 53–56.
[67] Maia, F., Almeida, M. do R., Gales, L., Kijjoa, A., Pinto, M.M.M., Saraiva, M.J., Damas, A.M., 2005. The binding of xanthone derivatives to transthyretin. Biochem Pharmacol 70, 1861–1869. https://doi.org/10.1016/j.bcp.2005.09.012.
[68] Marchesini, G.R., Meimaridou, A., Haasnoot, W., Meulenberg, E., Albertus, F., Mizuguchi, M., Takeuchi, M., Irth, H., Murk, A.J., 2008. Biosensor discovery of thyroxine transport disrupting chemicals. Toxicol Appl Pharmacol 232, 150–160. https://doi.org/10.1016/j.taap.2008.06.014.
[69] Marchesini, G.R., Meulenberg, E., Haasnoot, W., Mizuguchi, M., Irth, H., 2006. Biosensor recognition of thyroid-disrupting chemicals using transport proteins. Anal Chem 78, 1107–1114. https://doi.org/10.1021/ac051399i.
[70] Meerts, I.A.T.M., van Zanden, J.J., Luijks, E.A.C., van Leeuwen-Bol, I., Marsh, G., Jakobsson, E., Bergman, Å., Brouwer, A., 2000. Potent competitive interactions of some brominated flame retardants and related compounds with human transthyretin in vitro. Toxicol Sci 56, 95–104. https://doi.org/10.1093/toxsci/56.1.95.
[71] Montaño, M., Cocco, E., Guignard, C., Marsh, G., Hoffmann, L., Bergman, Å., Gutleb, A.C., Murk, A.J., 2012. New approaches to assess the transthyretin binding capacity of bioactivated thyroid hormone disruptors. Toxicol Sci 130, 94–105. https://doi.org/10.1093/toxsci/kfs228.
[72] Ouyang, X., Froment, J., Leonards, P.E.G., Christensen, G., Tollefsen, K.-E., de Boer, J., Thomas, K.V., Lamoree, M.H., 2017. Miniaturization of a transthyretin binding assay using a fluorescent probe for high throughput screening of thyroid hormone disruption in environmental samples. Chemosphere 171, 722–728. https://doi.org/10.1016/j.chemosphere.2016.12.119.
[73] Qin, W.-P., Li, C.-H., Guo, L.-H., Ren, X.-M., Zhang, J.-Q., 2019. Binding and activity of polybrominated diphenyl ether sulfates to thyroid hormone transport proteins and nuclear receptors. Environ Sci: Process Impacts 21, 950–956. https://doi.org/10.1039/C9EM00095J.
[74] Ren, X.-M., Qin, W.-P., Cao, L.-Y., Zhang, J., Yang, Y., Wan, B., Guo, L.-H., 2016. Binding interactions of perfluoroalkyl substances with thyroid hormone transport proteins and potential toxicological implications. Toxicol 366–367 32–42. https://doi.org/10.1016/j.tox.2016.08.011.
[75] Ren, X.-M., Yao, L., Xue, Q., Shi, J., Zhang, Q., Wang, P., Fu, J., Zhang, A., Qu, G., Jiang, G., 2020. Binding and activity of tetrabromobisphenol A Mono-Ether structural analogs to thyroid hormone transport proteins and receptors. Environ Health Perspect 128, 107008. https://doi.org/10.1289/EHP6498.
[76] Rosenmai, A.K., Winge, S.B., Möller, M., Lundqvist, J., Wedebye, E.B., Nikolov, N. G., Lilith Johansson, H.K., Vinggaard, A.M., 2021. Organophosphate ester flame retardants have antiandrogenic potential and affect other endocrine related endpoints in vitro and in silico. Chemosphere 263, 127703. https://doi.org/10.1016/j.chemosphere.2020.127703.
[77] Sandau, C.D., Meerts, I.A.T.M., Letcher, R.J., McAlees, A.J., Chittim, B., Brouwer, A., Norstrom, R.J., 2000. Identification of 4-hydroxyheptachlorostyrene in polar bear plasma and its binding affinity to transthyretin: a metabolite of octachlorostyrene? Environ Sci Technol 34, 3871–3877. https://doi.org/10.1021/es001134f.
[78] Simon, E., Bytingsvik, J., Jonker, W., Leonards, P.E.G., de Boer, J., Jenssen, B.M., Lie, E., Aars, J., Hamers, T., Lamoree, M.H., 2011. Blood plasma sample preparation method for the assessment of thyroid hormone-disrupting potency in effect-directed analysis. Environ Sci Technol 45, 7936–7944. https://doi.org/10.1021/es2016389.
[79] Simon, E., van Velzen, M., Brandsma, S.H., Lie, E., Løken, K., de Boer, J., Bytingsvik, J., Jenssen, B.M., Aars, J., Hamers, T., Lamoree, M.H., 2013. Effect-directed analysis to explore the polar bear exposome: identification of thyroid hormone disrupting compounds in plasma. Environ Sci Technol 47, 8902–8912. https://doi.org/10.1021/es401696u.
[80] van den Berg, K.J., 1990. Interaction of chlorinated phenols with thyroxine binding sites of human transthyretin, albumin and thyroid binding globulin. Chem Biol Interact 76, 63–75. https://doi.org/10.1016/0009-2797(90)90034-K.
[81] Viluksela, M., Heikkinen, P., Ven, L.T.M., van der, Rendel, F., Roos, R., Esteban, J., Korkalainen, M., Lensu, S., Miettinen, H.M., Savolainen, K., Sankari, S., Lilienthal, H., Adamsson, A., Toppari, J., Herlin, M., Finnilä, M., Tuukkanen, J., Leslie, H.A., Hamers, T., Hamscher, G., Al-Anati, L., Stenius, U., Dervola, K.-S., Bogen, I.-L., Fonnum, F., Andersson, P.L., Schrenk, D., Halldin, K., Håkansson, H., 2014. Toxicological profile of ultrapure 2,2′,3,4,4′,5,5′-Heptachlorbiphenyl (PCB 180) in adult rats. PLoS One 9, e104639. https://doi.org/10.1371/journal.pone.0104639.
[82] Weiss, J.M., Andersson, P.L., Zhang, J., Simon, E., Leonards, P.E.G., Hamers, T., Lamoree, M.H., 2015. Tracing thyroid hormone-disrupting compounds: database compilation and structure-activity evaluation for an effect-directed analysis of sediment. Anal Bioanal Chem 407, 5625–5634. https://doi.org/10.1007/s00216-015-8736-9.
[83] Xi, Y., Yang, X., Zhang, H., Liu, H., Watson, P., Yang, F., 2020. Binding interactions of halo-benzoic acids, halo-benzenesulfonic acids and halo-phenylboronic acids with human transthyretin. Chemosphere 242, 125135. https://doi.org/10.1016/j.chemosphere.2019.125135.
[84] Yang, X., Ou, W., Xi, Y., Chen, J., Liu, H., 2019. Emerging polar phenolic disinfection byproducts are high-affinity human transthyretin disruptors: an in vitro and in silico study. Environ Sci Technol 53, 7019–7028. https://doi.org/10.1021/acs.est.9b00218.
[85] Zhang, J., Begum, A., Brännström, K., Grundström, C., Iakovleva, I., Olofsson, A., Sauer-Eriksson, A.E., Andersson, P.L., 2016. Structure-based virtual screening protocol for in silico identification of potential thyroid disrupting chemicals targeting transthyretin. Environ Sci Technol 50, 11984–11993. https://doi.org/10.1021/acs.est.6b02771.
[86] Zhang, J., Kamstra, J.H., Ghorbanzadeh, M., Weiss, J.M., Hamers, T., Andersson, P.L., 2015. In silico approach to identify potential thyroid hormone disruptors among currently known dust contaminants and their metabolites. Environ Sci Technol 49, 10099–10107. https://doi.org/10.1021/acs.est.5b01742.
[87] Kovarich, S., Papa, E., Li, J., Gramatica, P., 2012. QSAR classification models for the screening of the endocrine-disrupting activity of perfluorinated compounds. SAR QSAR Environ Res 23, 207–220. https://doi.org/10.1080/1062936X.2012.657235.
[88] Good, P.I., Hardin, J.W., 2009. Common Errors in Statistics (And How to Avoid Them), Third Edition. ed. Wiley.

[89] Somack, R., Andrea, T.A., Jorgensen, E.C., 1982. Thyroid hormone binding to human serum prealbumin and rat liver nuclear receptor: kinetics, contribution of the hormone phenolic hydroxyl group, and accommodation of hormone side-chain bulk. Biochemistry 21, 163–170. https://doi.org/10.1021/bi00530a028.
[90] Nilsson, S., Rask, L., Peterson, P., 1975. Studies on thyroid hormone-binding proteins. II. Binding of thyroid hormones, retinol-binding protein, and fluorescent probes to prealbumin and effects of thyroxine on prealbumin subunit self association. J Biol Chem 250, 8554–8563. https://doi.org/10.1016/S0021-9258(19)40795-3.
[91] Smith, D.S., 1977. Enhancement fluoroimmunoassay of thyroxine. FEBS Lett 77, 25–27. https://doi.org/10.1016/0014-5793(77)80185-3.
[92] Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemom Intell Lab, Proc Multivar Stat Workshop Geol Geochem 2, 37–52. https://doi.org/10.1016/0169-7439(87)80084-9.
[93] O'Boyle, N.M., Banck, M., James, C.A., Morley, C., Vandermeersch, T., Hutchison, G.R., 2011. Open babel: an open chemical toolbox. J Chemin- 3, 33. https://doi.org/10.1186/1758-2946-3-33.
[94] Yap, C.W., 2011. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. J Comput Chem 32, 1466–1474. https://doi.org/10.1002/jcc.21707.
[95] Gramatica, P., Chirico, N., Papa, E., Cassani, S., Kovarich, S., 2013. QSARINS: a new software for the development, analysis, and validation of QSAR MLR models. J Comput Chem 34, 2121–2132. https://doi.org/10.1002/jcc.23361.
[96] Todeschini, R., Consonni, V., Maiocchi, A., 1999. The K correlation index: theory development and its application in chemometrics. Chemom Intell Lab 46, 13–29. https://doi.org/10.1016/S0169-7439(98)00124-5.
[97] Tropsha, A., Gramatica, P., Gombar, V.K., 2003. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. QSAR Comb Sci 22, 69–77. https://doi.org/10.1002/qsar.200390007.
[98] Kier, L.B., Hall, L.H., 1990. An electrotopological-state index for atoms in molecules. Pharm Res 7, 801–807. https://doi.org/10.1023/A:1015952613760.
[99] Todeschini, R., Consonni, V., 2000. Handbook of molecular descriptors. Wiley,.
[100] Moreau, G., Broto, P., 1980. The autocorrelation of a topological structure: a new molecular descriptor. Nouv J Chim 4, 359–360.
[101] Hall, L.H., Kier, L.B., 1995. Electrotopological state indices for atom types: a novel combination of electronic, topological, and valence state information. J Chem Inf Comput Sci 35, 1039–1045. https://doi.org/10.1021/ci00028a014.