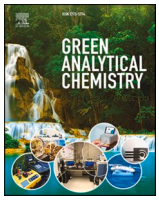
Green Analytical Chemistry 16 (2026) 100328

Contents lists available at ScienceDirect

Green Analytical Chemistry

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



QSPR-driven prediction of analyte permeability for advancing CP-MIMS applications

Enmanuel Cruz Muñoz, Veronica Termopoli *, Davide Ballabio, Marco Orlandi, Viviana Consonni

Milano Chemometrics and QSAR Research Group, Department of Earth and Environmental Sciences, University of Milano-Bicocca, P.za della Scienza 1, 20126 Milano, Italy

ARTICLE INFO

Keywords:
Condensed Phase Membrane Introduction Mass Spectrometry
QSPR
Classification modeling
Molecular similarity
PDMS permeability

ABSTRACT

Condensed Phase-Membrane Introduction Mass Spectrometry (CP-MIMS) is a sustainable and highly versatile approach within the framework of Direct Mass Spectrometry techniques, which enables real-time determination of target analytes by integrating sampling, purification, and mass-spectrometric analysis into a single step. Liquid or slurry samples are directly coupled to a mass spectrometer through a semipermeable membrane, with transport of analytes governed by their physicochemical and structural properties. At present, experimental testing is often required to determine whether a compound is suitable for CP-MIMS analysis, limiting the scalability of this approach. To overcome this limitation, computational strategies, such as Quantitative Structure-Property Relationship (QSPR), offer the possibility to rationalize and predict analyte behavior at the membrane interface, enabling a rapid recognition of permeant compounds.

We developed a QSPR classifier to predict whether a compound can permeate a polydimethylsiloxane membrane in the CP-MIMS system by using machine learning methods to relate molecular structural features to physicochemical properties. To improve robustness, the literature derived training set was expanded through experimental measurements and similarity searches in large molecular databases. The resulting model achieved high sensitivity and specificity in large-scale external validation. Experimental testing of QSPR-predicted permeant candidates confirmed the reliability of the model predictions. Moreover, a further experimental evaluation was performed on environmentally relevant contaminants under CP-MIMS conditions.

The proposed QSPR framework provides a reliable and automated *in-silico* tool for pre-screening of permeant analytes, supporting more efficient method development and improved sustainability for advancing CP-MIMS applications in several fields.

1. Introduction

In an era of increasing automation, there is a growing demand for analytical methodologies capable of detecting target analytes with minimal sample preparation and no need for extensive laboratory procedures, especially in complex matrices. Direct Mass Spectrometry (DMS) techniques have experienced rapid development due to their versatility across a wide range of sample types, streamlined workflows, and suitability for high throughput applications [1–3]. DMS techniques enable real-time analysis while fulfilling many principles of green chemistry since they eliminate extensive sample preparation procedures and avoid the use of large volume solvents and consumables.

Among DMS strategies, Condensed Phase-Membrane Introduction Mass Spectrometry (CP-MIMS) is a well-established technique that enables real-time determination and quantification of target analytes by integrating sampling, purification, and mass-spectrometric analysis into a single step [4]. It allows connecting liquid or slurry samples directly to a mass spectrometer (MS) through a semipermeable membrane Analytes that permeate the membrane are transported to the MS ion source, for subsequent detection, via a continuous flow of an organic solvent serving as the acceptor phase (AP). AP is commonly driven by a liquid chromatography (LC) pump, and it flows continuously through the lumen of a capillary hollow-fiber membrane (HFM), which is immersed in the sample matrix (donor phase). Generally, a polydimethylsiloxane membrane (PDMS) is employed, allowing permeation of neutral, hydrophobic analytes.

* Corresponding author.
E-mail address: veronica.termopoli@unimib.it (V. Termopoli).

https://doi.org/10.1016/j.greeac.2026.100328
Received 15 December 2025; Received in revised form 2 February 2026; Accepted 2 February 2026
Available online 3 February 2026

The main strength of CP-MIMS lies in the ability of the PDMS membrane to act as a barrier against matrix interferences, including ions, salts, and particulate matter [5]. This characteristic enables minimal or no sample preparation steps allowing fast and on-line determination. Moreover, CP-MIMS allows the possibility to adjust and monitor on the fly the donor phase pH. Vandergrift and co-workers [6] proposed a modified CP-MIMS approach that employs polymer inclusion membrane (PIM) systems: introducing a small percentage of a linear alkane as cosolvent (*i.e.*, heptane) in a methanolic AP together with a PDMS membrane creates an in-situ PIM. This configuration decreases response times and enhances sensitivity, as the diffusivity and membrane partitioning of the analytes are improved. It offers efficient screening across a wide range of compound classes in challenging matrices such as soils, oil sands process water, wastewater, seawater, groundwater, and effluents from hydrocarbon extraction while simultaneously enabling direct and cost-effective quantification [7–10]. The implementation of real-time monitoring for chemical reactions [11], microplastic-contaminant sorption [12], and the contaminant migration from food contact materials [13] has considerably shortened the overall laboratory duty cycle, enabling precise quantification of target analytes within a single-step analytical process. Moreover, the technique allows a fine tuning of several parameters, including real-time adjustment of the donor phase pH and the selection of membrane-compatible acceptor solvents, thereby preserving membrane integrity and facilitating the attainment of low detection limits [5].

Several publications provide comprehensive discussions on the fundamental concepts and operational aspects of CP-MIMS technique [14–16]. The process by which analytes migrate through the membrane is mainly controlled by Fick's diffusion law and is proportional to the product of the partitioning coefficient (K) of a permeant into the membrane and its diffusivity (D) through the membrane media. A central mechanistic feature of CP-MIMS is that the analytical response time ($t_{10\%\text{-}90\%}$) reflects the kinetics of analyte diffusion through the membrane. For PDMS membranes, which are the most common ones, more hydrophobic compounds that have a small hydrodynamic volume tend to permeate more efficiently [17,18]. Consequently, membrane transport represents a built-in selectivity mechanism, conferring partial analyte discrimination prior to ionization. However, this same dependence introduces variability: not all compounds are equally suited to CP-MIMS, and predicting permeability a priori remains challenging. Hence, the efficiency of CP-MIMS depends critically on the permeability of analytes across the membrane interface, which is governed by their physicochemical and structural properties.

At present, experimental testing is often required to determine whether a given compound or class of compounds is suitable to CP-MIMS analysis, limiting the scalability of the approach. To overcome this limitation, Quantitative Structure-Property Relationship (QSPR) modeling can provide a powerful computational strategy to rationalize and predict analyte behavior at the membrane interface. QSPR has extensive applications in the field of analytical chemistry, in particular to predict the analyte behavior in different chromatographic systems [19]. It relies on the principle that analytical properties are influenced by molecular structure features, which can be encoded into chemical-structural descriptors and modelled by machine learning methods [20]. If QSPR modeling is sufficiently accurate, it can be used to provide reliable predictions for untested chemicals. The rationale for this approach stems from established knowledge that membrane permeability is governed by several physicochemical and structural parameters, such as molecular weight, hydrodynamic volume, lipophilicity, functional groups capable of hydrogen bonding, and polarity. Recent studies by Duncan and Maguire *et al.* [18,21] combined CP-MIMS and QSPR to predict octanol-water partition coefficients of series of compounds, such as naphthenic acids or 6PPD-quinone and its structural related analogues.

In this study, we use QSPR to directly predict the capability of analytes to permeate across the membrane interface from the knowledge of their chemical-structure features. Establishing of such an approach is critical for extending the operational scope of DMS beyond its current boundaries. We addressed this challenge by designing an automated QSPR-based screening pipeline aimed at the rapid recognition of compounds potentially suitable for analysis with the CP-MIMS system.

2. Materials and methods

2.1. Reagents and chemicals

A detailed description of the chemicals, and the concentrations of the stock solutions is provided in Table S1 of the Supplementary Material. Acetonitrile (ACN), heptane, acetone, and methanol (MeOH) were LC-MS grade (purity >99%) and purchased from Merck (Milan, Italy). Ultrapure water was dispensed by a Milli-Q Integral 5 purification system (Merck, Italy). All stock solutions for each analyte were prepared gravimetrically. Stock solutions of theobromine, tadalafil, and trifluralin were prepared in ACN; diclofenac sodium salt in Milli-Q water; carbamazepine in a 60:40 ACN:H_2O (v:v); dibenzothiophene and 3-bromofluoranthene in acetone, and the remaining compounds were dissolved in MeOH.

2.2. CP-MIMS analysis

CP-MIMS measurements were performed using an LCMS-2010 system (Shimadzu, Japan), equipped with a binary pump, an SPD-M40 photodiode array (PDA) detector, an RF-20A fluorescence (FL) detector, and a single quadrupole mass spectrometer with an electrospray ionization (ESI) interface. The acceptor phase (AP),at a composition of 15%:85% (v:v) heptane in methanol, was delivered at 50 μL/min. A short C18 LC column (5 cm × 2.1 mm × 5 μm, Restek, Germany) was positioned upstream of the CP-MIMS probe to ensure stable backpressure. The probe was constructed from 2 cm length (exposed to the donor phase) of PDMS hollow fiber membrane (HFM, 170 μm thickness, 0.64 mm o.d., 0.30 mm i.d.; VWR, Italy) mounted on a 'U-Probe' interface, as described elsewhere. The PDA and FL detectors were connected in-line upstream of the mass spectrometer to enable simultaneous spectrophotometric and mass spectrometric monitoring of the permeant compounds. This set-up allows detection of all analytes capable of permeating, including those with poor ionization efficiency in the ESI interface. The PDA acquisition was set at 210–720 nm; fluorescence detection used excitation and emission ranges of 280–330 nm and 350–450 nm, respectively. MS parameters were optimized in flow injection analysis (FIA), injecting 2 μL of working standard solutions at 2 ng/μL in full SCAN mode, at a flow rate of 200 μL/min. The ESI source was operated in positive or negative ion mode depending on the ionization behavior of each compound. Instrumental parameters including detector voltage (0.5–3 kV), nebulizing N_2 flow (0.5–1.5 L/min), Curved Desolvation Line (CDL) temperature, and interface voltage were systematically varied, while the Heat Block (HB) temperature was maintained at 200°C. Optimized conditions for each analyte in negative (ESI^-) and positive (ESI^+) modes are summarized in Tables S1 of the Supplementary Material. Data acquisition of CP-MIMS experiments was performed in SIM mode, monitoring the protonated or deprotonated molecular ion.

For the experimental workflow, CP-MIMS probe was immersed in a 20 mL vial containing 15 mL ultrapure water (donor phase), stirred at 400 rpm with a magnetic stir bar (VELP Scientifica, Monza-Brianza, Italy) to maintain homogeneous mixing. The signal background in ultrapure water was acquired until stabilization (ca. 1-3 min), then, a defined volume of analyte stock solution was spiked into the donor phase using a mechanical micropipette. Table S1 of the Supplementary Material summarizes the analytical conditions for the aqueous donor phase, with basic and acidic water was adjusted to pH 8.8 and 2.7, respectively (Mettler Toledo pH meter, Mettler, Italy), to ensure charge-neutral analytes. The fortification volumes used for each compound

were also specified. On each analysis day, system performance was verified by fortifying 15 mL of ultrapure water with 30 μL of a stock solution of chlorobenzene, used as model compound, at a concentration of 500 ng/μL in acetone. This procedure ensures the absence of instrumental issues or membrane-related anomalies. Due to the negligible ionization in ESI of chlorobenzene, its signal was acquired with the PDA detector, monitoring the 270 nm wavelength.

2.3. CP-MIMS signal processing

Since CP-MIMS does not involve chromatographic separation, its output is a chronogram. The onset of permeation was detected as the rising signal front, typically occurring between 0.1 and 6 min depending on the analyte. Once steady state was achieved, the signal was recorded for 2–4 min before the fortified donor phase vial was replaced with a vial containing fresh solvent (ACN, MeOH, or H_2O:MeOH 50:50 (v:v)) to wash out the analyte and re-establish the background level.

Parameter calculations were carried out as follows: the steady state signal ($\overline{Y}_{steady}$) was obtained by averaging 60 s of the maximum reached signal after donor phase fortification. The background signal ($\overline{Y}_0$) was calculated by averaging the signal in a time range of 60 s after the complete equilibration of the membrane within the aqueous donor phase. Each analyte was tested in triplicate to evaluate repeatability. The results obtained for each tested compound were reported in Table S2 of the Supplementary Material. The time dependent signal $Y(t)$ was normalized to the 0–100 scale using the following equation:

$$Y'(t) = \frac{Y(t) - \overline{Y}_0}{\overline{Y}_{steady} - \overline{Y}_0} \cdot 100 \quad (1)$$

The signal response time ($t_{10\%\text{-}90\%}$) was defined as the interval required for the normalized signal $Y'(t)$ to increase from 10% to 90% of its steady-state intensity once the membrane was exposed to donor phase containing the analyte.

To facilitate signal managing, a dedicated MATLAB application with a graphical user interface was previously created and it can be freely downloaded from the Milano Chemometrics and QSAR Research Group website [22]. Given that the aim of this study was not quantitative determination but rather to identify which compounds could permeate the membrane, we considered as permeant (positive) only the molecules whose average steady-state signal was higher than the average background signal plus three times its standard deviation, otherwise they were classified as non-permeant (negative).

2.4. In-silico modelling

2.4.1. Data setup and curation

The dataset was initially built by collecting permeability data from the literature on CP-MIMS applications [6–9,11,12,17,18,23–33], resulting in a total of 82 compounds with 81 permeant (positive) analytes and only one non-permeant (negative) analyte. To address the permeant/non-permeant imbalance, four additional chemicals (i.e., resveratrol, theobromine, ivermectin B1, and vitamin K) expected to be non-permeant were tested; their lack of permeation was experimentally confirmed and these compounds were added to the negative class.

To further enrich the negative class, a structural similarity search was performed in the COCONUT (COlleCtion of Open Natural prodUcTs) database [34] using the five non-permeant compounds as queries. Compounds with high structural similarity were retrieved and further filtered using physicochemical criteria related to permeability, yielding a large set of potential non-permeant chemicals. Following a similar workflow, an external validation set of likely permeant compounds was extracted from COCONUT. Literature-derived permeant chemicals were used as queries to identify structurally similar compounds in the same database, which were then filtered using property thresholds associated with high permeability.

2.4.2. Molecular descriptors

A Quantitative Structure-Property Relationship (QSPR) approach was used to establish the relationship between chemical-structural features and permeability of analytes across the membrane interface. The chemical structures of the analytes were encoded as SMILES notations [35], which are symbolic representations of chemicals suitable for cheminformatics use. From the SMILES, an initial set of 164 molecular descriptors (MDs) was calculated. These descriptors encoded heterogeneous chemical-structural characteristics related to molecular size, steric hindrance, lipophilicity (commonly expressed by the octanol–water partition coefficient), the type and occurrence frequency of functional groups, the ability of hydrogen bonding, and global polarity [36]. In particular, the following types of molecular descriptors were calculated: topological indices [37], P_VSA-like descriptors [38], functional group counts [39], CATS 2D descriptors [40] and constitutional descriptors [36].

2.4.3. One-class classifier

We used a one-class modelling method [41] to characterize the class space of permeant (positive) chemicals, which is the target class of interest in this study. An advantage of this classification approach is the intrinsic definition of the model applicability domain, that is, the region in the chemical space where predictions can be considered reliable [42].

The classification model was developed by exploiting Potential Functions (PFs), also known as Kernel Density Estimators [41]. The classification of a new compound into the positive class is determined by means of the cumulative potential for the new compound, obtained by summing up the individual potentials of all positive training analytes, which decrease with distance from the analyte itself. The decision rule is based on a potential threshold: the new compound to be predicted is classified as positive if its cumulative potential is higher than a threshold defined as the 95th percentile of the potential values of training compounds (0.0145 in this study) [43]. The shape of the potential field depends on the chosen type of kernel and a smoothing parameter. In this study, we used the gaussian kernel and a smoothing value equal to 0.4, which was selected by internal cross-validation based on 5 venetian blind folds.

In order to select the optimal subset of molecular descriptors that define the chemical space, we coupled PFs with Genetic Algorithms (GAs), which is a well-known approach for descriptor selection. Starting from an initial randomly generated population of binary vectors encoding the inclusion or exclusion of descriptors (chromosomes), GAs simulate an evolutionary process aimed at minimizing the classification error estimated via cross-validation [44]. New chromosomes are generated through genetic operators such as crossover and mutation until the final population of chromosomes provides the optimal subset of molecular descriptors. The GA selection was carried out with a population size equal to 50, 10,000 runs, mutation and crossover rates equal to 0.01 and 0.5, respectively.

2.5. Software

Molecular descriptors were calculated by Dragon 7 software [45]. Classification models based on Potential Functions and Genetic Algorithms were calculated by means of the Classification Toolbox for MATLAB and in-house MATLAB functions [46]. Data to reproduce the results of this study are provided in the Table S2 of the Supplementary Material.

3. Results

3.1. Permeability dataset

3.1.1. Data collection from literature

From previously published CP-MIMS applications reported in Section 2.4.1, a total of 81 permeant (positive) analytes were retrieved along

with only a single non-permeant (negative) analyte (syringic acid). The permeant analytes include molecules from diverse chemical classes, with molecular weights ranging from approximately 90 to 400 Da. The list of compounds identified from the literature, detailing their names, CAS identifiers, SMILES notations, literature references and classification labels (positive: permeant, negative: non-permeant) is provided in Table S2 of the Supplementary Material.

3.1.2. Data integration from experiments to support statistical modeling

Since the literature mainly reports chemicals that are able to permeate, but statistical modelling also requires non-permeant chemicals to better learn the relationships between molecular structure features and ability to permeate, we experimentally tested the permeability of four additional chemicals (i.e., resveratrol, theobromine, ivermectin B_1, and vitamin K), which were expected not to permeate according to expert knowledge.

The experimental results confirmed that the four selected compounds are indeed non-permeant; they were therefore added to the dataset as negative compounds, together with syringic acid (previously retrieved from the literature). The negligible permeation of resveratrol is likely due to the low membrane/water partition coefficient, resulting from hydrogen-bond donation by the three phenolic hydroxyl groups and reduced diffusivity within the PDMS membrane. Likewise, theobromine does not permeate through the PDMS membrane. Its polar xanthine core contains multiple hydrogen-bond donors and acceptors, leading to poor partitioning into the hydrophobic PDMS matrix. Consequently, both compounds exhibit negligible permeation in the CP-MIMS system. Despite their hydrophobicity, ivermectin B_1 and vitamin K do not permeate through PDMS membranes under CP-MIMS conditions. For ivermectin B_1, the main reason can be attributed to its very large molecular size (MW=875.1 Da) that severely limits diffusion within the PDMS polymer. The high lipophilicity favors membrane partitioning but cannot compensate for the extremely low diffusivity, resulting in negligible steady-state signals, even in the presence of polymer inclusion membrane formation [6]. Even though Vitamin K is neutral and highly hydrophobic, its isoprenoid chain and conjugated ring system yield a very large hydrodynamic volume, leading to poor diffusion in PDMS. Consequently, both compounds do not exhibit permeation in the PDMS membrane. The experimental results for the four tested compounds are provided in Table S2 of the Supplementary Material, along with their CAS identifiers and SMILES notations.

3.1.3. Data enrichment by similarity search in molecular databases

Although one-class modelling formally relies solely on the positive class (permeant) to establish the decision boundary, a balanced representation of both negative and positive classes is generally preferable, as it allows a more reliable estimation of the model specificity, that is, the ability to avoid misclassifying negative chemicals as positive.

To enrich the negative class, we performed a structural similarity search in the COCONUT (COlleCtion of Open Natural prodUcTs) database [34], using the five available negative compounds as query molecules. The database contains over 400,000 unique chemicals from 52 different sources. Data were initially curated with solvent exclusion, salt removal and charge neutralization [47]. Moreover, compounds with SMILES that could not be parsed with RDKIT were removed [48]. The resulting dataset included 129,869 unique chemicals [47].

We used the SMILES representation of all the 129,869 COCONUT compounds, as well as for the 5 negative chemicals, to calculate binary extended connectivity fingerprints (ECFPs, 1024 bits), which provide a holistic and comprehensive representation of molecular structures [49]. Subsequently, we searched for COCONUT chemicals with Jaccard-Tanimoto similarity [50,51] coefficient higher than 0.4 with respect to at least one of the negative compounds: 18,008 chemicals, structurally-similar to the query non-permeant compounds, were finally retrieved from the molecular database.

To refine this selection, we next applied a filter based on molecular descriptors that are expected to be related with permeability of chemicals: molecular weight (MW), Moriguchi octanol-water partition coefficient (MlogP), the number of donor atoms for H-bonds (nHDon) and the number of acceptor atoms for H-bonds (nHAcc). Following rules similar to those of Lipinski's drug-like index that predicts oral bioavailability of compounds [52], non-permeant chemicals were screened based on at least one of the following criteria: MW $>$ 500 or MlogP $<$ 1 or nHDon $>$ 5 or nHAcc $>$ 10. After screening, a total of 9,705 chemicals were retained, which were expected not to be able to permeate.

In addition, to further assess the model ability to correctly classify positive (permeant) chemicals, we screened the COCONUT database again to search for positive compounds likely to be permeant, applying a workflow similar to the one used to retrieve negative chemicals. This time, we used the 81 literature-derived positive chemicals as queries and selected compounds with Jaccard-Tanimoto similarity over ECFPs greater than 0.4 to at least one positive query, obtaining a set of 34,243 chemicals. Then, to reduce this set of compounds and identify those more likely to be permeant, we applied the following property filter [53]: MLOGP $>$ 3 and MW $<$ 300 and nHDon $<$ 3 and nHAcc $<$ 6. Only the chemicals that satisfied all these criteria (3,955) were finally selected as the external validation set of positive chemicals.

The procedures for data collection, curation and enrichment are graphically represented in Fig. 1.

3.2. QSPR modelling

3.2.1. Selection of molecular descriptors and model calculation

The QSPR model to predict permeability from the chemical structures of compounds was calculated using a dataset comprised of 81 positive (permeant) chemicals, which were found in literature as

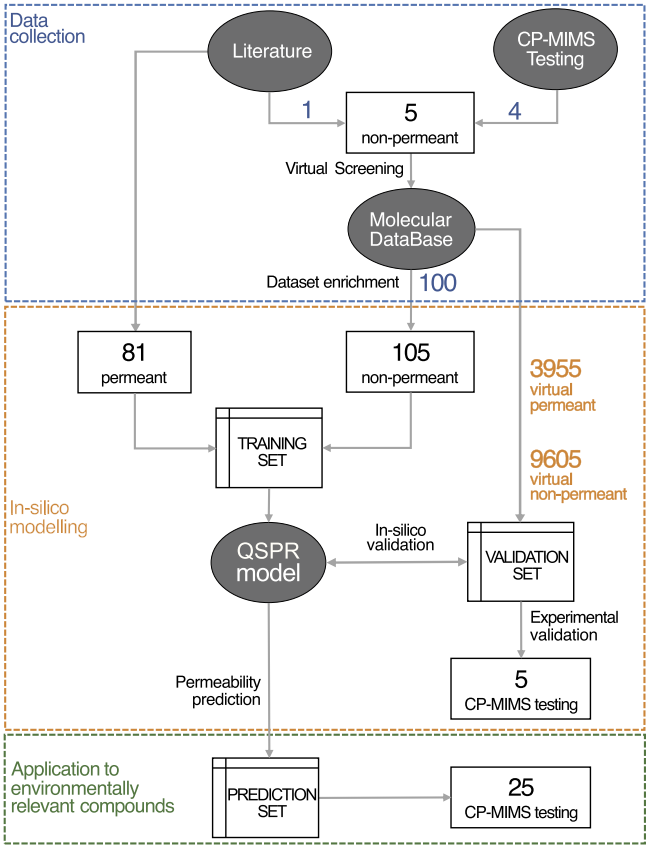


Fig. 1. Workflow diagram illustrating the adopted procedures for data collection, curation and enrichment.

previously explained, and 105 negative (non-permeant) chemicals (*i.e.*, 1 compound retrieved from literature, 4 experimentally tested and 100 randomly selected among the 9,705 negative chemicals from COCONUT database). On this dataset, the one-class classifier based on Potential Functions (PFs) was then applied to define the positive class decision boundary.

In addition, PFs were coupled with Genetic Algorithms (GAs) to select the optimal subset of molecular descriptors. The following 5 descriptors were selected to define the chemical space and build the classification rule: the superpendentic index (SPI) derived from a reduced topological distance matrix accounting only for distances of terminal atoms [54], the P_VSA–like descriptor based on the amount of van der Waals surface area (P_VSA_p_1) of the atoms with the lowest polarizability values [38], the number of aromatic hydroxyls (nArOH), the number of hydrogen bond–acceptor atoms (nHAcc), the CATS2-D_02_AL descriptor based on the counts of hydrogen bond–acceptor atoms and lipophilic atom-types at a topological distance of 2 bonds [40]. The selected molecular descriptors for all the training molecules are provided in Table S2 of the Supplementary Material.

The QSPR model based on PFs achieved very good accuracy in classification, with sensitivity (percentage of permeant chemicals correctly classified) equal to 93.8% and 90.1% in fitting and cross-validation, respectively, and specificity (% of non-permeant chemicals correctly classified) equal to 89.5% and 90.5% for both fitting and cross validation. This indicates the capability of the proposed modelling strategy to discriminate between chemicals which can permeate or not. Anyway, we applied further validation strategies to properly evaluate the predictive capability of PFs, that is, the accuracy when predicting chemicals which did not participate in the model calibration.

3.2.2. *In-silico model validation*

In order to further test the model specificity, the proposed QSPR model was applied to predict the 9,605 negative chemicals from COCONUT database, which were not previously used to train the model. The outcome of this validation test was more than satisfactory as 98.2% of chemicals were correctly predicted as negative (non-permeant). In addition, we assessed on the large set of 3,955 positive COCONUT-derived chemicals the QSPR ability to correctly classify positive (permeant) chemicals in a more robust way than using only the 81 positive literature-based chemicals. The QSPR model identified 86.8% of these potentially positive compounds as able to permeate, confirming a satisfactory predictive capability of the proposed model.

These validation results can be expected to be partially biased from the similarity search which was used to establish positive and negative chemicals used for testing; however, we consider these as promising results to support the quality of the proposed QSPR approach when predicting new chemicals. Anyway, we further estimated the predictive capability of the QSPR strategy with experimental testing, as discussed in the following paragraph.

A summary of statistical performance parameters of the proposed QSPR model is reported in Table 1.

Fig. 2 graphically represents the degree of separation between permeant (blue) and non-permeant (red) compounds in the chemical space defined by the molecular descriptors used for modelling. Since we employed 5 descriptors, we used Multidimensional Scaling (MDS) [55] to project compounds in a two-dimensional approximated chemical space. To calculate MDS coordinates, data were autoscaled and the Euclidean distance was used as distance metric.

Table 1
Statistical parameters for assessing the performance of the proposed QSPR classification model. Sensitivity and specificity refer to the percentage of permeant and non-permeant chemicals correctly classified, respectively.

	number of permeant (+)	number of non-permeant (–)	sensitivity (%)	specificity (%)
Fitting	81	105	93.8	89.5
Cross-validation	81	105	90.1	90.5
External validation	3,955	9,605	86.8	98.2

3.2.3. *Model validation by experimental testing*

Since the 3,955 COCONUT-derived positive analytes were only selected on the basis of their permeability potential, with no experimental evidence available in the literature, we performed supplementary experimental trials to further validate the QSPR model.

Among the 100 compounds predicted as permeant with the highest potential, the following 5 compounds were randomly selected to be experimentally tested: diphenylamine, dibenzothiophene, 2-amino-3,5-dibromobenzaldehyde, 3-bromofluoranthene, and cumene.

Pure standards of diphenylamine, dibenzothiophene, 2-amino-3,5-dibromobenzaldehyde, 3-bromofluoranthene, and cumene were analyzed via CP-MIMS. Due to their lower ionization efficiency in ESI, the permeation was measured with the FL detector for all analytes except for cumene, which was acquired using the PDA detector monitoring the 254 nm wavelength. All measurements were carried out following the conditions described in Section 2.2. Five out of five compounds were able to permeate, demonstrating the excellent prediction accuracy of the proposed model. The obtained results in terms of $\overline{Y}_{steady}$, $\overline{Y}_0$, $t_{10\%\text{-}90\%}$, and their relative normalized signal chronograms are shown in Table 2 and Fig. 3, respectively.

3.3. *QSPR prediction of environmentally relevant compounds with experimental testing*

In the last stage of this study, the QSPR model was applied to predict the permeability of 25 semi-volatile and non-volatile compounds to assess their suitability for CP-MIMS analysis. The selected compounds represent a broad range of environmentally relevant contaminants frequently occurred in environmental samples. They include pharmaceuticals and personal care products (ibuprofen, naproxen, acetylsalicylic acid, caffeine, theophylline, carbamazepine, diazepam, delorazepam, haloperidol, sildenafil citrate, tadalafil), pesticides (fipronil, chlorpyrifos, atrazine, trifluralin, oxybenzone), industrial and plastic-related chemicals (bisphenol A, 4-nitrophenol), and common fragrance or flavoring agents (benzaldehyde, methyl salicylate, β-citronellol, citral, linalool, geraniol). These compounds represent chemical classes differing in polarity, ionizability, molecular size, and functional groups, allowing assessment of CP-MIMS performance across a realistic spectrum of compounds. The experimental measurements of these molecules were carried out following the procedure described in the Section 2.2.

Experimental results obtained on the tested molecules are reported in Table 3, along with the QSPR model predicted labels, with the symbol + (positive) referring to the compounds able to permeate the PDMS membrane of CP-MIMS system and the symbol – (negative) to the non-permeant compounds. The molecular descriptors and the potential values calculated by the QSPR model are provided in Table S2 of the Supplementary Material. We considered permeating compounds those molecules exhibiting an average signal $\overline{Y}_{steady}$ clearly distinguishable from the average background signal $\overline{Y}_0$, as reported in the Section 2.3. The high standard deviation values observed for some analytes could be attributed to the lack of any signal pre-treatment during data processing. This approach was intentionally adopted to preserve the most realistic representation of analyte behavior.

Most compounds exhibited rapid permeation, with response times ranging from one to two minutes, thereby supporting the above-mentioned advantages of this technique, whose principal advantages are its high-speed and real-time analysis capability. In the cases of citral and geraniol, signal monitoring was performed using PDA (254 nm), due to the low efficiency in precursor ion formation. Theophylline, sildenafil

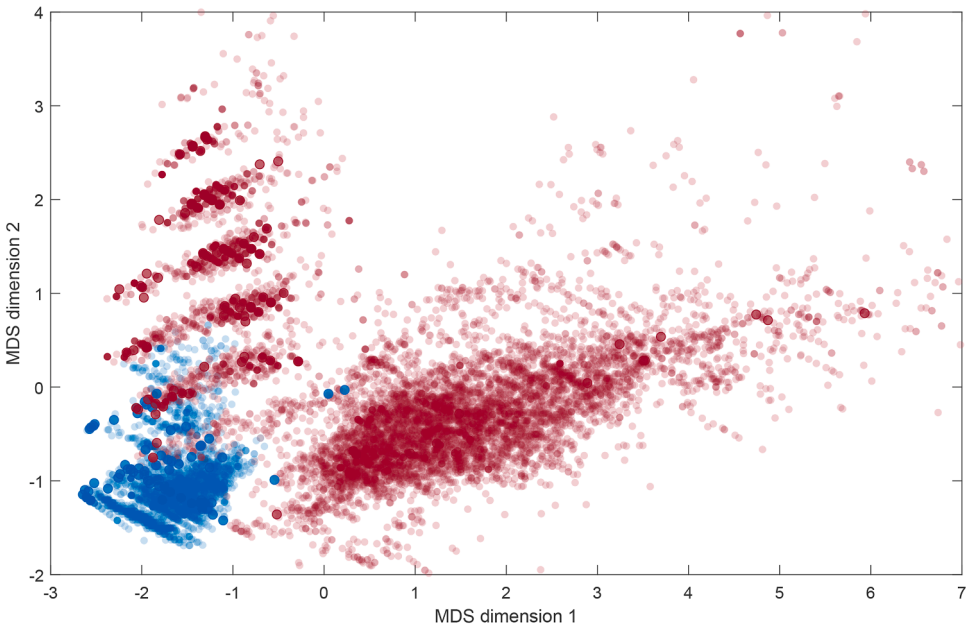


Fig. 2. Multidimensional Scaling plot: chemicals are colored according to the class labels (permeant in blue and non-permeant in red). The training chemicals (81 positive and 105 negative) are represented in dark color, while the chemicals used to validate the model (3,955 positive and 9,605 negative) in light color.

Table 2
Steady state signal ($\overline{Y}_{steady}$), background signal ($\overline{Y}_0$), and signal response time ($t_{10\%\text{-}90\%}$) for the five compounds experimentally tested for model validation.

ID	Name	CAS number	$\overline{Y}_{steady}$	$\overline{Y}_0$	$t_{10\%\text{-}90\%}$ *(min)*
EXP030	Diphenylamine	122-39-4	22 (± 1) ·10^4	11 (± 5) ·10	1.1 (± 0.5)
EXP031	Dibenzothiophene	132-65-0	61 (± 4) ·10^2	5 (± 3) ·10^2	1.5 (± 0.2)
EXP032	2-Amino-3,5-dibromobenzaldehyde	50910-55-9	11 (± 1) ·10^4	20 (± 2) ·10	1.2 (± 0.5)
EXP033	3-bromofluoranthene	13438-50-1	59 (± 4) ·10^3	19 (± 1) ·10^2	1.8 (± 0.1)
EXP034	Cumene	98-82-8	43 (± 4) ·10^2	54 (± 2) ·10	1.0 (± 0.4)

citrate, and haloperidol did not exhibit detectable steady-state signals in CP-MIMS using PDMS membranes. Their negligible permeation can be linked to unfavorable membrane/water partitioning and limited diffusivity within the polymer matrix. Theophylline, structurally related to theobromine, shows comparable behavior: the polar xanthine scaffold, rich in hydrogen-bond donors and acceptors, shows very low membrane affinity. Sildenafil citrate and haloperidol possess relatively large hydrodynamic volumes and rigid molecular structures that can limit diffusion through the membrane, resulting in undetectable permeation.

Based on the CP-MIMS experimental results shown in Table 3, we could therefore conclude that there is an overall good agreement between model predictions and experimental evidence; 22 out of the 25 environmentally relevant compounds were able to permeate the membrane and 19 of these 22 compounds (86%) were correctly predicted as positive (permeant) by the QSPR model, with only fipronil, trifluralin, and bisphenol A incorrectly predicted as non-permeant. However, the experimental test carried out on bisphenol A revealed that this compound did not permeate employing 100% of MeOH (v:v) as AP, but required the formation of a PIM. Therefore, this behavior is consistent with the QSPR model, which predicts no-permeability for bisphenol A depending on its physicochemical properties rather than on the experimental conditions. On the opposite, only 3 of the tested compounds were not able to permeate and 2 of them (67%) were correctly predicted as non-permeant chemicals.

Multidimensional Scaling (MDS) was used to represent the distribution of the environmentally relevant compounds (22 positive and 3 negative) in the chemical space defined by the training chemicals (81 positive and 105 negative). Also in this case, molecular descriptors were autoscaled and the Euclidean distance was used as distance metric. The score plot of the first two MDS dimensions is represented in Fig. 4: despite a minor overlap, a clear distinction between permeant (blue) and non-permeant (red) molecules can be seen. The compound exhibiting the fastest permeation was linalool; as illustrated in Fig. 4 a), its steep rising front of the normalized signal clearly indicates that permeation occurs very rapidly, with a of $t_{10\%\text{-}90\%}$ of 0.8 min. The slowest permeant compound was tadalafil which exhibited signal response times in the range of 4–5 minutes. This analyte possesses moderate to low diffusion coefficients, resulting in longer duty cycles during CP-MIMS acquisition. As illustrated in Fig. 4 b) the gentle rising front of the normalized signal clearly indicates that permeation occurs very slowly for tadalafil with a $t_{10\%\text{-}90\%}$ value of 5 min.

The difference in the time scales used in the chronograms of the faster and the slowest compounds is intentional and highlights the rapid permeation of linalool, resulting in a faster duty-cycle. The original MS recorded chronograms for these analytes are provided in Fig. S1 of the Supplementary Material. Each analytical cycle consisted of monitoring the signal rise until steady-state permeation was achieved, followed by the signal decay after rinsing the membrane with fresh MeOH, across three replicate measurements.

4. Conclusions

In this study, we integrated QSPR modeling with CP-MIMS to predict the membrane permeability of analytes directly from their molecular structure. By combining literature data, targeted CP-MIMS measurements, and extensive structural enrichment from a large database, we created a chemically diverse and statistically balanced dataset suitable for model development. The one-class classifier based on Potential

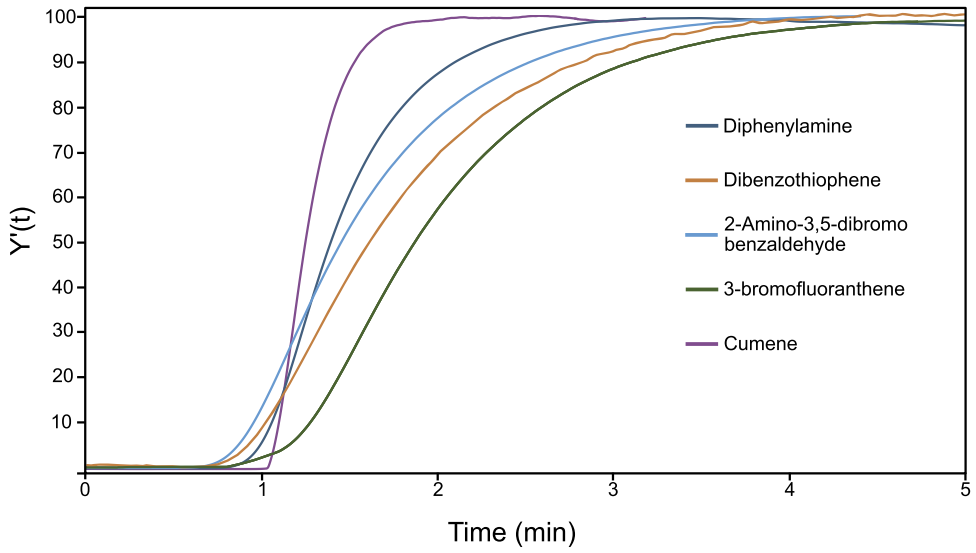


Fig. 3. Comparison of normalized signal chronograms for the five experimentally tested compounds for model validation.

Table 3
QSPR model predictions and experimental results of CP-MIMS measurements for 25 environmental relevant compounds; the symbol + (positive) refers to the compounds able to permeate, the symbol – (negative) to the non-permeant compounds.

ID	Name	CAS number	$\overline{Y}_{steady}$	$\overline{Y}_0$	$t_{10\%\text{-}90\%\ (min)}$	*Experimental permeation*	*QSPR prediction*
EXP001	Haloperidol	52-86-8	433 (±1) 10^2	421 (±1) 10^2	n.a.	–	–
EXP004	Sildenafil citrate	171599-83-0	121 (±4) 10^2	115 (±2) 10^2	n.a.	–	–
EXP007	Theophylline	58-55-9	72 (±2) 10^2	70 (±1) 10^2	n.a.	–	+
EXP008	Benzophenone-3 (oxybenzene)	131-57-7	30 (±2) 10^4	90 (±8) 10^2	1.6 (±0.3)	+	+
EXP009	Bisphenol A	80-05-7	74 (±3) 10^2	60 (±2) 10^2	2.6 (±0.1)	+	–
EXP010	Trifluralin	1582-09-8	115 (±2) 10^2	63 (±5) 10^2	1.64 (±0.06)	+	–
EXP011	4-Nitrophenol	100-02-7	193 (±6) 10^2	30 (±2) 10^2	0.8 (±0.2)	+	+
EXP012	Acetylsalicylic acid	50-78-2	99 (±5) 10^2	62 (±5) 10^2	1.1 (±0.4)	+	+
EXP013	Atrazine	1912-24-9	201 (±2) 10^3	7 (±2) 10^3	1.9 (±0.2)	+	+
EXP014	Benzaldehyde	100-52-7	37 (±1) 10^2	262 (±4) 10	1.1 (±0.3)	+	+
EXP015	Caffeine	58-08-2	72 (±1) 10^2	61 (±2) 10^2	4 (±1)	+	+
EXP016	Carbamazepine	298-46-4	269 (±3) 10^2	100 (±6) 10^2	3.6 (±0.1)	+	+
EXP017	Chlorpyrifos	2921-88-2	136 (±2) 10^3	279 (±3) 10^2	1.7 (±0.7)	+	+
EXP018	Citral	5392-40-5	161 (±3) 10^3	69 (±5) 10^2	1.03 (±0.05)	+	+
EXP019	Delorazepam	2894-67-9	71 (±1) 10^3	73 (±2) 10^2	1.9 (±0.1)	+	+
EXP020	Diazepam	439-14-5	165 (±2) 10^4	73 (±3) 10^3	2.1 (±0.5)	+	+
EXP021	Diclofenac	15307-86-5	214 (±3) 10^3	46 (±8) 10^2	1.44 (±0.05)	+	+
EXP022	Fipronil	120068-37-3	1633 (±3) 10^3	173 (±1) 10^2	1.8 (±0.8)	+	–
EXP023	Geraniol	624-15-7	70 (±7) 10^2	42 (±5)	1.02 (±0.07)	+	+
EXP024	Ibuprofen	58560-75-1	208 (±8) 10^3	9 (±5) 10^3	1.4 (±0.5)	+	+
EXP025	Linalool	78-70-6	17 (±4) 10^4	88 (±7) 10^2	0.8 (±0.7)	+	+
EXP026	Methyl Salicylate	119-36-8	276 (±8) 10^2	496 (±2) 10	1.0 (±0.1)	+	+
EXP027	Naproxen	23981-80-8	512 (±9) 10^2	105 (±2) 10^2	1.03 (±0.09)	+	+
EXP028	Tadalafil	304683-09-8	83 (±4) 10^2	60 (±3) 10^2	5.0 (±0.7)	+	+
EXP029	β-citronellol	106-22-9	15 (±1) 10^3	930 (±8) 10	1.7 (±0.1)	+	+

Functions, refined through Genetic Algorithms, identified a subset of six molecular descriptors capturing the structural features related to PDMS membrane permeation. The use of a one-class classification approach, defined exclusively on permeant molecules, allowed us to exploit the prevalence of permeant data in the literature, ensuring a reliable and representative definition of the permeant chemical space that is primarily driven by the quality of the available experimental data

The classification model exhibited high sensitivity and specificity across multiple validation levels, including external testing on chemicals retrieved from databases, as well as on analytes for which membrane permeability was experimentally determined. When applied to environmentally relevant contaminants, the model also provided accurate qualitative predictions for most compounds, with discrepancies highlighting complex cases that may reflect borderline structural features, alternative ionization behaviors, or analytical challenges. In the development of a CP-MIMS method, the choice of the ionization technique represents a key factor, as it directly affects detection sensitivity and selectivity for compounds belonging from different classes. While mass spectrometric detection offers several analytical advantages, permeation studies are independent of the ionization process enabling the analytical set-up to be tuned according to the physicochemical properties of the analytes or the instrumentation available in the laboratory. When mass spectrometric detection is not available, alternative approaches, such as spectrophotometric techniques, may be adopted. These results underline both the strengths and limitations of the CP-MIMS technique and emphasize the value of computational prescreening for method development. In fact, the proposed QSPR framework can enhance and foster application of CP-MIMS by screening analytes and early identifying those who can permeate. This strategy offers a practical and scalable add-on to anticipate analyte behavior at the membrane interface, yielding deeper insight into the mechanistic basis of structure–permeability relationships. By reducing the need for

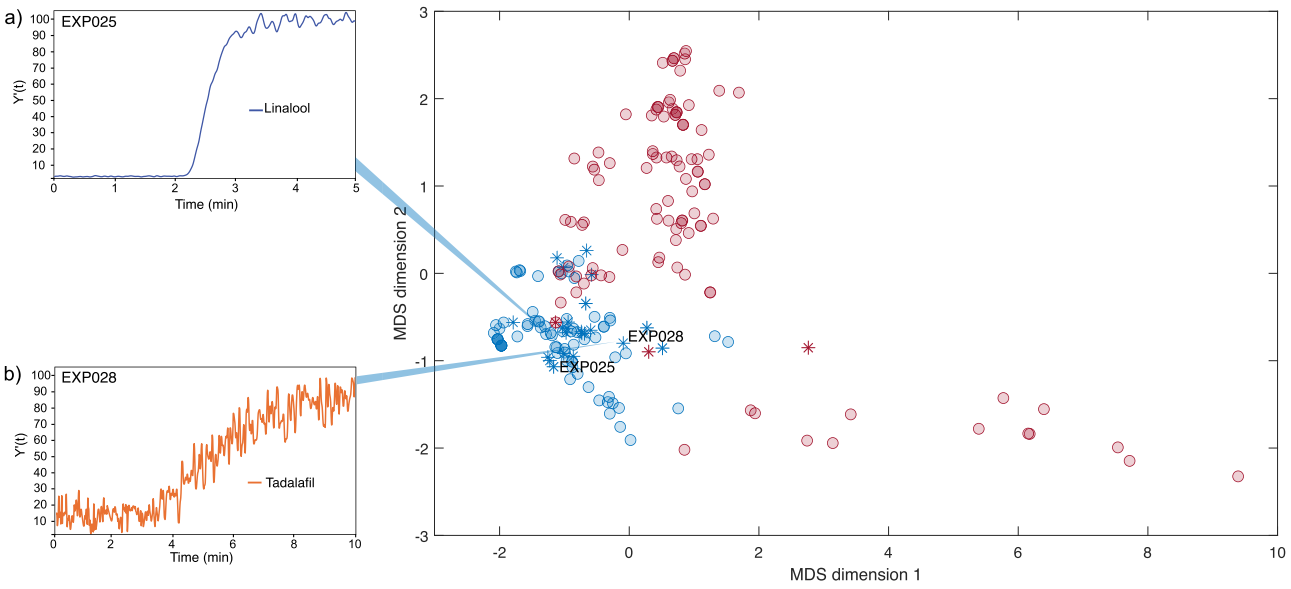


Fig. 4. Multidimensional Scaling score plot; chemicals are colored according to class labels (permeant in blue and non-permeant in red); training chemicals (81 positive and 105 negative) are represented with circles, environmental relevant compounds used to validate the model (22 positive and 3 negative) with asterisks. The arrows highlighting the labelled blue asterisks correspond to the faster (EXP025) and the slowest (EXP028) permeating compounds and their relative normalized signal chronograms: a) linalool and b) tadalafil.

experimental measurements, our approach improves sustainability and expands the applicability of the CP-MIMS system in greener analytical and environmental studies.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Enmanuel Cruz Muñoz: Writing – review & editing, Investigation, Formal analysis, Data curation. **Veronica Termopoli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Conceptualization. **Davide Ballabio:** Writing – original draft, Validation, Software, Investigation, Formal analysis. **Marco Orlandi:** Supervision, Funding acquisition. **Viviana Consonni:** Writing – original draft, Validation, Methodology, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare no competing financial interest.

Acknowledgements

Special thanks are due to Dr. Sofia Mescieri for her help with experimental operation and acquisition of a portion of the data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.greeac.2026.100328.

Data availability

Data are available in the Supplementary material.

References

[1] H.C. Huang, H.H. Chung, J.Y. Yu, B.R. Chen, M.Y. Wang, C.C. Hsu, Development and multicenter validation of on-site breast cancer diagnosis using paper spray ionization miniature mass spectrometry, Commun. Med. 5 (2025), https://doi.org/10.1038/s43856-025-00930-7.

[2] L.M. Schmidtke, L. Jiang, M. Dumlao, W.A. Donald, Direct ambient mass spectrometry for food, beverage, and agricultural sample analysis and research, Mass Spectrom. Rev. (2025), https://doi.org/10.1002/mas.70001.

[3] S. Mondal, U. Pandey, S. Chakrabarti, P. Pahchan, D. Koner, S. Banerjee, Rapid and reagent-free analysis of dried blood spot by paper spray mass spectrometry reveals sex: implications in forensic investigations, J. Proteome Res. 24 (2025) 2314–2323, https://doi.org/10.1021/acs.jproteome.4c00798.

[4] E.T. Krogh, C.G. Gill, Membrane introduction mass spectrometry (MIMS): a versatile tool for direct, real-time chemical measurements, J. Mass Spectrom. 49 (2014) 1205–1213, https://doi.org/10.1002/jms.3447.

[5] E.T. Krogh, C.G. Gill, Chapter Seven - Condensed Phase Membrane Introduction Mass Spectrometry – Continuous, Direct and Online Measurements in Complex Samples, in: A. Cappiello, P. Palma (Eds.), Advances in the Use of Liquid Chromatography Mass Spectrometry (LC-MS), Elsevier, 2018, pp. 173–203, https://doi.org/10.1016/bs.coac.2017.06.010.

[6] G.W. Vandergrift, E.T. Krogh, C.G. Gill, Polymer inclusion membranes with condensed phase membrane introduction mass spectrometry (CP-MIMS): improved analytical response time and sensitivity, Anal. Chem. 89 (2017) 5629–5636, https://doi.org/10.1021/acs.analchem.7b00908.

[7] J. Monaghan, Q. Xin, R. Aplin, A. Jaeger, N.E. Heshka, L.J. Hounjet, C.G. Gill, E. T. Krogh, Aqueous naphthenic acids and polycyclic aromatic hydrocarbons in a meso-scale spill tank affected by diluted bitumen analyzed directly by membrane introduction mass spectrometry, J. Hazard. Mater. 440 (2022), https://doi.org/10.1016/j.jhazmat.2022.129798.

[8] G.W. Vandergrift, J. Monaghan, E.T. Krogh, C.G. Gill, Direct analysis of polyaromatic hydrocarbons in soil and aqueous samples using condensed phase membrane introduction tandem mass spectrometry with low-energy liquid electron ionization, Anal. Chem. 91 (2019) 1587–1594, https://doi.org/10.1021/acs.analchem.8b04949.

[9] G.W. Vandergrift, E.T. Krogh, C.G. Gill, Direct, isomer-specific quantitation of polycyclic aromatic hydrocarbons in soils using membrane introduction mass spectrometry and chemical ionization, Anal. Chem. 92 (2020) 15480–15488, https://doi.org/10.1021/ACS.ANALCHEM.0C03259.

[10] V. Termopoli, M. Piergiovanni, D. Ballabio, V. Consonni, E.C. Muñoz, F. Gosetti, Condensed phase membrane introduction mass spectrometry: a direct alternative to fully exploit the mass spectrometry potential in environmental sample analysis, Sep. 2023 10 (2023) 10, https://doi.org/10.3390/SEPARATIONS10020139.

[11] V. Termopoli, E. Torrisi, G. Famiglini, P. Palma, G. Zappia, A. Cappiello, G. W. Vandergrift, M. Zvekic, E.T. Krogh, C.G. Gill, Mass Spectrometry Based Approach for Organic Synthesis Monitoring, Anal. Chem. 91 (2019) 11916–11922, https://doi.org/10.1021/acs.analchem.9b02681.

[12] M. Zvekic, G.W. Vandergrift, C.C. Tong, C.G. Gill, E.T. Krogh, Monitoring microplastic-contaminant sorption processes in real-time using membrane introduction mass spectrometry, Env. Sci. Process. Impacts 25 (2023) 1169–1180, https://doi.org/10.1039/d3em00083d.

[13] M. Piergiovanni, V. Termopoli, C. Maffezzoni, N. Riboni, V. Consonni, F. Bianchi, M. Mattarozzi, D. Ballabio, M. Careri, Condensed phase membrane introduction mass spectrometry: a new frontier for the real-time monitoring of hazardous chemical migration from food contact materials, Green Anal. Chem. 12 (2025), https://doi.org/10.1016/j.greeac.2024.100199.
[14] M.A. LaPack, J.C. Tou, C.G. Enke, Membrane mass spectrometry for the direct trace analysis of volatile organic compounds in air and water, Anal. Chem. 62 (1990) 1265–1271, https://doi.org/10.1021/ac00212a013.
[15] D.W. Janes, C.J. Durning, D.M. van Pel, M.S. Lynch, C.G. Gill, E.T. Krogh, Modeling analyte permeation in cylindrical hollow fiber membrane introduction mass spectrometry, J. Memb. Sci. 325 (2008) 81–91, https://doi.org/10.1016/j.memsci.2008.07.033.
[16] G. Hoch, B. Kok, A mass spectrometer inlet system for sampling gases dissolved in liquid phases, Arch. Biochem. Biophys 101 (1963) 160–170, https://doi.org/10.1016/0003-9861(63)90546-0.
[17] K.D. Duncan, M.D. Willis, E.T. Krogh, C.G. Gill, A miniature condensed-phase membrane introduction mass spectrometry (CP-MIMS) probe for direct and on-line measurements of pharmaceuticals and contaminants in small, complex samples, Rapid Commun. Mass Spectrom. 27 (2013) 1213–1221, https://doi.org/10.1002/rcm.6560.
[18] K.D. Duncan, J.A. Hawkes, M. Berg, B. Clarijs, C.G. Gill, J. Bergquist, I. Lanekoff, E. T. Krogh, Membrane sampling separates naphthenic acids from biogenic dissolved organic matter for direct analysis by mass spectrometry, Env. Sci. Technol. 56 (2022) 3096–3105, https://doi.org/10.1021/ACS.EST.1C07359.
[19] J. Stanstrup, S. Neumann, U. Vrhovšek, PredRet: Prediction of Retention Time by Direct Mapping between Multiple Chromatographic Systems, Anal. Chem. 87 (2015) 9421–9428, https://doi.org/10.1021/ACS.ANALCHEM.5B02287.
[20] A. Cherkasov, E.N. Muratov, D. Fourches, A. Varnek, I.I. Baskin, M. Cronin, J. Dearden, P. Gramatica, Y.C. Martin, R. Todeschini, V. Consonni, V.E. Kuz'Min, R. Cramer, R. Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, QSAR Modeling: where have you been? Where are you going to? J. Med. Chem. 57 (2014) 4977–5010, https://doi.org/10.1021/JM4004285.
[21] S.H. Maguire, M. Zvekic, A. Jaeger, J. Monaghan, E.T. Krogh, H.A. Wiebe, Physicochemical properties of tire-derived para-phenylenediamine quinones - a comparison of experimental and computational approaches, Env. Sci. Process. Impacts 27 (2025) 2550–2563, https://doi.org/10.1039/d5em00153f.
[22] Milano Chemometrics and QSAR Research Group Università degli Studi di Milano-Bicocca, MIMS Toolbox for MATLAB. https://michem.unimib.it/download/matlab-toolboxes/mims-toolbox-for-matlab/ (accessed December 5, 2025).
[23] J. Monaghan, A. Jaeger, A.R. Agua, R.S. Stanton, M. Pirrung, C.G. Gill, E.T. Krogh, A Direct Mass Spectrometry Method for the Rapid Analysis of Ubiquitous Tire-Derived Toxin N-(1,3-Dimethylbutyl)-N′-phenyl-p-phenylenediamine Quinone (6-PPDQ), Env. Sci. Technol. Lett. 8 (2021), https://doi.org/10.1021/acs.estlett.1c00794.
[24] M.D. Willis, K.D. Duncan, E.T. Krogh, C.G. Gill, Delicate polydimethylsiloxane hollow fibre membrane interfaces for condensed phase membrane introduction mass spectrometry (CP-MIMS), Rapid Commun. Mass Spectrom. 28 (2014) 671–681, https://doi.org/10.1002/RCM.6828.
[25] K.D. Duncan, E.P.B. McCauley, E.T. Krogh, C.G. Gill, Characterization of a condensed-phase membrane introduction mass spectrometry (CP-MIMS) interface using a methanol acceptor phase coupled with electrospray ionization for the continuous on-line quantitation of polar, low-volatility analytes at trace leve…, Rapid, Commun. Mass Spectrom. 25 (2011) 1141–1151, https://doi.org/10.1002/RCM.4967.
[26] K.D. Duncan, D.A. Volmer, C.G. Gill, E.T. Krogh, Rapid screening of carboxylic acids from waste and surface waters by ESI-MS/MS using barium ion chemistry and on-line membrane sampling, J. Am. Soc. Mass Spectrom 27 (2015) 443–450, https://doi.org/10.1007/S13361-015-1311-Y.
[27] S.A. Borden, H.N. Damer, E.T. Krogh, C.G. Gill, Direct quantitation and characterization of fatty acids in salmon tissue by condensed phase membrane introduction mass spectrometry (CP-MIMS) using a modified donor phase, Anal. Bioanal. Chem. 411 (2018) 291–303, https://doi.org/10.1007/S00216-018-1467-Y, 2018 411:2.
[28] K.D. Duncan, G.W. Vandergrift, E.T. Krogh, C.G. Gill, Ionization suppression effects with condensed phase membrane introduction mass spectrometry: Methods to increase the linear dynamic range and sensitivity, J. Mass Spectrom. 50 (2014) 437–443, https://doi.org/10.1002/JMS.3544.
[29] G.W. Vandergrift, W. Lattanzio-Battle, T.R. Rodgers, J.B. Atkinson, E.T. Krogh, C. G. Gill, Geospatial assessment of trace-level benzophenone-3 in a fish-bearing river using direct mass spectrometry, ACS ES&T Water 2 (2022) 262–267, https://doi.org/10.1021/ACSESTWATER.1C00205.
[30] G.W. Vandergrift, W. Lattanzio-Battle, E.T. Krogh, C.G. Gill, Condensed Phase Membrane Introduction Mass Spectrometry with In Situ Liquid Reagent Chemical Ionization in a Liquid Electron Ionization Source (CP-MIMS-LEI/CI), J. Am. Soc. Mass Spectrom 31 (2020) 908–916, https://doi.org/10.1021/JASMS.9B00143.
[31] J.F. Feehan, J. Monaghan, C.G. Gill, E.T. Krogh, Direct measurement of acid dissociation constants of trace organic compounds at nanomolar levels in aqueous solution by condensed phase–membrane introduction mass spectrometry, Env. Toxicol. Chem 38 (2019) 1879–1889, https://doi.org/10.1002/ETC.4519.
[32] T.M. Zarkovic, S.A. Borden, E.T. Krogh, C.G. Gill, A passive membrane system for on-line mass spectrometry reagent addition, Rapid Commun. Mass Spectrom. 37 (2023) e9487, https://doi.org/10.1002/RCM.9487.
[33] V. Termopoli, G. Famiglini, P. Palma, A. Cappiello, G.W. Vandergrift, E.T. Krogh, C.G. Gill, Condensed phase membrane introduction mass spectrometry with direct electron ionization: on-line measurement of PAHs in complex aqueous samples, J. Am. Soc. Mass Spectrom 27 (2016) 301–308, https://doi.org/10.1007/s13361-015-1285-9.
[34] M. Sorokina, P. Merseburger, K. Rajan, M.A. Yirik, C. Steinbeck, COCONUT online: collection of open natural products database, J. Cheminformatics 13 (2021), https://doi.org/10.1186/S13321-020-00478-9, 2021 13:12-.
[35] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, J. Chem. Inf. Comput. Sci. 28 (1988) 31–36, https://doi.org/10.1021/ci00057a005.
[36] R. Todeschini, V. Consonni, Molecular Descriptors for Chemoinformatics, 2009, Wiley-VCH Verlag GmbH & Co. KGaA, 2009, https://doi.org/10.1002/9783527628766.
[37] J. Devillers, A.T. Balaban, Topological Indices and Related Descriptors in QSAR and QSPR,, 1st Edition, CRC Press, 1999.
[38] P. Labute, A widely applicable set of descriptors, J. Mol. Graph. Model. 18 (2000) 464–477, https://doi.org/10.1016/S1093-3263(00)00068-1.
[39] A. Mauri, V. Consonni, M. Pavan, R. Todeschini, DRAGON SOFTWARE: AN EASY APPROACH TO MOLECULAR DESCRIPTOR CALCULATIONS, n.d.
[40] G. Schneider, W. Neidhart, T. Giller, G. Schmid, "Scaffold-Hopping" by topological pharmacophore search: a contribution to virtual screening. doi: 10.1002/(SICI)1521-3773(19991004)38:19.
[41] R.G. Brereton, One-class classifiers, J. Chemom 25 (2011) 225–246, https://doi.org/10.1002/CEM.1397.
[42] F. Sahigara, K. Mansouri, D. Ballabio, A. Mauri, V. Consonni, R. Todeschini, F. Sahigara, K. Mansouri, D. Ballabio, A. Mauri, V. Consonni, R. Todeschini, Comparison of different approaches to define the applicability domain of QSAR models, Molecules 17 (2012) 4791–4810, https://doi.org/10.3390/MOLECULES17054791, 2012, Vol. 17, Pages 4791-4810.
[43] M. Forina, C. Armanino, R. Leardi, G. Drava, A class-modelling technique based on potential functions, J. Chemom. 5 (1991) 435–453, https://doi.org/10.1002/CEM.1180050504.
[44] R. Leardi, R. Boggia, M. Terrile, Genetic algorithms as a strategy for feature selection, J. Chemom. 6 (1992) 267–281, https://doi.org/10.1002/CEM.1180060506.
[45] Milano Chemometrics and QSAR Research Group's personal version, Dragon 7 – Software for the calculation of molecular descriptors, Milano Chemometrics and QSAR Research Group's personal version.
[46] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models. PLS-DA, Anal. Methods 5 (2013) 3790–3798, https://doi.org/10.1039/C3AY40582F.
[47] D. Boldini, D. Ballabio, V. Consonni, R. Todeschini, F. Grisoni, S.A. Sieber, Effectiveness of molecular fingerprints for exploring the chemical space of natural products, J. Cheminformatics 16 (2024), https://doi.org/10.1186/S13321-024-00830-3, 2024 16:135-.
[48] G. Landrum, P. Tosco, B. Kelley, R. Rodriguez, D. Cosgrove, R. Vianello, Sriniker, P. Gedeck, G. Jones, E. Kawashima, N. Schneider, D. Nealschneider, A. Dalke, tadhurst-cdd, M. Swain, B. Cole, S. Turck, A. Savelev, A. Vaucher, M. Wójcikowski, H. Faara, I. Take, N. Maeder, R. Walker, V.F. Scalfani, D. Probst, K. Ujihara, A. Pahl, G. Godin, J. Lehtivarjo, rdkit/rdkit: 2025_09_1 (Q3 2025) Release, (2025). doi: 10.5281/zenodo.591637.
[49] D. Rogers, M. Hahn, Extended-Connectivity Fingerprints, J. Chem. Inf. Model. 50 (2010) 742–754, https://doi.org/10.1021/CI100050T.
[50] P. Jaccard, The distribution of the flora in the alpine zone, New Phytol. 11 (1912) 37–50, https://doi.org/10.1111/J.1469-8137.1912.
[51] D.J. Rogers, T.T. Tanimoto, A computer program for classifying plants, Sci. (1979) 132 (1960) 1115–1118, https://doi.org/10.1126/SCIENCE.132.3434.1115.
[52] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 46 (2001) 3–26, https://doi.org/10.1016/S0169-409X(00)00129-0.
[53] M. Congreve, R. Carr, C. Murray, H. Jhoti, A 'Rule of Three' for fragment-based lead discovery? Drug Discov. Today 8 (2003) 876–877, https://doi.org/10.1016/S1359-6446(03)02831-9.
[54] S. Gupta, M. Singh, A.K. Madan, Superpendentic Index: a novel topological descriptor for predicting biological activity, J. Chem. Inf. Comput. Sci. 39 (1999) 272–277, https://doi.org/10.1021/CI980073Q.
[55] J.B. Kruskal, Nonmetric multidimensional scaling: a numerical method, Psychometrika 29 (1964) 115–129, https://doi.org/10.1007/BF02289694, 1964 29:2.