Accepted Manuscript

Submission Date: 2024-07-29
Accepted Date: 2025-01-05
Accepted Manuscript online: 2025-01-23

Planta Medica

# Characterization of lipophilicity and blood partitioning of pyrrolizidine alkaloids and their N-oxides in vitro and in silico for toxicokinetic modeling

**Anja Lehmann, Manuel Haas, Julian Taenzer, Gerd Hamscher, Charlotte Kloft, Anja These, Christoph Hethey.**

Affiliations below.

**DOI: 10.1055/a-2523-3987**

**Please cite this article as:** Lehmann A, Haas M, Taenzer J et al. Characterization of lipophilicity and blood partitioning of pyrrolizidine alkaloids and their N-oxides in vitro and in silico for toxicokinetic modeling. Planta Medica 2025. doi: 10.1055/a-2523-3987

**Conflict of Interest:** Charlotte Kloft reports funding grants from an industry consortium (AbbVie Deutschland GmbH & Co. KG, Astra Zeneca, Boehringer Ingelheim Pharma GmbH & Co. KG, Grünenthal GmbH, F. Hoffmann-La Roche Ltd., Merck KGaA, Novo Nordisk and Sanofi) for the PharMetrX PhD program and from the Innovative Medicines Initiative-Joint Undertaking (‚DDMoRe'). The other authors declare no conflict of interest.

**Abstract:**
Lipophilicity and blood partitioning are important determinants for predicting toxicokinetics using physiologically-based toxicokinetic (PBTK) modeling. In this study, the logarithm of the n-octanol:water partition coefficient (logP) and the blood-to-plasma concentration ratio (R<sub>b</sub>) were for the first time experimentally determined for the pyrrolizidine alkaloids (PAs) intermedine, lasiocarpine, monocrotaline, retrorsine and their N-oxides (PANOs). Validated in vitro assays for logP (miniaturized shake-flask method) and R<sub>b</sub> (LC-MS/MS-based depletion assay) determination were compared to an ensemble of in silico models. Experimentally determined logP indicate a higher affinity of PAs and PANOs to the aqueous compared to the organic phase. Depending on the method, in silico determined logP overpredicted the experimental values by ≥ 1 log unit for 3 out of 4 PAs (SPARC), 4 out of 6 PAs and PANOs (CLOGP), 5 out of 8 PAs and PANOs (KowWIN) and 3 out of 8 PAs and PANOs (S+logP). R<sub>b</sub> obtained in vitro suggested a low binding affinity of PAs and PANOs towards red blood cells. For all 8 PAs and PANOs, in silico predicted R<sub>b</sub> deviated from experimental values by less than 50%. In conclusion, for PBTK modeling of PAs and PANOs, experimental logP should be preferred, while R<sub>b</sub> predicted by the acid/base classification model is a suitable surrogate for experimental data.

**Corresponding Author:**
Anja Lehmann, Freie Universität Berlin, Clinical Pharmacy and Biochemistry, Berlin, Germany, anja.lehmann@fu-berlin.de

**Contributors' Statement:** Conceptualization: A. Lehmann and C. Hethey; Methodology: A. Lehmann, M. Haas, A. These and C. Hethey; Validation: A. Lehmann, M. Haas, A. These and C. Hethey; Formal analysis: A. Lehmann and M. Haas; Investigation: M. Haas and J. Taenzer; Data curation: A. Lehmann; Writing - original draft: A. Lehmann; Writing – review & editing: M. Haas, J. Taenzer, G. Hamscher, C. Kloft, A. These and C. Hethey; Visualization: A. Lehmann; Supervision: G. Hamscher, C. Kloft, A. These and C. Hethey; Project administration: A. These

**Affiliations:**

Anja Lehmann, Freie Universität Berlin, Clinical Pharmacy and Biochemistry, Berlin, Germany
Anja Lehmann, German Federal Institute for Risk Assessment, Berlin, Germany
Manuel Haas, German Federal Institute for Risk Assessment, Berlin, Germany
[...]
Christoph Hethey, German Federal Institute for Risk Assessment, Berlin, Germany

# Characterization of lipophilicity and blood partitioning of pyrrolizidine alkaloids and their *N*-oxides *in vitro* and *in silico* for toxicokinetic modeling

**Anja Lehmann[1,2,**], Manuel Haas[1,**], Julian Taenzer[1], Gerd Hamscher[3], Charlotte Kloft[2], Anja These[1,**], Christoph Hethey[1,**]**

**Affiliation**

[1] German Federal Institute for Risk Assessment (BfR), Berlin, Germany

[2] Department of Clinical Pharmacy and Biochemistry, Institute of Pharmacy, Freie Universitaet Berlin, Berlin, Germany

[3] Institute of Food Chemistry and Food Biotechnology, Justus Liebig University Giessen, Giessen, Germany

**Correspondence**

***Anja Lehmann***

Department of Clinical Pharmacy and Biochemistry

Institute of Pharmacy

Freie Universitaet Berlin

Kelchstr. 31

12169 Berlin

Germany

anja.lehmann@fu-berlin.de

** Authors contributed equally.

**Abstract**

Lipophilicity and blood partitioning are important determinants for predicting toxicokinetics using physiologically-based toxicokinetic (PBTK) modeling. In this study, the logarithm of the *n*-octanol:water partition coefficient (log*P*) and the blood-to-plasma concentration ratio ($R_b$) were for the first time experimentally determined for the pyrrolizidine alkaloids (PAs) intermedine, lasiocarpine, monocrotaline, retrorsine and their *N*-oxides (PANOs). Validated *in vitro* assays for log*P* (miniaturized shake-flask method) and $R_b$ (LC-MS/MS-based depletion assay) determination were compared to an ensemble of *in silico* models. Experimentally determined log*P* indicate a higher affinity of PAs and PANOs to the aqueous compared to the organic phase. Depending on the method, *in silico* determined log*P* overpredicted the experimental values by ≥ 1 log unit for 3 out of 4 PAs (SPARC), 4 out of 6 PAs and PANOs (CLOGP), 5 out of 8 PAs and PANOs (KowWIN) and 3 out of 8 PAs and PANOs (S+logP). $R_b$ obtained *in vitro* suggested a low binding affinity of PAs and PANOs towards red blood cells. For all 8 PAs and PANOs, *in silico* predicted $R_b$ deviated from experimental values by less than 50%. In conclusion, for PBTK modeling of PAs and PANOs, experimental log*P* should be preferred, while $R_b$ predicted by the acid/base classification model is a suitable surrogate for experimental data.

**Keywords:** Log*P*, Octanol:water partition coefficient, Blood-to-plasma ratio, Drug distribution, Pharmacokinetics, PBTK modeling

## Introduction

Pyrrolizidine alkaloids (PAs) are a large group of secondary plant metabolites of which the 1,2-unsaturated PAs are of particular concern due to their hepatotoxicity, genotoxicity and potential carcinogenicity [1]. PAs and their corresponding *N*-oxides (PANOs) usually coexist in plants, while the latter are reduced to toxic PAs in both intestine and liver [2]. The relative toxicity of structurally different PA congeners is substantially influenced by their individual kinetics, i.e. by their absorption, distribution, metabolism and excretion (ADME) characteristics [3]. Physiologically-based toxicokinetic (PBTK) modeling is a powerful tool to explore the ADME characteristics of PAs within an organism. PBTK models have been successfully used for PAs to quantitatively predict *in vivo* liver toxicity and *in vivo* genotoxicity from *in vitro* toxicity data [4,5]. In conjunction with other physico- and biochemical properties, lipophilicity and blood partitioning are important determinants in PBTK models for predicting whole body toxicokinetics.

Since the *in vivo* analysis of tissue distribution requires animal testing that apart from ethical concerns is costly and time intensive, mechanistic models for *a priori* prediction of tissue distribution have been established for PBTK modeling. In these mechanistic approaches, e.g. by Rodgers and Rowland [6,7] or by Schmitt [8,9], the composition of the different tissues and the interaction of compounds with the tissue constituents is taken into consideration. For prediction of compound partitioning into the tissue lipids, a measure of lipophilicity is needed as parameter in the models.

The *n*-octanol:water partition coefficient $K_{o:w}$ is a measure of lipophilicity in absence of relevant speciation and refers to the neutral form of a substance. $K_{o:w}$ is defined as the concentration ratio in an organic phase, typically *n*-octanol, and an aqueous phase:

$$K_{o:w} = P = \frac{C_{organic}}{C_{aqueous}} \quad (1)$$

$K_{o:w}$ is usually expressed as base 10 logarithm log*P*.

For ionizable substances, where neutral and charged species co-exist to relevant extent in solution, the corresponding distribution ratio *D* specifically accounts for neutral and ionized species [10]:

$$D = \frac{C_{neutral,organic} + C_{ionized,organic}}{C_{neutral,aqueous} + C_{ionized,aqueous}} \quad (2)$$

Since the degree of ionization depends on pH, so does *D*. *D* is usually given as base 10 logarithm log*D*. The distribution coefficient log*D* can be calculated for a specific pH from log*P* and the acid dissociation constant p*K*a using the Henderson-Hasselbalch relation:

$$\log D(pH) = \log P - \log\left(1 + 10^{(pH - pKa)}\Delta i\right) \quad (3)$$

with Δi = 1 for acids and Δi = -1 for bases [11].

PAs are weak bases, since the nitrogen atom of the tertiary amine becomes protonated under neutral and acidic conditions. PANOs also show weak basic behavior attributed to protonation of the oxygen atom of the *N*-oxide under acidic conditions. The weak basic characteristic of PAs and PANOs is confirmed by *in silico* predicted p*K*a values (see Table 1S).

Experimentally, the log*P* value can be determined using direct or indirect methods [12]. Direct experimental methods, such as the OECD shake-flask method [13] or miniaturized adaptations [14], obtain log*P* directly from the concentration ratio of the compound partitioned between the organic and aqueous phase. In indirect experimental methods, like the reversed-phase high-performance liquid chromatography (RP-HPLC) [15], log*P* is estimated through the compound retention in a hydrophobic stationary phase. During experimental log*P* determination for ionizable substances, the impact of pH on the ionization state needs to be considered so that partitioning exclusively refers to their neutral form. In addition, a vast number of *in silico* methodologies have been developed to predict lipophilicity. While substructure-based models derive log*P* from cutting molecules into fragments or down to the

single-atom level, property-based models use empirical approaches, the 3D-structure of a molecule or topological descriptors [16]. The lipophilicity of a compound is related to many biological properties such as solubility and membrane permeability, thereby strongly affecting *in vivo* compound ADME characteristics, its pharmacodynamic and toxicological profile [17].

Certain compounds are highly bound to or distributed into blood cells. Like plasma protein binding, this can significantly reduce a compound's free plasma concentration, i.e. the fraction that is available for its pharmacological or toxicological action and its elimination. Blood partitioning, therefore, can strongly determine the toxicokinetic and -dynamic profile of a compound. The blood-to-plasma concentration ratio ($R_b = C_{blood}/C_{plasma}$) is a measure of compound distribution within whole blood at equilibrium [18].

The importance of a precisely determined $R_b$ value can be illustrated on the example of the drugs thioridazine and maprotiline. For thioridazine, which shows a $R_b$ value of 0.55 [19] and therefore almost no distribution into blood cells, thioridazine exposure calculated from plasma equals twice the exposure calculated from whole blood. For maprotiline, which is distributed into blood cells showing a $R_b$ value of 2.1 [19], maprotiline exposure determined from plasma is half the exposure determined from whole blood. In the latter case, clearance calculated from plasma data would highly overestimate whole blood clearance if $R_b$ was not considered and could even exceed hepatic blood flow.

Experimentally, the $R_b$ value can be obtained *in vitro* by separate measurement of compound concentrations in equilibrating plasma and whole blood (or blood cells), or more laborious *in vivo* within a pharmacokinetic study [20]. In addition, *in silico* methodologies for $R_b$ prediction have been developed. Some methodologies are based on mechanistic prediction of distribution into red blood cells [6,8], while others comprise regression models that relate a variety of molecular descriptors to the $R_b$ value using statistical algorithms such as partial

least squares regression and artificial neural networks [18,21,22]. In absence of experimental data, it is often assumed $R_b$ = 1 for neutrals and bases, and $R_b$ = 1-hematocrit = 0.55 for acids and zwitterions (acid/base classification) [22].

Due to the need to characterize lipophilicity and blood partitioning of PAs for PBTK modeling, the aim of this study was to determine log*P* and $R_b$ values of representative 1,2-unsaturated PA congeners. Therefore, PAs of different ester types were analyzed including the monoester intermedine, the cyclic diesters monocrotaline and retrorsine, and the open-chained diester lasiocarpine. Additionally, the corresponding PANOs were included in the study. Log*P* and human $R_b$ values were determined *in vitro* using established experimental methods and compared to published literature values. The performance of established *in silico* models for prediction of log*P* and $R_b$ values was assessed for PAs and PANOs. Log*P* values were predicted using a substructure- and property-based model and $R_b$ values were predicted by an acid/base classification model and a mechanistic model. Finally, *in silico* predictions were evaluated against experimental values.

**Results**

Log*P* values were determined by bidirectional equilibration from *n*-octanol to water and vice versa (see section *Determination of n-octanol:water partition coefficient*). Method evaluation with reference substances showed convincing linear correlation ($R^2$ = 0.988) between determined log*P* values and literature log*P* values (Fig. 1aS). Except for warfarin, determined log*P* values did not differ by more than 0.2 log units from those reported in literature demonstrating a high accuracy (Table 2S). The experimental method was precise as indicated by relative standard deviations below 10% for sulfanilamide, tolbutamide and warfarin (caffeine: 20%, theophylline: 47%).

For all PAs and PANOs, measured log*P* values were negative and ranged between -1.93 (intermedine) and -0.302 (lasiocarpine *N*-oxide) (Table 1 and Fig.1, circles). There were no significant differences in log*P* values depending on the direction of equilibration, i.e. starting from the organic phase or starting from the aqueous phase. Log*P* values of PANOs were between 0.02 log units (retrorsine *N*-oxide) and 0.2 log units (lasiocarpine *N*-oxide) higher compared to log*P* values of respective PAs.

Log*P* predicted by KowWIN deviated from experimental values by < 1 log unit for monocrotaline, monocrotaline *N*-oxide and retrorsine *N*-oxide and by ≥ 1 log unit for the other PAs and PANOs (Table 1 and Fig. 1, asterisks). Log*P* calculated by the CLOGP method overpredicted experimental values by < 1 log unit for retrorsine and retrorsine *N*-oxide and by ≥ 1 log unit for the other PAs and PANOs. With regard to SPARC, log*P* were overpredicted by < 1 log unit for monocrotaline and by ≥ 1 log unit for the other PAs. Note that log*P* predictions were not available for PANOs in SPARC. Log*P* predicted by the S+logP method (ADMET Predictor) deviated from experimental values by < 1 log unit for all PANOs and retrorsine and by ≥ 1 log unit for intermedine, lasiocarpine and monocrotaline.

The $R_b$ value was determined using the LC-MS/MS-based depletion assay that measures the compound concentration (i) in plasma that has been equilibrating with red blood cells and (ii) in a plasma reference (see section *Determination of blood-to-plasma concentration ratio*). The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Charité University Medicine Berlin (EA4/183/19, 12 November 2019). Method evaluation with reference substances showed good linear correlation ($R^2 = 0.908$) between determined $R_b$ and literature $R_b$ (Fig. 1bS). Except for imipramine and warfarin, determined $R_b$ did not deviate by more than 10% from previously published $R_b$, which shows the high accuracy of the method (Table 2S). Relative standard deviations below 10% indicate that the experimental method was precise. An effect

of the compound concentration on $R_b$ was not observed with the range of 4 to 1000 ng/mL, exemplarily determined for imipramine, caffeine and warfarin (Fig. 2S).

Measured $R_b$ values of PAs and PANOs were ranging between 0.689 (intermedine *N*-oxide) and 1.12 (intermedine) (Table 2 and Fig. 2, circles). The $R_b$ value of the PAs intermedine, monocrotaline and retrorsine was identical ($R_b$ = 1.1). Except for lasiocarpine, determined $R_b$ values of PAs were 30% to 40% higher compared to that of their corresponding *N*-oxide. Within the range of 4 to 1000 ng/mL no concentration-dependency was observed, exemplarily determined for retrorsine and its *N*-oxide (Fig. 3a and b).

*In silico* predictions of $R_b$ according to the acid/base classification model deviated from experimentally determined $R_b$ by less than 10% for monocrotaline and retrorsine and by more than 10% but less than 50% for the other PAs and PANOs showing a tendency for over prediction (Table 2 and Fig. 2, asterisks). With regard to the mechanistic model, *in silico* predictions of $R_b$ deviated from experimentally determined $R_b$ by less than 10% for monocrotaline *N*-oxide and by more than 10% but less than 50% for the other PAs and PANOs with a tendency for under prediction.

**Discussion**

We here report for the first time lipophilicity and blood partitioning of representative PAs and their corresponding *N*-oxides, both being important predictors of compound ADME characteristics and key parameters in PBTK models. Measures of lipophilicity and blood partitioning, the logarithm of the *n*-octanol:water partition coefficient (log*P*) and the blood-to-plasma concentration ratio ($R_b$), were determined using published experimental assays as well as established *in silico* methods.

For all PAs and PANOs, log*P* values were negative showing that the compounds have a higher affinity to the aqueous phase. The observed lipophilicity of analyzed PAs can be

ranked as intermedine < monocrotaline < retrorsine < lasiocarpine. Regarding the PA ester type, these results suggest the following lipophilicity order, which is however limited by the small number of PAs assessed: monoester < cyclic diester < open-chained diester. In reversed-phase liquid chromatography, log*P* is correlated with the retention time, i.e. hydrophilic compounds elute first and lipophilic last [15]. Our observed lipophilicity ranking is supported by the elution order of 22 structurally different PAs in the reversed-phase chromatogram [23]: In tendency, monoesters and cyclic diesters have shorter retention times, which implies that they have a lower lipophilicity compared to open-chained diesters. The experimental results indicate, except for lasiocarpine *N*-oxide, that PANOs are slightly more lipophilic compared to their respective PAs. This is opposed to *N*-oxidation of pyridines, aromatic weak bases with a chemical structure related to PAs, which was reported to increase hydrophilicity [24].

Log*P* values calculated using *in silico* methodologies showed a trend for overprediction for PAs and PANOs. Predictions by substructure-based methods deviated upward from the experimental values by ≥ 1 log unit for 5 out of 8 (63%) PAs and PANOs (KowWIN) and 4 out of 6 (67%) PAs and PANOs (CLOGP). Predictions by property-based methods deviated upwards by ≥ 1 log unit for 3 out of 4 (75%) PAs (SPARC) and 3 out of 8 (38%) PAs and PANOs (S+logP). The best performance was achieved with S+logP, which showed particularly good results for all PANOs, but not PAs.

For comparison, KowWIN, SPARC and S+logP (CLOGP not available) showed better accuracy of log*P* for reference substances than for PAs and PANOs, i.e. deviations were < 1 log unit for 5 out of 5 (100%, KowWIN) and 4 out of 5 (80%, SPARC, S+logP) reference substances (also see Fig. 3S). The accuracy of log*P* prediction methods is strongly dependent on the compounds included in a training dataset and is limited by the complexity of molecules, in particular by their size [16].

Exposure-related kinetic parameters, such as clearance, especially in relation to blood flow and volume of distribution can be misleading if interpreted based on plasma instead of whole blood data, which can be prevented if the $R_b$ value of a compound is known. $R_b$ values below 1 ($C_{plasma} > C_{blood}$) indicate stronger binding to plasma proteins than to red blood cells, while $R_b$ values above 1 ($C_{blood} > C_{plasma}$) show that binding to blood cells is larger than plasma protein binding [18]. For PAs and PANOs, determined $R_b$ values were in the range between 0.689 and 1.12 with $R_b$ values of *N*-oxides in tendency being smaller compared to PAs. These results suggest that PAs and PANOs have a low binding affinity towards red blood cells and distribution within red blood cells is not expected. This is in line with the absence of a concentration-dependency of the $R_b$ value as was exemplarily shown for retrorsine and its *N*-oxide. By contrast, for compounds that are highly bound or distributed into red blood cells concentration-dependent saturation of binding or active uptake would have been expected [25,26].

In the absence of experimental data, the $R_b$ value is often assumed to be 1 for neutrals and bases or 0.55 (1-hematocrit) for acids and zwitterions [22]. Since PAs are weak bases and PANOs are neutrals under physiological conditions, their $R_b$ value is 1 according to the acid/base classification model. While predictions by this acid/base classification model were slightly superior to predictions by the mechanistic tissue distribution model, both methods showed predictions close to experimental values with deviations below 1.5-fold for all PAs and PANOs. In particular, $R_b$ values deviated from the experimental values by less than 10% for 2 out of 8 (acid/base classification model) and 1 out of 8 (mechanistic model) PAs and PANOs, and by more than 10% but less than 50% for 6 out of 8 (acid/base classification model) and 7 out of 8 (mechanistic model) PAs and PANOs. It was reported that the performance of more complex artificial neural network (ANN) models was superior to that of mechanistic models for tissue distribution [18,22]. However, the application of ANN models

is less feasible, since in addition to log*P* and fraction unbound in plasma $f_{u,p}$ more compound-specific prior information (> 10 molecular descriptors) is required.

For PAs and PANOs, we recommend predicting tissue distribution with experimentally determined log*P* values, since the outlined deviations of *in silico* predictions from validated *in vitro* measurements were ≥1 log unit. A direct experimental method for log*P* determination, such as the miniaturized shake-flask method used in this study, should be preferred over indirect methods to avoid erroneous results [27]. While *in silico* prediction of log*P* values was insufficient, the *in silico* predicted $R_b$ value using the acid/base classification model (deviations below 1.5-fold) can be regarded as suitable surrogate for PAs and PANOs when experimental data on blood partitioning are not available.

## Material and Methods

*Material*

Intermedine (purity > 95%) and lasiocarpine (purity > 97%) were purchased from Phytoplan Diehm & Neuberger GmbH. Retrorsine (purity ≥ 90%) was obtained from Sigma-Aldrich Chemie GmbH. Monocrotaline (purity ≥ 90%), intermedine *N*-oxide (≥ 90%), monocrotaline *N*-oxide (≥ 90%) and retrorsine *N*-oxide (purity ≥ 90%) were received from Phytolab GmbH & Co. KG. Lasiocarpine *N*-oxide (purity 96.7%) was purchased from Cfm Oskar Tropitzsch GmbH. Caffeine (purity ≥ 99%), imipramine hydrochlorid (purity ≥ 99%), sulfanilamide (purity ≥ 99%), theophylline (purity ≥ 99%), tolbutamide (purity 99.8%), warfarin (purity ≥ 98%) and quinidine (purity ≥ 80%) were obtained from Sigma-Aldrich. 1-Octanol Chromasolv (purity ≥ 99%) was purchased from Fisher Scientific GmbH. Water LiChrosolv, acetonitrile LiChrosolv (purity ≥ 99.9%) and methanol LiChrosolv (purity ≥ 99.9%) were obtained from Merck KGaA. Ammonium formate Chromanorm (purity ≥ 99%) was provided by VWR International GmbH.

*Collection of whole blood and plasma samples*

Blood samples were collected from two healthy donors (female, 44 years; male, 25 years). Collection of blood samples was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the Charité University Medicine Berlin (EA4/183/19, 12 November 2019). Whole blood was collected from cubital vein into blood collection tubes with heparin (S-Monovette Lithium heparin, Sarstedt AG & Co. KG). To obtain plasma, collection tubes were centrifuged for 10 min at 11,180 g and 8°C and the plasma was separated from blood cells. Whole blood and plasma samples were stored at 4°C until analysis.

*Determination of the n-octanol:water partition coefficient P*

The *n*-octanol:water partition coefficient *P* was determined according to the miniaturized shake-flask method [14]. Presaturated solutions of *n*-octanol and water were prepared. Therefore, equal volumes of *n*-octanol and water were mixed thoroughly in a separating funnel. After equilibration for 24 h, the organic and aqueous phases were separated for further use. The compounds were dissolved either in presaturated water or in presaturated *n*-octanol to measure bidirectional partitioning from the aqueous to the organic phase and vice versa. 350 µL of compound solution were added to a reaction tube and volume was made up to 500 µL with the same phase to achieve a concentration of 1,000 ng/mL. 500 µL of the other phase were added to obtain a total volume of 1 mL. Reaction tubes were shaken overnight at 500 rpm and 22°C. After centrifugation for 20 min at 16,099 g and 20°C, separated phases were removed using a glass syringe to avoid phase contamination. 200 µL of each sample were transferred to glass vials and analyzed by LC-

MS/MS. Experiments were performed in triplicate for equilibration from the organic to the aqueous phase ($n$ = 3) and from the aqueous to the organic phase ($n$ = 3).

Log*P* values were calculated as logarithm of the mass-spectrometric peak area ratio (Eq. 4). Analyte concentrations were assumed proportional to peak areas since volumes of aqueous and organic phase were identical and samples were not diluted.

$$\log P = \log \frac{Peak\ area(Analyte \in n-octanol)}{Peak\ area(Analyte \in water)} \quad (4)$$

The experimental method was validated with the five structurally different reference substances caffeine, sulfanilamide, theophylline, tolbutamide and warfarin. Linear correlation between determined and literature log*P* values was analyzed visually and numerically using the coefficient of determination $R^2$. PAs and PANOs, as well as reference compounds were neutral (≥ 96%) under experimental conditions (see Table 1S).

The performance of *in silico* methodologies for prediction of the log*P* value was assessed against experimentally determined log*P* values for PAs and PANOs. To this end, two substructure-based prediction methods, KowWIN [28] and CLOGP (values published in [29]), were compared to two property-based methods, SPARC [30] and S+logP (ADMET Predictor [31]).

*Determination of blood-to-plasma concentration ratio*

Conventionally, the partitioning of a compound between blood and plasma is determined by separate analysis of compound concentrations in equilibrating plasma and whole blood or red blood cells. Therefore, separate standards in their respective matrix, i.e. in plasma and whole blood or red blood cell lysates are used. In the here used LC-MS/MS-based depletion assay, the blood-to-plasma concentration ratio $R_b$ was determined by measurement of compound concentrations in the plasma sample against a defined plasma reference [19].

A whole blood sample and a plasma reference sample of the same volume were spiked with the compound to achieve a final concentration of 1,000 ng/mL. Additional concentrations were assessed at 4, 20, 100 and 500 ng/mL to analyze concentration dependency of $R_b$. The whole blood was incubated on a shaker for 60 min at 37°C to allow equilibration of the compound between plasma and red blood cells. Then, the whole blood was centrifuged for 10 min at 11,180 g and 8°C to yield the plasma. The plasma reference was treated the same way as the whole blood regarding incubation and centrifugation. Then, 50 µL of plasma were collected from the separated whole blood and from the plasma reference, respectively. The collected plasma samples were quenched with 450 µL of ice-cold acetonitrile, mixed with a vortex mixer and centrifuged at 21,913 g for 10 min to remove plasma protein. Supernatants were collected and analyzed using LC-MS/MS. Experiments for PAs and PANOs were performed with $n$ = 6 for PAs and PANOs and $n$ = 4 for reference substances.

The $R_b$ value was calculated using the analyte peak area responses of the plasma samples (Eq. 5). Since volumes and matrix of the reference plasma sample and the equilibrating plasma sample were identical, analyte concentrations are proportional to the peak areas.

$$R_b = \frac{Peak\ area(Analyte \in plasma\ reference)}{Peak\ area(Analyte \in equilibrating\ plasma)} \quad (5)$$

The experimental method was validated with the six structurally diverse reference substances caffeine, imipramine, theophylline, tolbutamide, warfarin and quinidine. Linear correlation between determined and literature $R_b$ was assessed visually and numerically ($R^2$).

The performance of *in silico* models for of the $R_b$ value was evaluated against the experimentally determined $R_b$ values for PAs and PANOs. To this end, an acid/base classification model and a mechanistic model for $R_b$ value prediction were used. In the acid/base classification model, the $R_b$ value is assumed 1 for neutrals and bases or 0.55 (1-hematocrit) for acids and zwitterions [22]:

$$R_b = \begin{cases} 1 \; for \; neutrals \wedge bases \\ 0.55 \; for \; acids \wedge zwitterions \end{cases} \quad (6)$$

PAs and PANOs are weak bases (see Table 1S for p*K*a predictions). Under physiological conditions, i.e. at pH 7.4, PAs are present as bases, while PANOs are uncharged and therefore present as neutrals. Hence, according to the acid/base classification model PAs and PANOs are assigned an $R_b$ value of 1.

In the mechanistic model, $R_b$ values are calculated from red blood cell-to-unbound plasma partition coefficient $K_{rbc,pla}$ by the Open Systems Pharmacology (OSP) Suite (https://www.open-systems-pharmacology.org/, accessed 28 Oct 2023) based on a mechanistic tissue distribution model. In this model, compound distribution into red blood cells is approximated by the distribution into the cellular constituents water, lipids and proteins, instead:

$$R_b = (K_{rbc,pla} \cdot f_{u,p} - 1) \cdot hct - 1 \quad (7)$$

$$K_{rbc,pla} = (f_{rbc,water} + f_{rbc,lipids} \cdot 10^{\log P} + f_{rbc,prot} \cdot K_{prot}$$

with $K_{rbc,pla}$ = red blood cell-to-unbound plasma partition coefficient; $f_{u,p}$ = fraction unbound in plasma; hct = hematocrit; $f_{rbc,water}$ = volume fraction of red blood cells that is water; $f_{rbc,lipids}$ = volume fraction of red blood cells that is lipids; log*P* = logarithm to base 10 of the *n*-octanol:water partition coefficient; $f_{rbc,prot}$ = volume fraction of red blood cells that is proteins; $K_{prot}$ = water-protein partition coefficient.

Log*P* values were experimentally determined (see Table 1). $f_{u,p}$ values were informed by our in-house dataset (intermedine: 0.890, intermedine *N*-oxide: 1.08, retrorsine: 0.600, retrorsine *N*-oxide: 0.960; rapid equilibrium dialysis) or were predicted (lasiocarpine: 0.227, lasiocarpine *N*-oxide: 0.237, monocrotaline: 0.456, monocrotaline *N*-oxide: 0.461 [32]).

*Sample analysis by liquid chromatography and high-resolution mass spectrometry*

The chromatographic measurements were performed using an UltiMate 3000 Ultra-High-Performance Liquid Chromatography (UHPLC) system (Thermo Fisher Scientific) with a C18 Hypersil Gold column (150 mm x 2.1 mm; 1.9 µm particle size) and guard column (Thermo Fisher Scientific). The column temperature was set at 40°C, while the flow rate and injection volume were adjusted to 0.3 mL/min and 2 µL, respectively. UHPLC was carried out by a binary mobile phase consisting of water (A) and methanol (B) as mobile phases, whereby 0.1% formic acid and 5 mM ammonium formate in each mobile phase was integrated. The gradient system was configured as follows: 0-7.0 min A: 95%, 7.0-7.5 min A: 50%, 7.5-7.6 min A: 20%, 7.6-10.1 min A: 0%, 10.1-15 min A: 95%.

High-resolution mass spectrometry was conducted using an Orbitrap mass spectrometer (Thermo Fisher Scientific) in a positive ionization mode (precursor ions $[M+H]^+$) to determine ion scans of reference substances caffeine, imipramine, sulfanilamide, theophylline, tolbutamide, quinidine and warfarin. Ion scans (ddMS2) with collision energies in a range of 10-35 eV and a resolution of 17,500 were recorded, whereby three fragments (product ions) with the highest relative abundances of each substance were identified (Table 3S). Representative ddMS2 data of the reference substances achieved with 25 eV are shown in Fig. 4S-10S. For further analyses, samples containing reference substances were measured by data-independent acquisition (vDIA) consisting of full scans for quantification and MS2 data by fragmentation of four mass range windows (m/z: 50-150, 130-300, 280- 400 and 380-600) with a fixed collision energy of 25 eV for verification. Data were acquired and processed with TraceFinder 4.1 and Xcalibur 4.0 (Thermo Fisher Scientific).

Low-resolution mass spectrometry was carried out using Agilent 6495 Triple Quadrupole system combined with an Agilent 1290 Infinity II LC System (Agilent Technologies) to measure PAs and PANOs. Compounds were ionized by electrospray ionization (ESI) in a

positive mode and collision energies were individually adjusted. For identification, three product ions of each analyte were selected in each measurement (Table 3S). Data were acquired and processed with Mass Hunter Workstation (Agilent Technologies).
Guidance [33] criteria for recovery (70-120%) and intraday precision (≤ 20%) were tested and fulfilled for reference substances in matrices water, *n*-octanol and human plasma (also see Table 4S).

**Supporting Information**

Linear regression of log*P* values **(Fig 1aS)**, and $R_b$ values **(Fig 1bS)** of reference substances experimentally determined in this work versus literature, the effect of the concentration on the $R_b$ value of reference substances **(Fig. 2S)**, experimentally determined and predicted log*P* values of reference substances **(Fig. 3S)**, representative ddMS2 data of reference substances **(Fig. 4S-10S)**, predicted p*K*a values and predominant charge state at experimental pH (log*P* determination) of PAs, PANOs and reference substances **(Table 1S)**, log*P* values and $R_b$ values of reference substances **(Table 2S)**, MS/MS transitions and parameters **(Table 3S)** and mass spectrometric peak areas and calculated recovery of reference substances for log*P* determination **(Table 4S)** are available as Supporting Information.

**Contributors' statement**

Conceptualization: A. Lehmann and C. Hethey; Methodology: A. Lehmann, M. Haas, A. These and C. Hethey; Validation: A. Lehmann, M. Haas, A. These and C. Hethey; Formal analysis: A. Lehmann and M. Haas; Investigation: M. Haas and J. Taenzer; Data curation: A. Lehmann; Writing - original draft: A. Lehmann; Writing – review & editing: M. Haas, J. Taenzer, G. Hamscher, C. Kloft, A. These and C. Hethey; Visualization: A. Lehmann;

Supervision: G. Hamscher, C. Kloft, A. These and C. Hethey; Project administration: A. These

## Data availability

All data are available in the GitHub repository https://github.com/al901010/Supplement_Lipophilicity_BloodPartitoning_PAs.


## Acknowledgments

The authors thank Anja Gessele for her skillful technical support. Open Access funding enabled and organized by Projekt DEAL.


## Conflicts of Interest

Charlotte Kloft reports funding grants from an industry consortium (AbbVie Deutschland GmbH & Co. KG, Astra Zeneca, Boehringer Ingelheim Pharma GmbH & Co. KG, Grünenthal GmbH, F. Hoffmann-La Roche Ltd., Merck KGaA, Novo Nordisk and Sanofi) for the PharMetrX PhD program and from the Innovative Medicines Initiative-Joint Undertaking ('DDMoRe'). The other authors declare no conflict of interest.

## References


[1] Dusemund B, Nowak N, Sommerfeld C, Lindtner O, Schäfer B, Lampen A. Risk assessment of pyrrolizidine alkaloids in food of plant and animal origin. Food Chem Toxicol 2018; 115: 63–72

[2] Mengbi Y, Ma J, Ruan J, Ye Y, Fu P, Lin G. Intestinal and hepatic biotransformation of pyrrolizidine alkaloid N-oxides to toxic pyrrolizidine alkaloids. Arch Toxicol 2019; 93: 2197–2209

[3] Widjaja F, Alhejji Y, Rietjens IMCM. The Role of Kinetics as Key Determinant in Toxicity of Pyrrolizidine Alkaloids and Their N-Oxides. Planta Med 2022; 88: 130–143

[4] Chen L, Peijnenburg A, de Haan L, Rietjens IMCM. Prediction of in vivo genotoxicity of lasiocarpine and riddelliine in rat liver using a combined in vitro-physiologically based kinetic modelling-facilitated reverse dosimetry approach. Arch Toxicol 2019; 93: 2385–2395

[5] Lehmann A, Geburek I, Hessel-Pras S, Hengstler JG, Albrecht W, Mielke H, Müller-Graf C, Yang X, Kloft C, Hethey C. PBTK modeling of the pyrrolizidine alkaloid retrorsine to predict liver toxicity in mouse and rat. Arch Toxicol 2023; 97: 1319–1333

[6] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci 2005; 94: 1259–1276

[7] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci 2006; 95: 1238–1257

[8] Schmitt W. General approach for the calculation of tissue to plasma partition coefficients. Toxicol Vitro Int J Publ Assoc BIBRA 2008; 22: 457–467

[9] Schmitt W. Corrigendum to: „General approach for the calculation of tissue to plasma partition coefficients“ [Toxicology in Vitro 22 (2008) 457-467] (DOI:10.1016/j.tiv.2007.09.010). Toxicol Vitro - TOXICOL VITRO 2008; 22: 1666–1666

[10] Kah M, Brown CD. LogD: Lipophilicity for ionisable compounds. Chemosphere 2008; 72: 1401–1408

[11] Aliagas I, Gobbi A, Lee M-L, Sellers BD. Comparison of logP and logD correction models trained with public and proprietary data sets. J Comput Aided Mol Des 2022; 36: 253–262

[12] Soares JX, Santos Á, Fernandes C, Pinto MMM. Liquid Chromatography on the Different Methods for the Determination of Lipophilicity: An Essential Analytical Tool in Medicinal Chemistry. Chemosensors 2022; 10: 340

[13] OECD. Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method. Paris, France: OECD Publishing; 1995

[14] Bharate SS, Kumar V, Vishwakarma RA. Determining Partition Coefficient (Log P), Distribution Coefficient (Log D) and Ionization Constant (pKa) in Early Drug Discovery. Comb Chem High Throughput Screen 2016; 19: 461–469

[15] OECD. Test No. 117: Partition Coefficient (n-octanol/water), HPLC Method. Paris, France: OECD Publishing; 2022

[16] Mannhold R, Poda GI, Ostermann C, Tetko IV. Calculation of Molecular Lipophilicity: State-of-the-Art and Comparison of LogP Methods on more than 96,000 Compounds. J Pharm Sci 2009; 98: 861–893

[17] Testa B, Crivori P, Reist M, Carrupt P-A. The influence of lipophilicity on the pharmacokinetic behavior of drugs: Concepts and examples. Perspect Drug Discov Des 2000; 19: 179–211

[18] Paixão P, Gouveia LF, Morais JAG. Prediction of drug distribution within blood. Eur J Pharm Sci 2009; 36: 544–554

[19] Yu S, Li S, Yang H, Lee F, Wu J-T, Qian MG. A novel liquid chromatography/tandem mass spectrometry based depletion method for measuring red blood cell partitioning of pharmaceutical compounds in drug discovery. Rapid Commun Mass Spectrom 2005; 19: 250–254

[20] Liu X-R, Wu K-C, Huang Y, Sun J-B, Ke X-Y, Wang J-C, Lu W-L, Zhang X, Zhang Q. In vitro and in vivo studies on plasma-to-blood ratio of paclitaxel in human, rabbit and rat blood fractions. Biol Pharm Bull 2008; 31: 1215–1220

[21] Uchimura T, Kato M, Saito T, Kinoshita H. Prediction of human blood-to-plasma drug concentration ratio. Biopharm Drug Dispos 2010; 31: 286–297

[22] Mamada H, Iwamoto K, Nomura Y, Uesawa Y. Predicting blood-to-plasma concentration ratios of drugs from chemical structures and volumes of distribution in humans. Mol Divers 2021; 25: 1261–1270

[23] Geburek I, Preiss-Weigert A, Lahrssen-Wiederholt M, Schrenk D, These A. In vitro metabolism of pyrrolizidine alkaloids – Metabolic degradation and GSH conjugate formation of different structure types. Food Chem Toxicol 2020; 135: 110868

[24] Caron G, Carrupt P-A, Testa B, Ermondi G, Gasco A. Insight into the Lipophilicity of the Aromatic N-Oxide Moiety. Pharm Res 1996; 13: 1186–1190

[25] Snoeck E, Jacqmin P, Van Peer A, Danhof M, Ver Donck K, Van Belle H, Woestenborghs R, Crabbé R, Van Gool R, Dupont A, Heykants J. The implications of non-linear red blood cell partitioning for the pharmacokinetics and pharmacodynamics of the nucleoside transport inhibitor draflazine. Br J Clin Pharmacol 1996; 42: 605–613

[26] Dash RP, Veeravalli V, Thomas JA, Rosenfeld C, Mehta N, Srinivas N. Whole blood or plasma: what is the ideal matrix for pharmacokinetic-driven drug candidate selection? Future Med Chem 2021; 13: 157–171

[27] Colson RO. Unexpected Results of Some Simple Exercises in Equilibrium Melting Based on Experimentally Determined Partition Coefficients. Int Geol Rev 1998; 40: 936–943

[28] US EPA. Estimation Programs Interface Suite™ for Microsoft® Windows. 2019;

[29] Merz K-H, Schrenk D. Interim relative potency factors for the toxicological risk assessment of pyrrolizidine alkaloids in food and herbal medicines. Toxicol Lett 2016; 263: 44–57

[30] ARChem LLC. SPARC Automated Reasoning in Chemistry. 2019; Im Internet: http://www.archemcalc.com/sparc.html

[31] SimulationsPlus. ADMET Predictor 12. ADMET Property Estimation and Model Building. 2024;

[32] Watanabe R, Esaki T, Kawashima H, Natsume-Kitatani Y, Nagao C, Ohashi R, Mizuguchi K. Predicting Fraction Unbound in Human Plasma from Chemical Structure: Improved Accuracy in the Low Value Ranges. Mol Pharm 2018; 15: 5302–5311

[33] EU Reference Laboratories. Guidance document on analytical quality control and method validation procedures for pesticide residues and analysis in food and feed. SANTE/11813/2017. 2017; Im Internet: https://www.eurl-pesticides.eu/userfiles/file/EurlALL/SANTE_11813_2017-fin.pdf; Stand: 26.10.2023

**Figure Legends**

**Fig. 1.** Experimentally determined and predicted log*P* values of PAs and PANOs. Experimental data are shown as mean ± standard deviation (SD) of equilibration from the organic to the aqueous phase ($n$ = 3, yellow circles) and from the aqueous to the organic phase ($n$ = 3, blue circles). Log*P* predictions were not available for intermedine *N*-oxide, monocrotaline *N*-oxide (CLOGP) and for PANOs (SPARC).

**Fig. 2.** Experimentally determined and predicted $R_b$ values of PAs and PANOs. Experimental data are shown as mean ± standard deviation (SD) with $n$ = 6.

**Fig. 3.** Effect of the concentration (4, 20, 100, 500 and 1000 ng/mL) on the $R_b$ value on the example of the PA retrorsine **(a)** and its *N*-oxide **(b)**. Experimental data are shown as mean ± standard deviation (SD) with $n$ = 6 ($n$ = 3: retrorsine 500 ng/mL).

**Table 1.** Experimentally determined and predicted log*P* values of PAs and PANOs.

| PAs and PANOs | **log*P*** | | | | | |
|---|---|---|---|---|---|---|
| | **Experimental**[a] | | ***In silico* prediction** | | | |
| | | | substructure-based | | property-based | |
| | | | KowWIN [28] | CLOGP [29] | SPARC [30] | S+logP [31] |
| Intermedine | -1.93 | ± 0.168 | 0.91 | -0.96 | -0.85 | -0.045 |
| Intermedine *N*-oxide | -1.36 | ± 0.0346 | 0.15 | -[b] | -[b] | -0.875 |
| Lasiocarpine | -0.487 | ± 0.0220 | 2.43 | 0.65 | 1.21 | 1.59 |
| Lasiocarpine *N*-oxide | -0.302 | ± 0.0269 | 1.66 | 0.77 | -[b] | 0.319 |
| Monocrotaline | -1.92 | ± 0.0727 | -1.18 | -0.93 | -1.01 | -0.686 |
| Monocrotaline *N*-oxide | -1.61 | ± 0.0352 | -1.95 | -[b] | -[b] | -1.56 |
| Retrorsine | -1.26 | ± 0.0610 | -0.25 | -0.62 | -0.24 | -0.298 |
| Retrorsine *N*-oxide | -1.24 | ± 0.0421 | -1.01 | -0.50 | -[b] | -1.26 |

[a]Mean ± standard deviation of equilibration from the organic to the aqueous phase ($n$ = 3) and from the aqueous to the organic phase ($n$ = 3).
[b]Log*P* predictions not available.

**Table 2.** Experimentally determined and predicted $R_b$ values of PAs and PANOs.

| PAs and PANOs | $R_b$ | | | |
|---|---|---|---|---|
| | **Experimental**[a] | | ***In silico* prediction** | |
| | | | Acid/base classification model[b] (Eq. 6) | Mechanistic model (Eq. 7) |
| Intermedine | 1.12 | ± 0.0853 | 1.00 | 0.807 |
| Intermedine *N*-oxide | 0.689 | ± 0.0499 | 1.00 | 0.867 |
| Lasiocarpine | 0.736 | ± 0.0438 | 1.00 | 0.601 |
| Lasiocarpine *N*-oxide | 0.897 | ± 0.0718 | 1.00 | 0.605 |
| Monocrotaline | 1.07 | ± 0.0865 | 1.00 | 0.672 |
| Monocrotaline *N*-oxide | 0.709 | ± 0.0430 | 1.00 | 0.674 |
| Retrorsine | 1.08 | ± 0.0859 | 1.00 | 0.717 |
| Retrorsine *N*-oxide | 0.736 | ± 0.228 | 1.00 | 0.830 |

[a]Mean ± standard deviation with $n$ = 6.
[b]At physiological pH of 7.4 PAs are basic, while PANOs are neutral.

# Supporting Information

**Characterization of lipophilicity and blood partitioning of pyrrolizidine alkaloids and their *N*-oxides *in vitro* and *in silico* for toxicokinetic modeling**

**Anja Lehmann[1,2,**], Manuel Haas[1,**], Julian Taenzer[1], Gerd Hamscher[3], Charlotte Kloft[2], Anja These[1,**], Christoph Hethey[1,**]**

**Affiliation**

[1] German Federal Institute for Risk Assessment (BfR), Berlin, Germany

[2] Department of Clinical Pharmacy and Biochemistry, Institute of Pharmacy, Freie Universitaet Berlin, Berlin, Germany

[3] Institute of Food Chemistry and Food Biotechnology, Justus Liebig University Giessen, Giessen, Germany

**Correspondence**

***Anja Lehmann***

Department of Clinical Pharmacy and Biochemistry

Institute of Pharmacy

Freie Universitaet Berlin

Kelchstr. 31

12169 Berlin

Germany

anja.lehmann@fu-berlin.de

[**] Authors contributed equally.

## Contents

**Fig. 1S.** Linear regression of log*P* values **(a)** and $R_b$ values **(b)** of reference substances experimentally determined in this work versus literature.

**Fig. 2S.** Effect of the concentration (4, 20, 100, 500 and 1000 ng/mL) on the $R_b$ value of reference substances.

**Fig. 3S.** Experimentally determined and predicted log*P* values of reference substances.

**Fig. 4S.** ddMS2 spectrum of caffeine.

**Fig. 5S.** ddMS2 spectrum of imipramine.

**Fig. 6S.** ddMS2 spectrum of sulfanilamide.

**Fig. 7S.** ddMS2 spectrum of theophylline.

**Fig. 8S.** ddMS2 spectrum of tolbutamide.

**Fig. 9S.** ddMS2 spectrum of quinidine.

**Fig. 10S.** ddMS2 spectrum of warfarin.

**Table 1S.** Predicted highest basic (strongest base) and lowest acidic (strongest acid) p*K*a values, as well as calculated charged species at experimental pH (log*P* determination).

**Table 2S.** Log*P* values and $R_b$ values of reference substances experimentally determined in this work versus literature.

**Table 3S.** MS/MS transitions and parameters.

**Table 4S.** Mean peak areas (LC-MS/MS) and percentage recovery of log*P* determination.

**Fig. 1S.** Linear regression of log*P* values **(a)** and $R_b$ values **(b)** of reference substances experimentally determined in this work versus literature. The red line represents the regression line. The inner dashed line represents the line of unity, while the outer lines in (a) indicate an error range of ± 1 log unit.

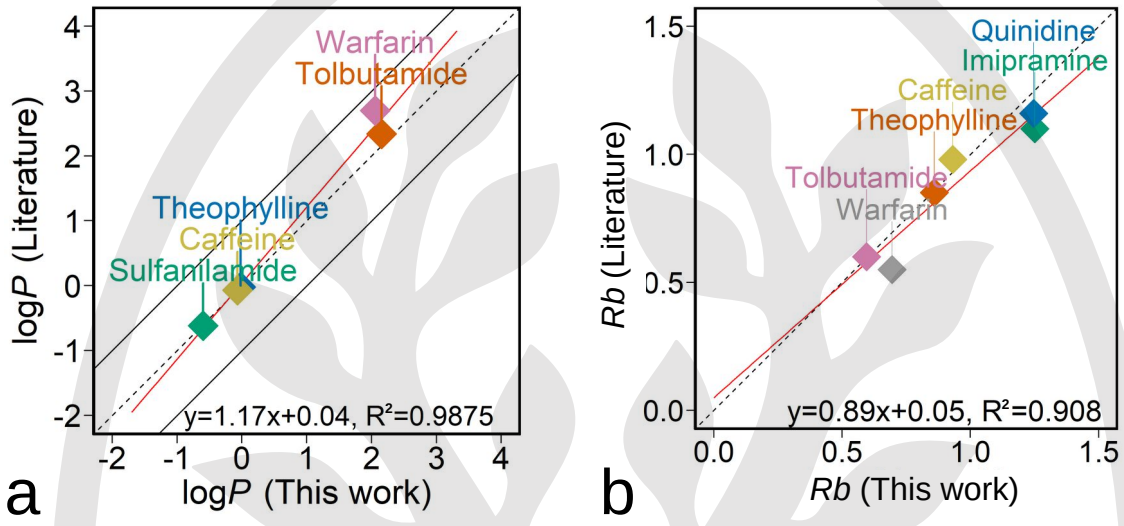


**Fig. 2S.** Effect of the concentration (4, 20, 100, 500 and 1000 ng/mL) on the $R_b$ value of reference substances. Experimental data are shown as mean ± standard deviation (SD) with $n$ = 3 (caffeine, imipramine) or $n$ = 2 (warfarin).

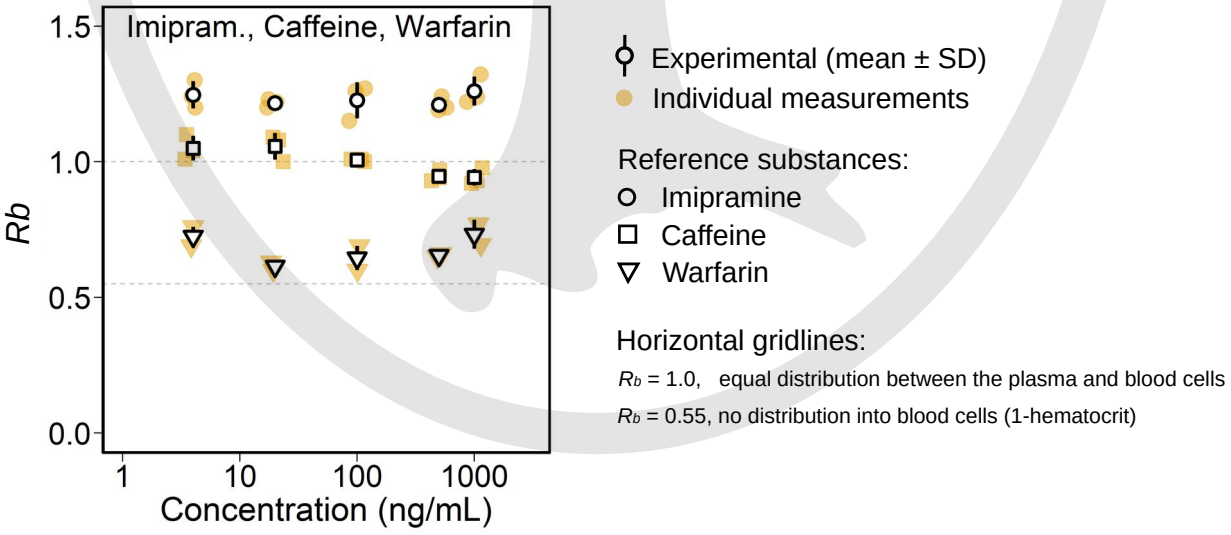

**Fig. 3S.** Experimentally determined and predicted log*P* values of reference substances. Experimental data are shown as mean ± standard deviation (SD) of equilibration from the organic to the aqueous phase (*n* = 3, yellow circles) and from the aqueous to the organic phase (*n* = 3, blue circles).

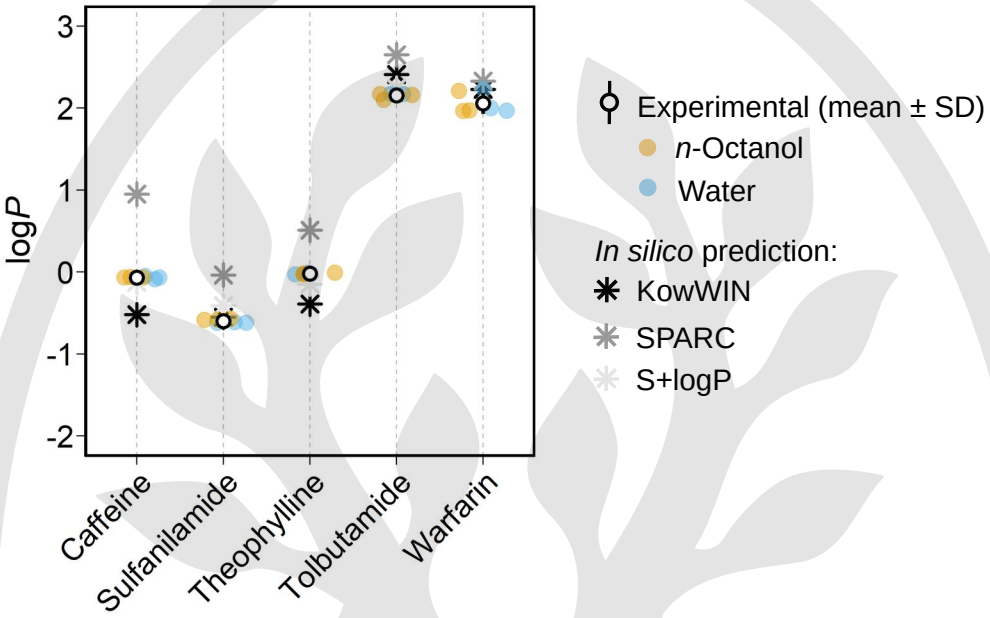


**Fig. 4S.** ddMS2 spectrum of caffeine.

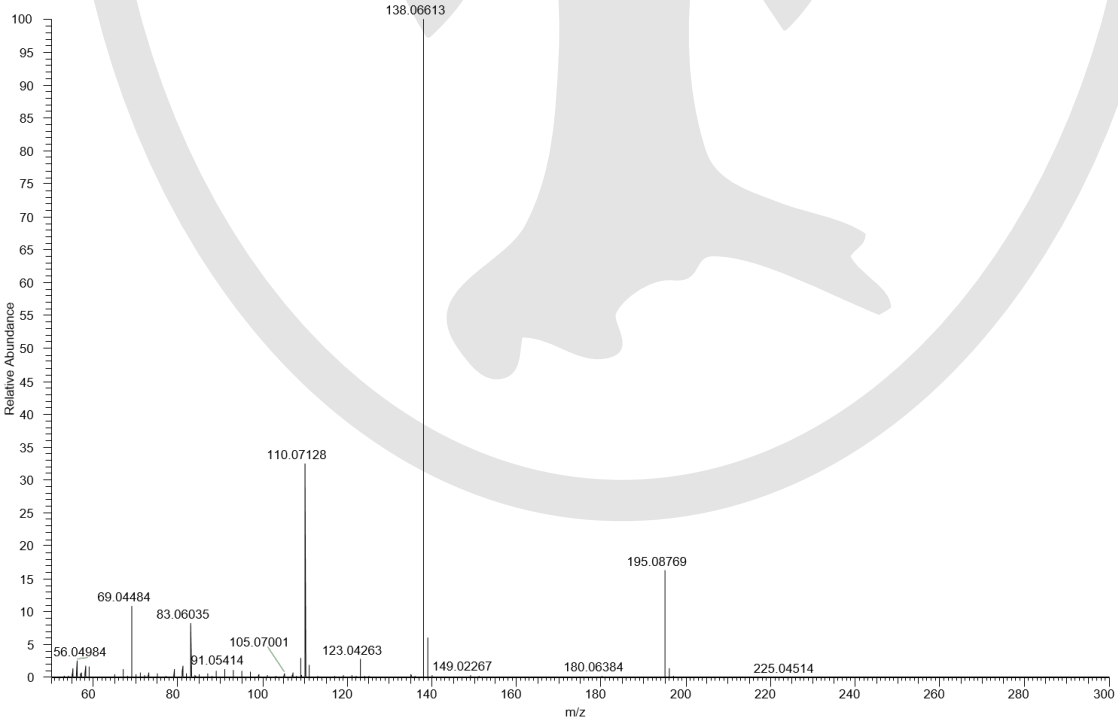

**Fig. 5S.** ddMS2 spectrum of imipramine.

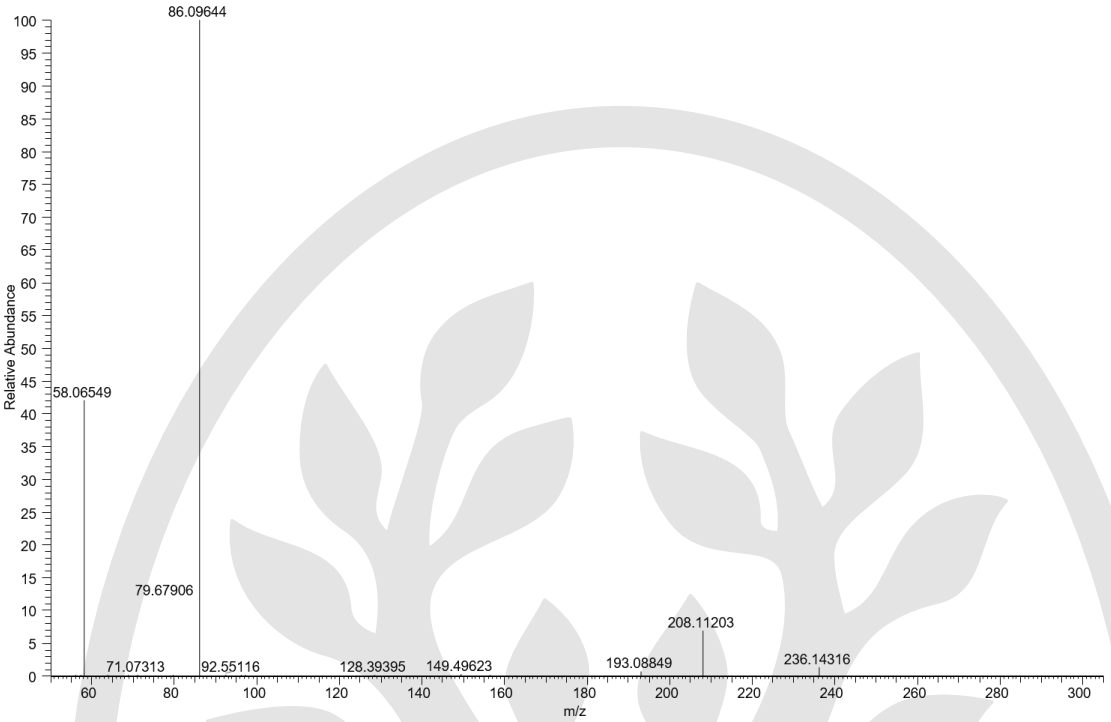


**Fig. 6S.** ddMS2 spectrum of sulfanilamide.

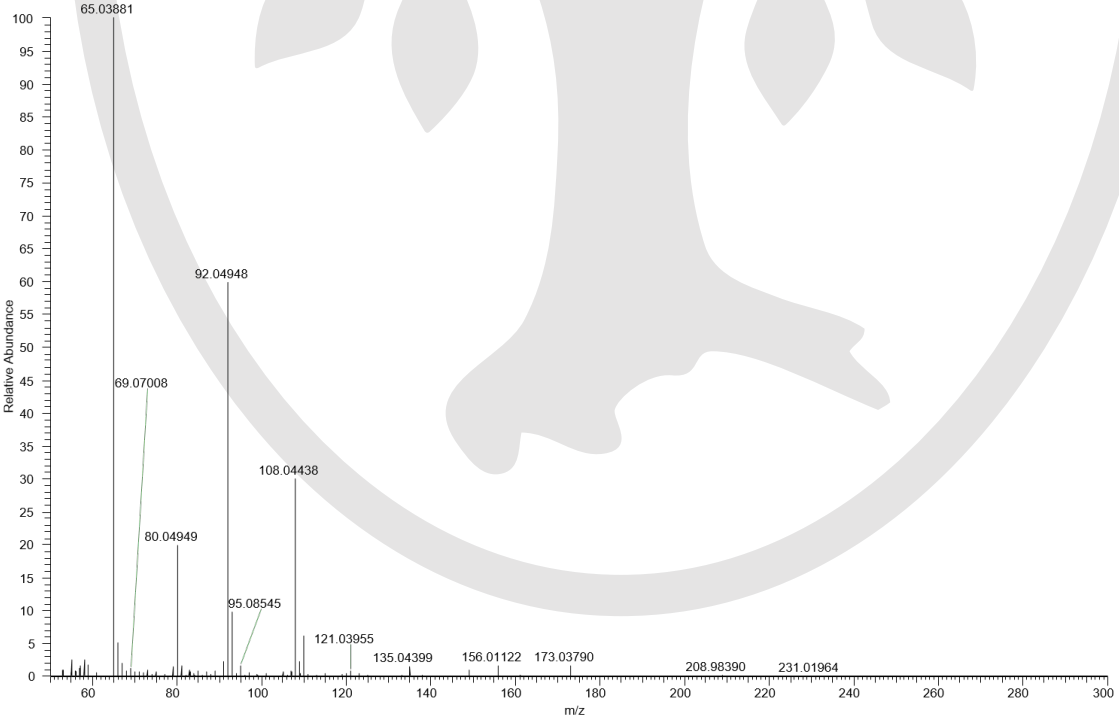

**Fig. 7S.** ddMS2 spectrum of theophylline.

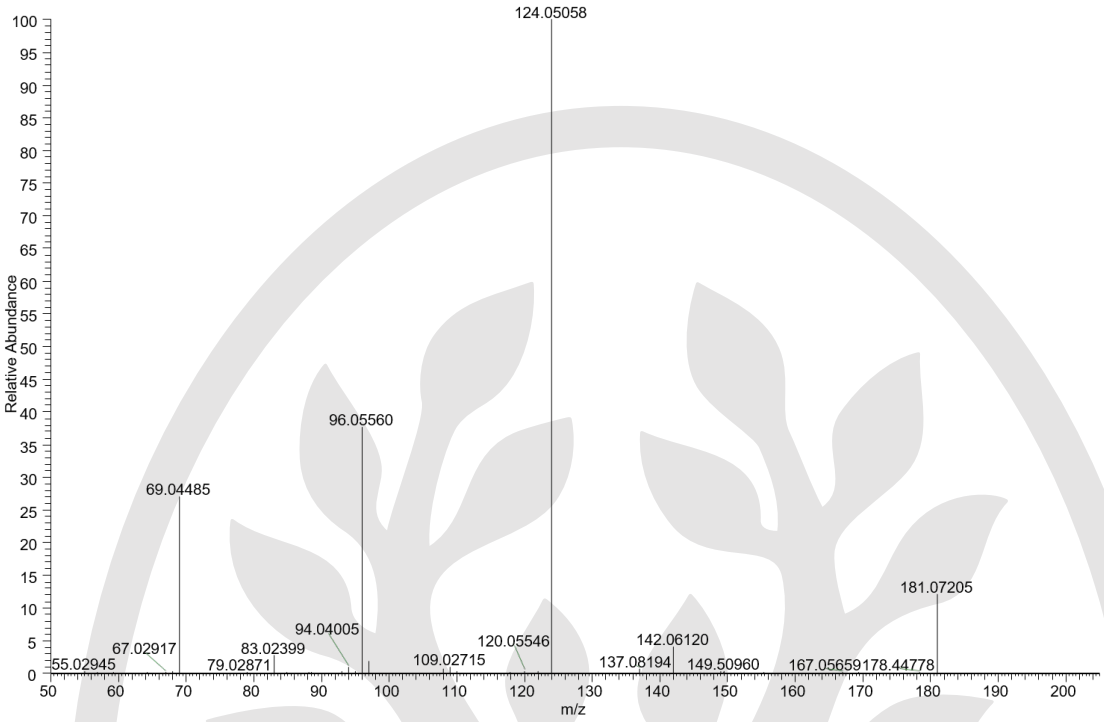


**Fig. 8S.** ddMS2 spectrum of tolbutamide.

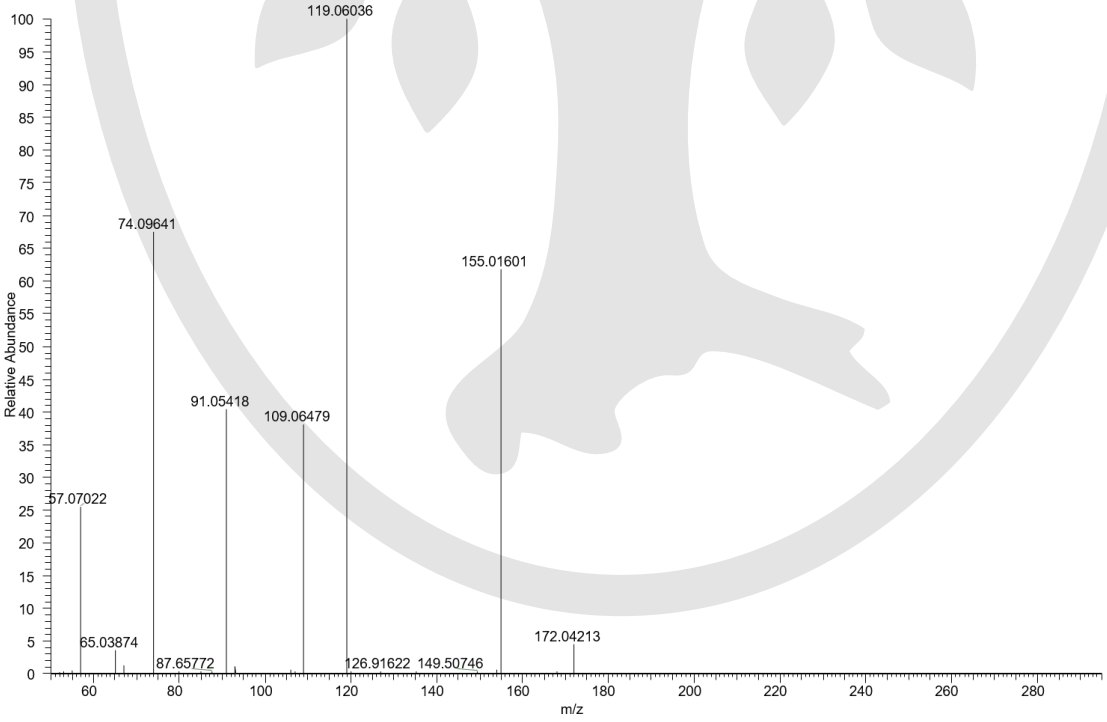

**Fig. 9S.** ddMS2 spectrum of quinidine.

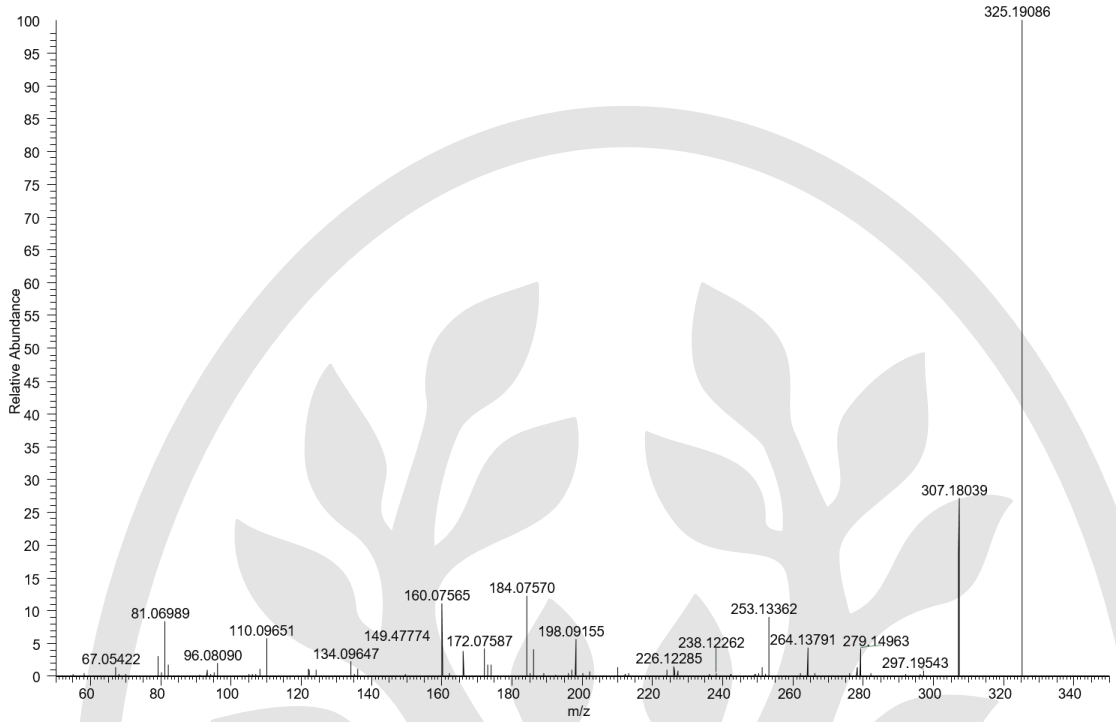


**Fig. 10S.** ddMS2 spectrum of warfarin.

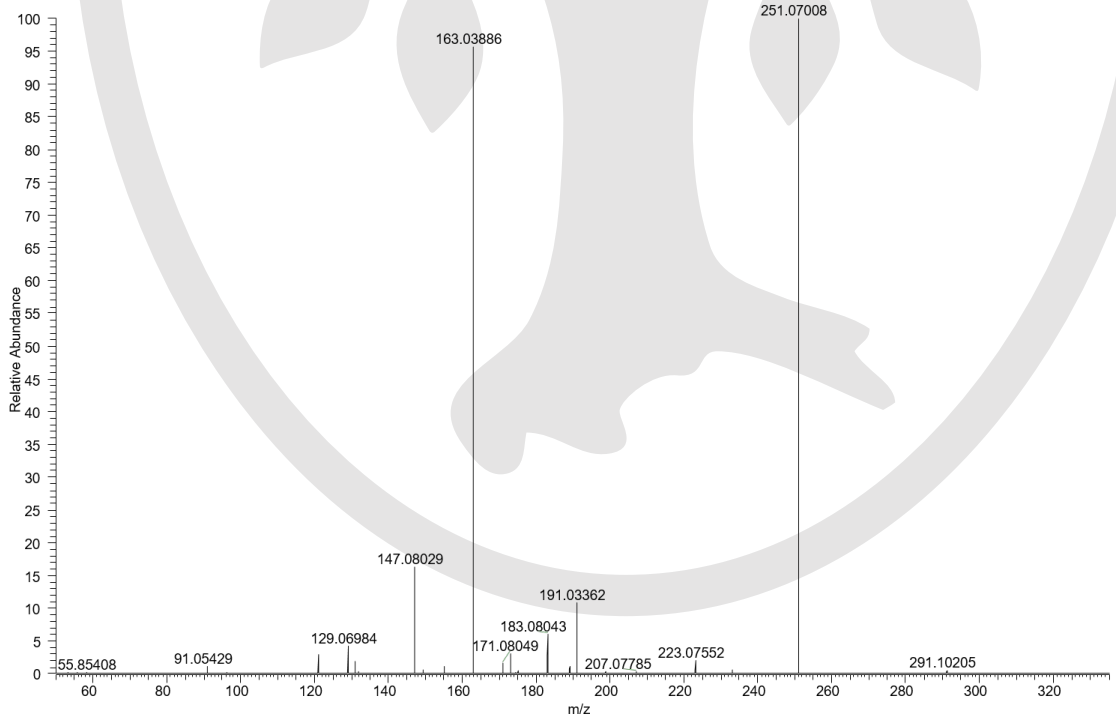

**Table 1S.** Predicted highest basic (strongest base) and lowest acidic (strongest acid) p*K*a values, as well as calculated charged species at experimental pH (log*P* determination).

| Compound | **Highest basic and lowest acidic p*K*a**[a] | **pH and predominant charge state at experimental conditions** | |
|---|---|---|---|
| | base; acid | pH[b] | charge state[c] |
| **Pyrrolizidine alkaloids and *N*-oxides** | | | |
| Intermedine | 7.90; - | 9 | neutral (97%) |
| Intermedine *N*-oxide | 4.67; - | 6 | neutral (97%) |
| Lasiocarpine | 6.44; - | 8 | neutral (96%) |
| Lasiocarpine *N*-oxide | 4.65; - | 6 | neutral (96%) |
| Monocrotaline | 6.82; - | 8 | neutral (97%) |
| Monocrotaline *N*-oxide | 4.60; - | 6 | neutral (97%) |
| Retrorsine | 6.94; - | 8 | neutral (97%) |
| Retrorsine *N*-oxide | 4.62; - | 6 | neutral (97%) |
| **Reference substances** | | | |
| Caffeine | 1.43; - | 3 | neutral (98%) |
| Sulfanilamide | 1.81; 10.6 | 8 | neutral (98%) |
| Theophylline | 1.32; 9.31 | 9 | neutral (98%) |
| Tolbutamide | -; 5.21 | 4 | neutral (98%) |
| Warfarin | -; 5.09 | 4 | neutral (97%) |

[a]p*K*a values predicted by ADMET Predictor [1].

[b]The experimental pH was calculated as pH = p*K*a + log($C_{\text{conjugate base}}/C_{\text{acid}}$) based on the nominal compound concentrations and autoprotolytic water activity with $C_{\text{conjugate base}}$ (µM) = 1(ng/mL)/Molar mass (g/mol) and $C_{\text{acid}}$ (µM) = 0.1 for bases and vice versa for acids. pH for amphoteric substances sulfanilamide and theophylline were calculated based on their strongest acid/base property.

[c]The fraction neutral $f_n$ was calulated as $f_n = 1/(1 + 10^{(\text{pH - p}Ka)})$ for acids and $f_n = 1 - (1/(1 + 10^{(\text{pH - p}Ka)}))$ for bases.

**Table 2S.** Log*P* values and $R_b$ values of reference substances experimentally determined in this work versus literature.

| Reference substance | **log*P*** | | | ***R*$_b$** | | |
|---|---|---|---|---|---|---|
| | **This work**[a] | | **Literature** | **This work**[b] | | **Literature** |
| Caffeine | -0.0683 | ± 0.0133 | -0.07 [2] | 0.931 | ± 0.0287 | 0.98 [3] |
| Imipramine | - | | - | 1.25 | ± 0.105 | 1.1 [4] |
| Quinidine | - | | - | 1.25 | ± 0.0971 | 0.92 [5], 1.4 [6] |
| Sulfanilamid | -0.598 | ± 0.0223 | -0.62 [7] | - | | - |
| Theophylline | -0.0197 | ± 0.00932 | -0.02 [7] | 0.860 | ± 0.00600 | 0.85 [8] |
| Tolbutamide | 2.16 | ± 0.0274 | 2.34 [9] | 0.595 | ± 0.0262 | 0.6 [10] |
| Warfarin | 2.06 | ± 0.128 | 2.7 [11] | 0.695 | ± 0.0436 | 0.55 [4] |

[a]Mean ± standard deviation of equilibration from the organic to the aqueous phase ($n$ = 3) and from the aqueous to the organic phase ($n$ = 3).

[b]Mean ± standard deviation with $n$ = 4 ($n$ = 3: caffeine; $n$ = 2: warfarin).

**Table 3S.** MS/MS transitions and parameters. QT: quantifier, QL: qualifier.

| Compound | **Precursor ion (*m*/*z*)** | **Product ions (*m*/*z*) QT; QL; QL** | **Collision energy (eV) QT; QL; QL** |
|---|---|---|---|
| **Reference substances** | | | |
| Caffeine | 195.1 | 138.1; 110.7; 108.7 | 15; 20; 10 |
| Imipramine | 281.2 | 86.1; 208.1; 193.1 | 15; 25; 35 |
| Sulfanilamide | 172.2 | 108.0; 156.0; 92.1 | 10; 10; 10 |
| Theophylline | 181.1 | 124.1; 142,1; 96.1 | 25; 25; 30 |
| Tolbutamide | 271.1 | 119.1; 155.0; 91.1 | 25; 20; 35 |
| Quinidine | 325.2 | 186.1; 307.2; 253.1 | 35; 35; 35 |
| Warfarin | 309.1 | 251.1; 163.0; 147.1 | 25; 15; 15 |
| **Pyrrolizidine alkaloids and their *N*-oxides** | | | |
| Intermedine | 300.2 | 120.1; 156.1; 138.1 | 26; 22; 30 |
| Intermedine *N*-oxide | 316.0 | 111.1; 172.1; 138.1 | 30; 30; 30 |
| Lasiocarpine | 412.2 | 120.1; 336.2; 220.1 | 18; 22; 30 |
| Lasiocarpine *N*-oxide | 428.2 | 136.1; 352.1; 254.1 | 38; 30; 26 |
| Monocrotaline | 326.2 | 120.1; 280.1; 194.1 | 40; 26; 30 |
| Monocrotaline *N*-oxide | 342.2 | 118.1; 138.1; 120.1 | 38; 30; 40 |
| Retrorsine | 352.2 | 120.2; 324.1; 138.2 | 34; 38; 30 |
| Retrorsine *N*-oxide | 368.2 | 120.1; 136.1; 118.1 | 38; 34; 40 |

**Table 4S.** Mean peak areas (LC-MS/MS) and percentage recovery of log*P* determination measured with sample size $n = 3$ at three different days by two distinct analysts (a, b) accounting for the

starting phase of equilibration (water, *n*-octanol). Relative dispersion reported as coefficient of variation (%CV). Intraday precision given as %CV of mean peak areas.

| Reference substance | **Analyst** | **Starting phase of equilibration** | **Day 1** | | **Day 2** | | **Day 3** | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean peak area (%CV) | %Recovery[a] (%CV) | Mean peak area (%CV) | %Recovery[a] (%CV) | Mean peak area (%CV) | %Recovery[a] (%CV) |
| Caffeine | a | water | $1.9{\cdot}10^5$ (1) | 121 (0) | $2.7{\cdot}10^7$ (2) | 114 (1) | $2.9{\cdot}10^7$ (2) | 116 (1) |
| | | *n*-octanol | $1.7{\cdot}10^5$ (1) | | $2.3{\cdot}10^7$ (0) | | $2.3{\cdot}10^7$ (1) | |
| | b | water | $1.9{\cdot}10^5$ (2) | 115 (1) | $2.7{\cdot}10^7$ (4) | 114 (2) | $2.8{\cdot}10^7$ (1) | 113 (1) |
| | | *n*-octanol | $1.7{\cdot}10^5$ (4) | | $2.3{\cdot}10^7$ (2) | | $2.4{\cdot}10^7$ (3) | |
| Sulfanilamide | a | water | $7.9{\cdot}10^7$ (1) | 117 (2) | $7.0{\cdot}10^7$ (2) | 104 (1) | $7.8{\cdot}10^7$ (2) | 115 (2) |
| | | *n*-octanol | $1.9{\cdot}10^7$ (3) | | $1.7{\cdot}10^7$ (1) | | $1.9{\cdot}10^7$ (2) | |
| | b | water | $7.4{\cdot}10^7$ (5) | 111 (4) | $6.5{\cdot}10^7$ (1) | 98 (2) | $7.2{\cdot}10^7$ (5) | 107 (5) |
| | | *n*-octanol | $1.9{\cdot}10^7$ (2) | | $1.7{\cdot}10^7$ (2) | | $1.9{\cdot}10^7$ (2) | |
| Theophylline | a | water | $4.9{\cdot}10^4$ (0) | 92 (1) | $8.2{\cdot}10^6$ (1) | 130 (1) | $8.5{\cdot}10^5$ (3) | 103 (0) |
| | | *n*-octanol | $4.5{\cdot}10^4$ (1) | | $7.9{\cdot}10^6$ (2) | | $8.5{\cdot}10^5$ (1) | |
| | b | water | $4.7{\cdot}10^4$ (2) | 95 (2) | $8.0{\cdot}10^6$ (2) | 120 (2) | $8.2{\cdot}10^5$ (2) | 104 (2) |
| | | *n*-octanol | $4.5{\cdot}10^4$ (2) | | $7.9{\cdot}10^6$ (2) | | $7.6{\cdot}10^5$ (1) | |
| Tolbutamide | a | water | $1.0{\cdot}10^5$ (2) | 118 (2) | $3.8{\cdot}10^4$ (19) | 108 (1) | $3.8{\cdot}10^4$ (12) | 108 (1) |
| | | *n*-octanol | $1.5{\cdot}10^7$ (2) | | $5.6{\cdot}10^6$ (1) | | $5.5{\cdot}10^6$ (1) | |
| | b | water | $9.9{\cdot}10^4$ (3) | 114 (1) | $4.4{\cdot}10^4$ (13) | 102 (1) | $3.6{\cdot}10^4$ (6) | 95 (2) |
| | | *n*-octanol | $1.5{\cdot}10^7$ (0) | | $5.5{\cdot}10^6$ (1) | | $5.2{\cdot}10^6$ (1) | |
| Warfarin | a | water | $2.2{\cdot}10^5$ (18) | 85 (3) | $2.9{\cdot}10^7$ (10) | 99 (2) | $3.5{\cdot}10^7$ (14) | 106 (3) |
| | | *n*-octanol | $2.1{\cdot}10^7$ (10) | | $4.9{\cdot}10^9$ (2) | | $3.2{\cdot}10^9$ (9) | |
| | b | water | $2.5{\cdot}10^5$ (3) | 99 (1) | $3.2{\cdot}10^7$ (32) | 99 (0) | $3.4{\cdot}10^7$ (3) | 97 (2) |
| | | *n*-octanol | $2.2{\cdot}10^7$ (1) | | $5.0{\cdot}10^9$ (0) | | $3.1{\cdot}10^9$ (2) | |

[a]Recovery compares the total quantity of analyte present in both phases with the quantity of analyte originally introduced: %Recovery = 100% · (Peak area (Analyte in *n*-octanol) + Peak area (Analyte in water))/ Peak area (Analyte recovery sample).

## References

[1] SimulationsPlus. ADMET Predictor 12. ADMET Property Estimation and Model Building. 2024;

[2] Mirrlees MS, Moulton SJ, Murphy CT, Taylor PJ. Direct measurement of octanol-water partition coefficients by high-pressure liquid chromatography. J Med Chem 1976; 19: 615–619

[3] Gaohua L, Abduljalil K, Jamei M, Johnson TN, Rostami-Hodjegan A. A pregnancy physiologically based pharmacokinetic (p-PBPK) model for disposition of drugs metabolized by CYP1A2, CYP2D6 and CYP3A4. Br J Clin Pharmacol 2012; 74: 873–885

[4] Obach RS. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: An examination of in vitro half-life approach and nonspecific binding to microsomes. Drug Metab Dispos Biol Fate Chem 1999; 27: 1350–1359

[5] Sawada Y, Hanano M, Sugiyama Y, Iga T. Prediction of the disposition of nine weakly acidic and six weakly basic drugs in humans from pharmacokinetic parameters in rats. J Pharmacokinet Biopharm 1985; 13: 477–492

[6] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci 2005; 94: 1259–1276

[7] Pyka A, Babuśka M, Zachariasz M. A comparison of theoretical methods of calculation of partition coefficients for selected drugs. Acta Pol Pharm 2006; 63: 159–167

[8] Ebden P, Banks J, Peel T, Buss DC, Routledge PA, Spragg BP. The disposition of theophylline in blood in chronic obstructive lung disease. Ther Drug Monit 1986; 8: 424–426

[9] Singh BN. A quantitative approach to probe the dependence and correlation of food-effect with aqueous solubility, dose/solubility ratio, and partition coefficient (Log P) for orally active drugs administered as immediate-release formulations. Drug Dev Res 2005; 65: 55–75

[10] Perkins EJ, Posada M, Kellie Turner P, Chappell J, Ng WT, Twelves C. Physiologically Based Pharmacokinetic Modelling of Cytochrome P450 2C9-Related Tolbutamide Drug Interactions with Sulfaphenazole and Tasisulam. Eur J Drug Metab Pharmacokinet 2018; 43: 355–367

[11] Hansch C, Leo A, Hoekman D. Exploring QSAR - Hydrophobic, Electronic, and Steric Constants. p.161. Washington, DC, USA: American Chemical Society; 1995