ERGO

Breaking down the wall between human health and environmental testing of endocrine disrupters: EndocRine Guideline Optimisation

Deliverable Report

[D4.1 Report on in silico bioavailability triggers applied to TDs]

WP:	Modelling and Biotransformation
WP number:	WP 4; Deliverable 4.1
WP Leader:	Helmholtz Centre for Environmental Research (UFZ)
Responsible task leader:	Gerrit Schüürmann and Ralph Kühne
Responsible Partner:	Helmholtz Centre for Environmental Research (UFZ)
Author:	Ralph Kühne, Friederike Leßmöllmann and Gerrit Schüürmann
Type of Deliverable:	*Report [X]* *Demonstrator [ ]* *Prototype [ ]* *Other [ ]*
Dissemination level:	*Public [X]* *Restricted Designated Group [ ]* *Prog. Participants [ ]* *Confidential (consortium) [ ]*

Programme: H2020
Grant agreement no: 825753

SDU

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 825753 (ERGO). This output reflects only the author's view and the European Union cannot be held responsible for any use that may be made of the information contained therein.

Table of contents

Summary

For a set of 1761 different chemicals known to be thyroid hormone disruptors, QSAR models to predict their bioavailability have been examined and adapted. The Lipinski's rule-of-5 model (Lipinski et al. 1979), a Caco-2 approach in terms of the apparent permeability (BMU 2015), and a model for the brain/blood partitioning (Abraham et al. 2006) have been taken into account. Since compounds yielding effects in respective in vitro assays are confirmed to be bioavailable, they are expected to be predicted as bioavailable by these models.

Lipinski's rule-of-5 model can be applied to 1715, for the remaining compounds, mainly inorganic salts, there is no octanol/water partition coefficient available. The model yields 96 (about 6 %) false negatives within the 1715 chemicals. Modification of the model and limiting the applicability domain reduces this number to 7 chemicals (0.4 %).

The Caco-2 model cannot be used for thyroid disruption bioavailability predictions with regard to thyroid hormone disruption in a reasonable manner.

The blood brain barrier permeation model is applicable to 1693 chemicals. It yields predictions of good bioavailability for the majority of them. With the recommended threshold of 10% partitioning into blood 53 (3 %) false negatives remain.

Combination of the adapted rule-of-5 model and the blood/brain partitioning threshold yields correct bioavailability predictions for all chemicals with valid results from the estimation models.

Introduction

Endocrine disrupters can only take effect if they reach their target within an organism. Failing due to the absence of a MOA-specific bioavailability will prevent them from triggering adverse effects. Thus, in order to assess the potential of a chemical for acting as an endocrine disruptor, it is important to obtain information about its thyroid hormone disruption specific bioavailability. To this end, suitable bioavailability measures and thresholds are required. Compounds yielding effects in respective in vitro assays are confirmed to be bioavailable, and thus are expected to be predicted as available by bioavailability models.

Basically, bioavailability results from physicochemical partitioning and transport processes. In consequence, it is supposed to be related to physicochemical properties, and in turn these properties depend on the chemical structures of the compounds. Therefore, QSAR models to predict properties from chemical structures are preferred tools for a screening level bioavailability prediction in the context of potential endocrine disruption.

Lipinski's rule-of-5 (Lipinski et al. 1997) profiles to detect poor absorption or permeation, the Caco-2 membrane permeation (BMU 2015), and the blood-brain barrier permeation (Abraham et al. 2006) have been examined for their potential to characterize the bioavailability of thyroid disruptors. Lipinski's rule-of-5 applies thresholds for the octanol/water partition coefficient, the molar weight, the number of hydrogen bond donors, and the number of hydrogen bond acceptors to distinguish between good and poor absorption/permeation of cells.

The Caco-2 assay is an in vitro model for gastrointestinal absorption processes using cell lines from a human colon adenocarcinoma. It is supposed to address active and passive absorption processes. Taking quantitative Caco-2 assay results in terms of the apparent permeability, classifications into low, medium, and high permeability are suggested in the literature (Pham-The et al. 2013).

Blood brain barrier permeation is an important property in drug design for compounds targeting the central nervous system. The blood/brain partition coefficient provides a quantitative measure for it. There are thresholds available (Kunwittaya et al. 2013) for the ability to pass this barrier at all, and higher thresholds for a readily or optimum permeation.

In any case, QSAR prediction models to obtain these properties have been applied. Since the bioavailability assessment is supposed to be a general screening step within a decision support system for thyroid hormone disruption, the requirement of additional assays and experimental measurements would be disadvantageous. Furthermore, for all of these properties the number of available experimental data is rather limited, and the respective chemicals are not identical to the substances currently known for adverse effects to thyroid hormones. All calculations were carried out within the software system ChemProp (UFZ 2019).

In order to examine the potential for thyroid hormone disruption specific bioavailability predictions, only compounds known to be active in at least one thyroid hormone disrupting mode could be taken into account for the study. Inactive chemicals may still be bioavailable, while potentially active compounds with regard to their chemical nature may not trigger effects without sufficient bioavailability.

From the database established in Task 4.1, data for all compounds known to be active in at least one mode and with available chemical structures were used in this studies.

Data Set

Data set creation. Compounds known to be active as thyroid receptor antagonists, AhR receptor (human liver) agonists, thyrotropin-releasing hormone receptor agonists or antagonists, retinoic acid receptor agonists or antagonists, and RXR signalling pathway agonists have been taken into account. In many cases, data from more than one experimental study for the same compound and the same mode of action were available. These data were compared in terms of their classification between being active or inactive, and in case of any ambiguities the respective compounds were excluded from further consideration for the particular mode of action.

Chemical structures. In total, the set of active compounds comprised 1761 different chemicals. 359 Chemicals contained more than one molecule and/or ions. Mostly, this concerned hydrates and hydrohalogenides (mostly hydrochlorides), halogenides, sulfates, and ammonium salts. In each of these cases, the main compound supposed to trigger the endocrine disrupting effect has been extracted. These extracted structures were applied for all property calculations.

Modes of action. One chemical (proflavin) of the final set is active against all the five targets, there are 37 compounds active to four targets, 253 to three targets, 644 to two targets, and for the remaining 826 substances only one activity was observed. However, the absence of an activity to another target does not necessarily mean inactivity to this other target, because not all compounds were tested for each mode of action. With regard to the individual modes, 881 compounds act as thyroid receptor antagonists, 500 as AhR receptor (human liver) agonists, 303 as thyrotropin-releasing hormone receptor disruptors, 1064 as retinoic acid receptor disruptors, and 278 as RXR signalling pathway agonists.

There were no mode of action-specific differences in the performance of the bioavailability models however, and thus this report focusses on an overall investigation and calibration without regard to specific mode of actions.

The characterisation of the original chemicals, the chemical structures of the molecules applied for property calculations, and the thyroid hormone disrupting activities for the different modes of action are given in UFZ 2020, Tab. 1. The SMILES codes (Daylight 2019a) of the applied chemical structures are listed in UFZ 2020, Tab. 2.

Chemical domain. With regard to the chemical composition of the data set in terms of different atom types, there are 25 hydrocarbons, 31 halogenated hydrocarbons, 422 molecules additionally containing O atoms, 246 additionally containing N atoms, and 615 with both of them. Furthermore, 369 compound additionally contain S or P or both of them, and in 53 molecules other atom types can be found.

Concerning complexity, beside the 25 hydrocarbons without functional groups there are 323 compound with one single functional group, 189 chemicals with multiple occurrences of one type of functional groups, 590 compound with two different types of functional groups, and 631 chemicals with more than two types of functional groups. 3 substances contain special structures not matching to this scheme.

Regarding polarity, 109 chemicals are non- or weakly polar. 958 Compounds are polar with hydrogen bond donors, they typically but not necessarily also contain hydrogen bond acceptors. 608 polar compound have hydrogen bond acceptors but no donors, and there are 86 polar chemicals without hydrogen bond donors or acceptors.

The chemical domain is further characterised by an atom-centred fragments (ACF) approach (Kühne et al. 2009). The resulting ACF tables are available in ChemProp and listed in terms of SMARTS codes (Daylight 2019b) in UFZ 2020, Tab. 3 and Tab. 4. There are 2201 1st order and 9518 2nd order ACFs.

Lipinski's Rule-of-5

Original rule set. The original Rule-of-5 model (Lipinski et al. 1997) provides predictions whether poor absorption or permeation is expected for a chemical in a drug discovery setup. Enhanced probability for poor absorption/permeation is indicated if there are more than 5 H bond donors HD, if there are more than 10 H bond acceptors HA, if the molar weight MW exceeds 500 D, or if the logarithmic octanol/water partition coefficient log K_{ow} is higher than 5. When applying this rule set as a threshold, poor absorption or permeation is predicted if at least two of this criteria are fulfilled.

In the ChemProp (UFZ 2019) implementation, H bond donors are looked for as OH and NH groups, and H bond acceptors are considered to be any O and N atoms. Log K_{ow} is calculated by a consensus of several fragment methods from literature, and the molar weight calculation from the structure is straight-forward.

The application of this rule set to the 1761 active chemicals yields 1715 valid results, for the remaining 46 compounds no log K_{ow} could be estimated. These 46 chemicals (mostly inorganic compounds) were excluded from the further investigations. The calculated individual properties and the comparisons to the thresholds are given in UFZ 2020, Tab. 5.

In its original form the model already performs fairly well, but 96 of the 1715 chemicals (5.6 %) would be indicated as probably not bioavailable. In almost all cases, MW was one of the triggers. Only one compound (Carmin acid, $C_{22}H_{20}0O_{13}$) with the MW of 492 D contained 9 HD and 13 HA. However, 21 of these substances (1.2 % of all) fulfilled two additional conditions beside the molar weight.

The comparison to the individual triggers for all chemicals is shown in Fig. 1 and 2. The original threshold are exceeded for HA by 15 (1 %), for HD by 48 (3 %), for MW by 127 (7 %), and for log K_{ow} by 48 (3 %) chemicals.

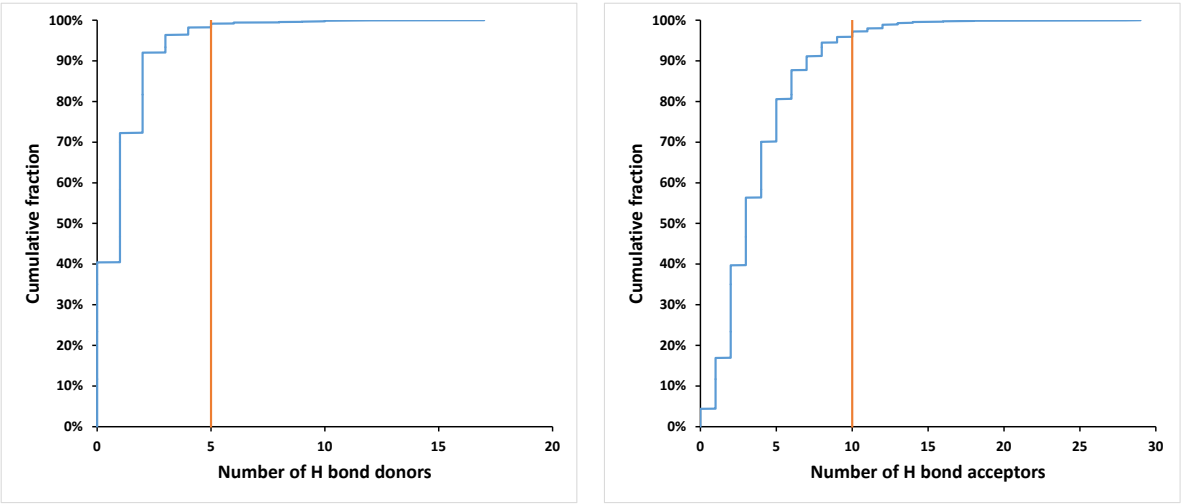

Fig. 1. Cumulative distribution of the number of H bond donors (left) and acceptors (right). The red lines denote the suggested bioavailability threshold of 5 or 10, respectively.

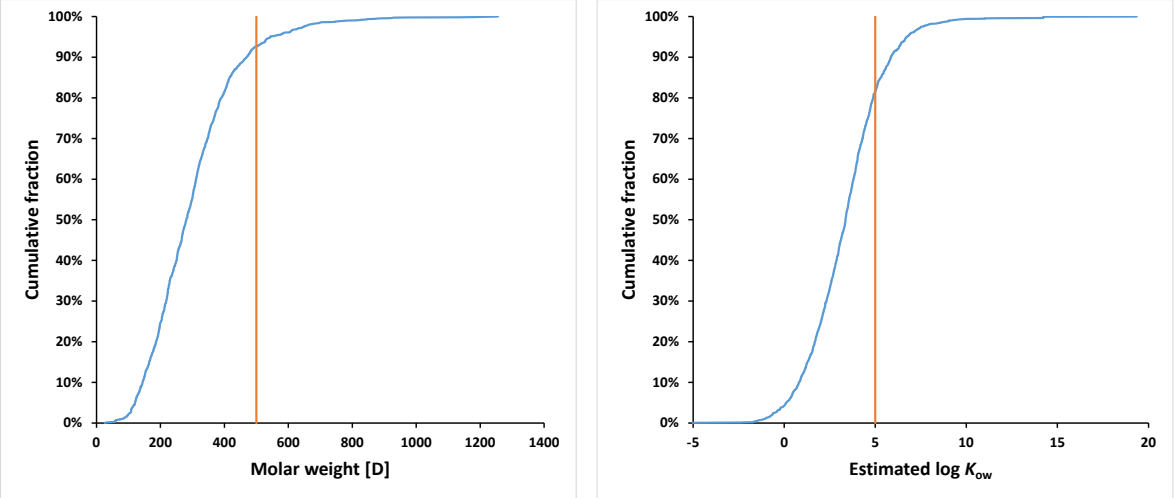

Fig. 2. Cumulative distribution of the molar weight (left) and the logarithmic octanol/water partition coefficient (right). The red lines denote the suggested bioavailability threshold of 500 D or 5, respectively.

Discussion and model adaption. Clearly, the molar weight limit of 500 D is the worst threshold with respect to thyroid hormone disruption bioavailability for thyroid hormone disruption. Increasing this threshold to 600 D will already reduce the number of unavailability prediction from 96 to 66 (i.e. from 5.6 % to 3.8 % of all). A threshold of 750 D yields 34 (2 % of all) predictions for unavailability.

Changing the log K_{ow} does not appear to be reasonable. Beside the rather moderate effect on decreasing the number of unavailability predictions, significantly higher partition coefficients tend to be unrealistic and may be artefacts of the estimation models to some extent. Increasing the HA or HB thresholds would also not significantly enhance the performance.

Interestingly, requiring the exceedance of three instead of 2 of the threshold also notably decreases the number of chemicals predicted to show poor absorption/permeation. Only 21 substances (1.2 %) would remain. Combined with a MW threshold increased to 600 D only 16 chemicals (0.9 %) remain, and with the 750 D threshold this number decreases to 9 compounds (0.5 %).

Coming back to the original model, some of the property values appear to be extremely high and may exceed the applicability domain of this approach. Even though there is no explicit domain characterisation available, inspection and comparison of the individual values suggest to define some upper limits. Any occurrence of a HA value above 15, HD above 20, MW above 1000 D, or log K_{ow} above 10 should indicate the rule set should not be applied for the respective chemical, and no prediction of bioavailability should be given in these cases. 13 chemicals (1 %) of the compounds exceed one or more of these limits.

The number of predictions for unavailability by the original model only decreases from 96 to 92 chemicals (5.4 % of all), but with the most conservative approach of 3 rule matches and MW > 750 only 7 chemicals (0.4 %) remain.

Other Bioavailability Approaches

Caco-2. The Caco-2 assay is an in vitro model for gastrointestinal absorption processes using cell lines from a human colon adenocarcinoma. It is supposed to address active and passive absorption processes. The quantitative outcome of Caco-2 assays is the apparent permeability Papp. It is a kinetic property in the dimension of length or path per time unit. According to literature (Pham-The 2013), P_{app} values below 0.7×10^{-6} cm/s are considered as low permeability, values of 0.7×10^{-6} cm/s < P_{app} < 16×10^{-6} cm/s as medium permeability, and P_{app} >16×10^{-6} cm/s as high permeability.

To estimate P_{app}, the approach suggested in BMU 2015 has been applied. Two models were available, one employing log K_{ow}, MW and the total polar surface area TPSA (Ertl et al. 2000)

$$\text{Log } P_{app} = 0.17 \log K_{ow} - 0.3 \text{ MW}/100 - 0.7 \text{ TPSA}/100 \quad (1)$$

with MW in D and TPSA in Å^2. The other model applies LSER parameters, also known as Abraham parameters H bond acidity A, H bond basicity B, and polarity/polarizability S calculated according to Platts et al. 1999.

$$\text{Log } P_{app} = 0.88\, A - 0.80\, B + 0.11\, S \quad (2)$$

Both model yield the logarithm of P_{app} with P_{app} in 10^{-6} cm/s. For a conservative approach, the maximum of both model results has been taken into account for each chemical.

Valid estimation results could be obtained or 1717 compounds. The results are listed in UFZ 2020, Tab. 6. Fig.3 (left) shows the cumulative distribution of the estimated Papp values together with the permeability threshold. The plot is limited at right to the most interesting part of the value range. As can be seen the majority of the values (1433 items, 83 %) are below the threshold, and the almost linear slope of the distribution does not point to any reasonable lower limit for an adapted threshold. As the Caco-2 assay is designed to mainly address oral and gastrointestinal uptake, it does not appear to be a suitable approach for thyroid hormone disruptor bioavailability. Furthermore, the performance of the P_{app} estimation model is at screening level only, and this limitation increases the uncertainty of potential bioavailability predictions.

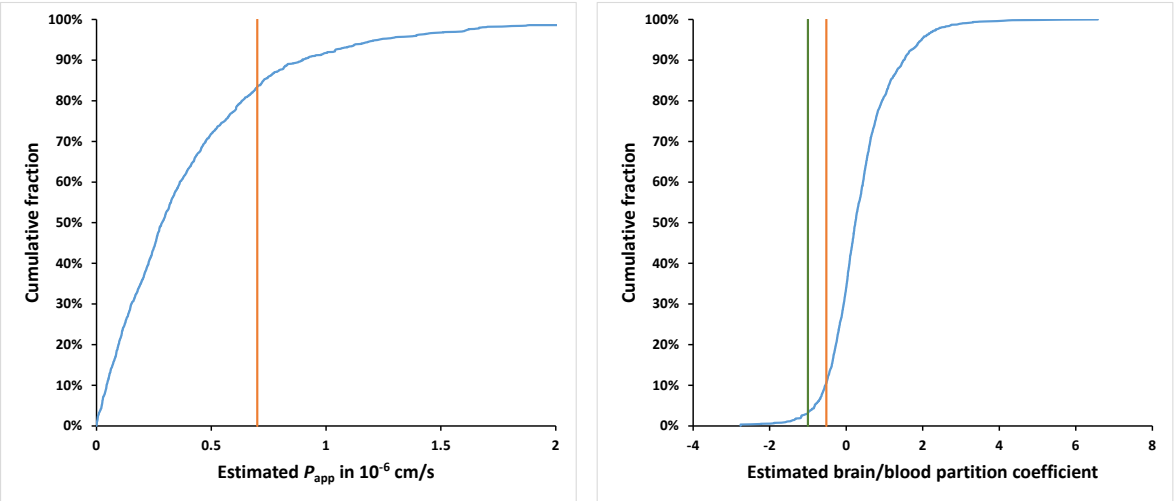

Fig. 3. Cumulative distribution of the Caco-2 apparent permeability P_{app} (left) and the logarithmic brain/blood partition coefficient (right). The red lines denote the suggested bioavailability threshold

of 0.7×10^{-6} cm/s or -0.52, respectively. The left plot is cropped at right, 100 % is reached at $9.3\ 7 \times 10^{-6}$ cm/s. In the right plot, the green lines gives a more conservative threshold of -1.

Blood brain barrier permeation. The blood brain barrier penetration is an important property in drug design for compounds targeting the central nervous system. The blood/brain (or brain/blood) partition coefficient provides a quantitative measure for it. In term of brain/blood partition, values above 30 %, i.e. -0.52 in logarithmic terms, are suggested as threshold for good brain penetration, but also values around 10% (-1 in logarithms) are discussed (Kunwittaya et al. 2013)

The estimation of the brain/blood partition coefficient employs a LSER (Abraham) model (Abraham et al. 2006)

$$- \log K_{brain/blood} = 0.526 + 0.185\ E + -0.596\ S + -0.623\ A + -0.630\ B + 0.630\ V + -1.210\ I_c \quad (3)$$

Here, *A*, *B*, and *S* again characterise the H bond acidity, H bond basicity, and polarity/polarizability. In addition, *E* is related to the electronic state, and *V* is a characteristic volume (Platts et al. 1999, Abraham et al. 1987). I_c is 1 for carboxylic acids and 0 otherwise.

The cumulative distribution of the 1693 valid results is shown at right in Fig. 3, the results are listed in UFZ 2020, Tab. 6. The threshold of -0.52, here similar to P_{app} applied as an upper limit, is exceeded for 176 chemicals (10 %), while 53 (3%) compounds are below -1. In conclusion, the logarithmic estimated blood/brain partition coefficient with the threshold of -1 can serve as a complementing approach to assess the thyroid hormone disruption specific bioavailability.

Combined approach. The prediction of good bioavailability can be further increased by combining the rule-of-5 approach with the brain/blood partition coefficient. When combining the original rule-of-5 model without domain limitations with the blood/blain partitioning coefficient threshold already leaves only 6 remaining compounds (0.4 %) considered as not bioavailable. Application of the adapted rule-of-5 model including the domain limitations together with the blood/brain threshold considers all active chemicals as bioavailable and thus no false negatives remain. There may still be false positives, however this would be less critical in terms of a conservative (i.e. save with regard to absence of thyroid disruption) approach.

Recommendation. In practice, if either the adapted rule-five model or the brain/blood partitioning models predict bioavailability, thyroid hormone specific bioavailability is assumed. Only if both model do not predict bioavailability, the compound can be considered as not reaching its target organ for a possible effect.

References

Abraham MH, Ibrahim A, Zhao Y, Acree WEJ 2006. A data base for partition of volatile organic compounds and drugs from blood/plasma/serum to brain, and an LFER analysis of the data. *J. Pharm. Sci.* 95: 2091-2100.

Abraham MH, McGowan JC 1987. The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography. *Chromatographia* 23: 243-246.

BMU 2015.
https://www.bmu.de/fileadmin/Daten_BMU/Pools/Forschungsdatenbank/fkz_3711_63_405_1_lipophilie_bioakkumulation_bf.pdf
Accessed Dec 2020.

Daylight 2019a.
https://daylight.com/smiles/
Accessed Dec 2020.

Daylight 2019b.
https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html
Accessed Dec 2020.

Ertl P, Rohde B, Selzer P 2000. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* 43: 3714-3717.

Kühne R, Ebert R-U, Schüürmann G 2009. Chemical domain of QSAR models from atom-centered fragments. *J. Chem. Inf. Model.* 96: 2660-2669

Kunwittaya S, Nantasenamat C, Treeratanapiboorn L, Srisarin A, Isaranka.Na-Ayudhuya C, Prachaysittikul V 2013. Influence of logBB cut-off on the prediction of blood-brain barrier permeability. *Biomedical and Applied Technology Journal*. 1: 16-34

Lipinski CA, Lombardo F, Dominy BW, Feeney PL 1997. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliver. Rev.* 23: 3-25

Pham-The H, González-Álvarez I, Bermejo M, Garrigues T, Le-Thi-Thu H, Cabrera-Pérez MÁ. 2013. The use of rule-based and QSPR approaches in ADME profiling: A case study on Caco-2 permeability. *Mol. Inf.* 32: 459-479.

Platts JA, Butina D, Abraham MH, Hersey A 1999. Estimation of molecular linear free energy relation descriptors using a group contribution approach. *J. Chem. Inform. Comput. Sci.* 39: 835-845.

UFZ Department of Ecological Chemistry 2019. ChemProp 6.7.1
http://www.ufz.de/ecochem/chemprop

UFZ Department of Ecological Chemistry 2020. In silico bioavailability triggers applied to thyroid hormone disrupters. Supporting Information. *To be published. Currently available to the consortium and to the EU office on request.*