Environmental Science Processes & Impacts

PAPER

View Article Online
View Journal | View Issue

Cite this: *Environ. Sci.: Processes Impacts*, 2017, **19**, 307

Predicting the phospholipophilicity of monoprotic positively charged amines†

S. T. J. Droge,‡*[a] J. L. M. Hermens,[a] S. Gutsell,[b] J. Rabone[b] and G. Hodges[b]

Received 15th November 2016
Accepted 23rd January 2017

DOI: 10.1039/c6em00615a

rsc.li/process-impacts

The sorption affinity of eighty-six charged amine structures to phospholipid monolayers (log K_{IAM}) was determined using immobilized artificial membrane high-performance liquid chromatography (IAM-HPLC). The amine compounds covered the most prevalent types of polar groups, widely ranged in structural complexity, and included forty-seven pharmaceuticals, as well as several narcotics and pesticides. Amine type specific corrective increments were used to align log K_{IAM} data with bilayer membrane sorption coefficients (K_{MW}(IAM)). Using predicted sorption affinities of neutral amines, we evaluated the difference (scaling factor Δ_{MW}) with the measured log K_{MW}(IAM) for cationic amines. The Δ_{MW} values were highly variable, ranging from −2.37 to +2.3 log units. For each amine type, polar amines showed lower Δ_{MW} values than hydrocarbon based amines ($C_xH_yN^+$). COSMOmic software was used to directly calculate the partitioning coefficient of ionic structures into a phospholipid bilayer ($K_{DMPC-W,cation}$), including quaternary ammonium compounds. The resulting root mean square error (RMSE) between log $K_{DMPC-W,cation}$ and log K_{MW}(IAM) was 0.83 for all eighty-six polar amines, and 0.47 for sixty-eight $C_xH_yN^+$ amines. The polar amines were then split into five groups depending on polarity and structural complexity, and corrective increments for each group were defined to improve COSMOmic predictions. Excluding only the group with sixteen complex amine structures (≥4 polar groups, M_w > 400, including several macrolide antibiotics), the resulting RMSE for corrected $K_{DMPC-W,cation}$ values improved to 0.45 log units for the remaining set of 138 polar and $C_xH_yN^+$ amines.

Environmental impact

Ionisable organic compounds, and particularly permanently charged compounds, pose a great challenge to accurate environmental fate/risk modelling, because most models have been designed to deal with neutral compounds only. Predictions for accumulation of ionogenic chemicals into tissues of exposed organisms define the starting point of toxicodynamic evaluations, such as adverse outcome pathways and bioaccumulation factors. The current study provides a large and structurally diverse experimental data set on the sorption affinity of organic cations to phospholipid membranes. The IAM-HPLC based data form a solid data set to evaluate the applicability domain of quantum-chemistry based molecular predictions of the cell membrane–water partition coefficients. We provide insight into the need for IAM-HPLC measurements and advanced molecular simulations over currently used octanol–water based predictions for different types of organic cations.

1. Introduction

According to an analysis of the 1999 World Drug Index (WDI; 51 596 compounds),[1] 63% of all drugs have ionisable groups between pH 2 and 12. Of these ionisable drugs, 42.9% are compounds with a single basic group.[1] Analysis of a smaller set of ionisable drugs indicated that 77% of the basic groups have a dissociation constant (pK_a) larger than 7.5.[2] Many narcotics (*e.g.* amphetamines, heroin, cocaine) are ionisable bases with a pK_a larger than 7.5.[2,3] These "strongly" basic drugs are thus predominantly protonated (positively charged, cationic) at the physiological pH in most tissues of target organisms (humans, cattle, poultry, *etc.*). Basic amines are also components in many other types of products, such as pesticides, biocides, herbicides, cleaning agents and industrial chemicals.[4] In contaminated environments, trace level concentrations of these, often bioactive, ionisable amines equilibrate with tissues of non-target organisms.[5] Chemical risk assessment models need to adequately capture the partitioning behaviour of these ionisable amines between different abiotic compartments such as soil/sediment and water,[6–9] and partitioning into biotic compartments, in order to predict concentrations at the primary site of toxic action.[10]

[a]*Institute for Risk Assessment Sciences, Utrecht University, Yalelaan 104, 3508 TD Utrecht, The Netherlands*

[b]*Safety and Environmental Assurance Centre, Unilever, Sharnbrook, Bedford, UK*

† Electronic supplementary information (ESI) available: Additional tables with information on chemical properties, uncorrected and corrected predictions and measurements and figures with solvent mixture data. See DOI: 10.1039/c6em00615a

‡ Current address: UvA-IBED, Postbus 94248, 1090 GE Amsterdam, The Netherlands. E-mail: steven.droge@gmail.com; Tel: +31 205257437.

Toxicokinetic modelling of drugs and xenobiotic contaminants depends in many ways on their lipophilicity, as this relates to retention time in the body, accumulation in certain tissues and passive membrane permeation. For most neutral compounds, lipophilicity is fairly well predicted by the octanol–water partition coefficient (log K_{OW}).[11] For largely ionised bases, however, one has to take into account that the lipophilic properties of the neutral and charged species can differ substantially. The octanol–water distribution coefficient D_{OW} of organic cations is typically around a factor of a thousand lower than that of the corresponding neutral species,[12–14] and this most likely reflects the distribution coefficient between neutral adipose lipids (triglycerides) and water ($K_{adipose–W}$). Differences in sorption affinities to cell membrane phospholipids (K_{MW}), however, are reported to be lower for the ionized bases by only a factor of thirty or less.[12–14] Ionic compounds interact the strongest with the membrane at the oppositely charged phospholipid moieties in the headgroup, and the orientation of the rest of the molecule aligns with the hydrophobicity profile of the ordered phospholipids in the bilayer, which can strongly differ from the favourable depth and orientation of the neutral form.[15,16] Given the conceptual differences between partitioning of ionic species into octanol (*i.e.* neutral bulk solvent) and partitioning into membranes (*i.e.* an ordered, zwitterionic phase), there is no simple correlation to be expected between the pH-dependent octanol–water distribution coefficient D_{OW} and the membrane–water distribution coefficient (D_{MW}) for ionisable bases.[17–19] Although there is a close link between the sorption affinity to phospholipid membranes and passive membrane permeability for neutral species, this is likely not the case for ionic species.[17] Hydrophobic cations may sorb relatively strongly to the phosphate anions in the head group region, while at the same time the passive diffusion through the hydrophobic membrane interior may be energetically unfavourable.[17]

The D_{MW} of ionizable compounds is thus an important parameter for any toxicological modeling effort. Measurements of the D_{MW} are ideally performed using standardized protocols, but no OECD or ASTM guidelines are available. Artificial unilamellar bilayer liposomes with a relatively homogenous size distribution can be generated from the pure phospholipid material through repeated extrusion steps.[20] Such dissolved liposomes are widely considered as representative model membranes to determine D_{MW}, and can be prepared in any aqueous buffer to determine the pH-dependency of D_{MW} for IOCs.[21] Liposomes are not readily separated from the aqueous phase by centrifugation, though, and therefore dialysis systems or passive samplers are often deployed to distinguish freely dissolved concentrations and membrane-sorbed concentrations.[11] Alternatively, solid supported lipid membranes (SSLM), such as the commercially available 96 well plate TRANSIL assay, deploy phospholipid bilayers that are non-covalently surrounding macroporous silica, which allow for a clean supernatant by mild centrifugation.[22] However, systematic SSLM studies measuring the D_{MW} for ionizable compounds are scarce.[23] Chromatographic measurements with phospholipid coated column packing (IAM-HPLC) have recently been systematically evaluated for organic cations,[24,25] and considered a valuable tool as long as pH and salinity of the eluent conditions are carefully considered. It appears that measurements with ionogenic compounds on IAM-HPLC columns may suffer from confounding pH-dependent surface charges,[21,24] with the IAM material appearing net neutral only at pH $\sim$ 5, positively charged at a lower pH, and negatively charged at a higher pH, including in physiological buffers.

Also several predictive tools to derive D_{MW} values for ionogenic compounds have become available, which are important in order to screen the properties of large chemical data bases in chemical risk assessment or drug discovery processes. A review of liposomal K_{MW} data of ionic and corresponding neutral species reported in publicly available literature listed data for 56 acids and 36 bases.[16,19] Based on this data set, several scaling factors (Δ_{MW}) have been derived for different classes of ionic compounds (*e.g.* phenolic acids, secondary amines) that represent average differences between the log K_{MW} of ionic and corresponding neutral species:[26,27]

$$\log K_{MW,ion} = \log K_{MW,neutral} - \Delta_{MW} \quad (1)$$

where the log K_{MW} for neutral species is predicted using a log K_{OW} based-regression:[11]

$$\log K_{MW,ion} = (1.01 \log K_{OW,neutral} + 0.12) - \Delta_{MW} \quad (2)$$

The pH-dependent membrane–water distribution coefficient (D_{MW}) for ionisable compounds is:

$$D_{MW} = f_N K_{MW,neutral} + (f_N - 1) K_{MW,ion} \quad (3)$$

where f_N is the fraction of chemical in the neutral form, which at a given pH is:

$$f_N = (1 + 10^{(pK_a - pH)})^{-1} \text{ for bases, and}$$
$$f_N = (1 + 10^{(pH - pK_a)})^{-1} \text{ for acids} \quad (4)$$

Eqn (1)–(4) improve risk assessment of IOCs compared to a solely K_{OW} based approach, because they aim to include the actual contribution of the ionic species in the calculation of partition coefficients. However, the limited training data set showed considerable ranges in Δ_{MW} for each of the five types of ionogenic organic compounds (IOCs) considered, *i.e.* primary, secondary, and tertiary amines, phenolic acids, and other acids. Refining the Δ_{MW} values to more detailed structural features requires substantially larger data sets. In addition to such log K_{OW} derived predictions of D_{MW}, the computational tool COSMOmic became available to directly predict the K_{MW} for ionic species in 2014, using calculations of the molecular interactions between quantum-chemically optimized input structures.[16] A liposome-water partitioning data set of 75 ionisable compounds was used to fit an internal membrane potential to align the simulated K_{MW} data with experimental K_{MW} data for anions, cations and neutral compounds.[16] The direct prediction of the cationic partitioning coefficient by COSMOmic could replace the Δ_{MW}-approach based on log K_{OW} input values in eqn (1)–(4). Besides reducing the uncertainty related to the algorithms used to predict K_{OW}, the K_{OW}–K_{MW} regression, and the specificity of Δ_{MW} for certain types of IOCs, COSMOmic has

several advantages that render the predictions more mechanistically sound. COSMOmic simulations run with detailed 3-dimensional structural information for the most relevant conformers of the solute, taking electrostatic interactions into account, using quantum-chemistry-based calculations of the surface charge distribution.[28–30] Hence the influences that neighbouring charged and polar functional groups exert are accounted for. However, COSMOmic also has drawbacks in that the several tuneable parameters and conformer input structures can influence the output of calculations, and as yet there is no clear consensus on how to control this variability. Also, the COSMOtherm software package in which COSMOmic is a sub-module is a commercial product, while QSAR based algorithms like K_{OW} and the Δ_{MW}-approach can be readily embedded in freely available webtools (*e.g.* QSAR-toolbox).

In the current study, we used IAM-HPLC to obtain experimental sorption data for phospholipid monolayer coatings with 86 charged amine structures encompassing a wide variety of functional polar groups. Two preceding studies have demonstrated that the K_{MW} of organic cations can also be approximated with IAM-HPLC chromatographic retention data.[24,25] The IAM-HPLC data were obtained at pH 5 for seventy different amine structures that contained only charged nitrogen and hydrocarbon moieties ("$C_xH_yN^+$"), and COSMOmic predictions compared well with these data. Furthermore, for a broad series of cationic surfactants (also $C_xH_yN^+$), the IAM-HPLC data corresponded closely with sorption data obtained with TRANSIL solid supported lipid membranes (SSLM), in which non-covalently bonded phospholipid bilayers are used.[23] The main differences between IAM-HPLC and SSLM for these $C_xH_yN^+$ cations were linked to contributions from the methyl groups attached to the charged nitrogen ("*N*-methylation"). Hence, corrective increments for *N*-methylation were derived to align the IAM-HPLC based values with the SSLM D_{MW} measurements ($\delta_{IAM\text{-}SSLM}$). These $\delta_{IAM\text{-}SSLM}$ values are expected to be applicable to all charged amines, which is part of the current investigation. Interestingly, COSMOmic predictions of the contribution of *N*-methylation to D_{MW} followed a similar trend to the SSLM data with the effect only slightly exaggerated for primary amines.

Together with the IAM-HPLC data set on $C_xH_yN^+$ amines, the current IAM-HPLC data set on 86 polar amines will be used to evaluate the Δ_{MW}-approach, and to refine the chemical application domain of COSMOmic as a predictive tool in the screening of large sets of ionisable chemicals in lower tier risk assessment.

2. Materials and methods

2.1 Test compounds

Eighty-six amine structures were purchased as the free base or the salt from various suppliers (see ESI Table S1†). Each contained a single ionisable amine moiety, and was grouped and numbered according to the polarity of the main polar groups and complexity of the structures, as listed in Table 1 and with structures in Table 2: monopolar amines (simple esters, ethers, halogenated amines), dipolar amines (simple hydroxylated amines), amines with polar N moieties (anilines, amides, carbamate, N-heterocyclic, guanidines, sulfonamides), polycyclic amine structures, and complex multifunctional amine structures (≥4 polar groups, and/or M_w > 400, and/or larger rings than C_7). All compounds had a purity of >95% except for two macrolides (#85; erythromycin A (<12% erythromycin B), and #86 tylosin hemitartrate (91.4%)). Stock solutions were prepared as ~10 mg in 1 mL methanol (Biosolve BV, Valkenswaard, NL), in some cases in acidified ethanol. If experimental pK_a values and K_{OW} values were not available, ChemAxon (http://www.chemicalize.org, April 2016) was used to predict these parameters (Table 1). Nine of the chemicals were quaternary ammonium compounds (*i.e.* the amine group always has a charge); seventy amines were strong bases with pK_a values >7.0; four had pK_a values between 5.0 and 7.0, and three had pK_a values between 3.0 and 5.0.

2.2 IAM column, eluents, detection and analysis

The same IAM-HPLC settings were used as reported previously for the $C_xH_yN^+$ amines.[25] Briefly, a 100 × 4.6 mm IAM.PC.DD2 column (Regis Technologies, Inc., Morton Grove, IL, USA) with an IAM.PC.DD2 10/300 guard cartridge was operated at a flow rate of 1.0 mL min^{-1} (23 ± 2 °C) of pH 5.0 or 3.0 buffered aqueous eluent (without organic solvent), with an additional 8.0 g L^{-1} NaCl (137 mM) and 0.2 g L^{-1} KCl (2.7 mM) added to the 10 mM buffer. Amines with aromatic moieties were mostly detected by using a UV-diode array (Agilent 1100 system) at various wavelengths. Amines that lacked a convenient UV absorption profile were detected using LC-MS/MS (MDS Sciex API 3000) in the 10 mM buffer alone, with a split flow injection of <5%. Triplicate IAM-injections (20 μL) were run on the same day for most compounds. A series of at least 3 different eluent mixtures with aqueous buffer and acetonitrile (≤30%) were applied in the case of strongly sorbing amines and extrapolated to obtain the retention capacity factor (K_{IAM}) for buffered aqueous eluent without organic solvent.

3-Nitroaniline was the neutral reference compound used to check IAM consistency during UV-diode array detection. Pure water (MilliQ, Millipore Merck) was used as a neutral non-retained tracer (t_0) when using UV-diode array detection. Tryptamine was used as a reference cation to align retention on UV-diode array and LC-MS/MS systems. The peak apex of the eluted peak (t_r) on both detectors was used to calculate the retention capacity factors (k_{IAM}):

$$k_{IAM} = t_r/(t_r - t_0) \quad (5)$$

The intrinsic sorption coefficient of organic cations to the IAM monolayer ($K_{IAM,intr}$) at pH 5.0 was obtained by multiplying k_{IAM} by the solvent/sorbent phase ratio of 18.9 for the IAM.PC.DD2 column:[25,31]

$$k_{IAM,intr}\ (pH\ 5) = k_{IAM} \times 18.9 \quad (6)$$

For weaker bases with pK_a of 3.5–7, measurements were performed at pH 3.0, 5.0 and 7.4 (10 mM phosphate buffer), all at a total salinity of 0.15 M. The pH 3.0 values were corrected with the $\Delta_{pH5-pH3}$ electrostatic factor[25,31] (additionally confirmed in this study):

$$\log K_{IAM,intr}\ (pH\ 5) = \log(k_{IAM}\ (pH\ 3) \times 18.9) + \Delta_{pH5-pH3} \quad (7)$$

Table 1 Properties of amines (A) (1–60), (B) (61–86), IAM-HPLC phospholipid affinity of cations corrected with $\delta_{IAM\text{-}SSLM}(\log(K_{IAM,intr}) = \log(18.9 \times k_{IAM}) + \delta_{IAM\text{-}SSLM})$, COSMOmic calculated affinity (K_{DMPC-W}-TUHH, incl. potential, no offset)

A (1–60)

#	Name	pK_a[a]	IOC type	Molecular formula $C_xH_yN^+$ + rest	Rest moieties[b]	Δ_{pH5-3}[c]	log K_{MW}(IAM)[d]	log K_{DMPC-W} TUHH[e]	log K_{OW}[f]
Halogenated/nonpolar amines									
1	Deprenyl	8.67	3°	$C_{13}H_{17}N$	C≡C		1.97	1.48	2.90*
2	4-Fluoroaniline	4.65*	1°	C_6H_6N–F	F		1.05	2.01	1.15*
3	3-Chloroaniline	3.52*	1°	C_6H_6N–Cl	Cl		1.97	2.34	1.88*
4	4-Chlorobenzylamine	9.44	1°	C_7H_8N–Cl	Cl	0.06	1.66	2.53	1.70
5	3,4-Dichlorobenzylamine	9.34	1°	C_7H_7N–Cl_2	Cl × 2	0.07	2.24	3.06	2.31
6	Sertraline	9.85	2°	$C_{17}H_{17}N$–Cl_2	Cl × 2		4.32	3.47	5.15
Simple monopolar amines									
7	Diphenhydramine	8.98*	3°	$C_{17}H_{21}N$–O	COC	0.12	2.88	3.28	3.27*
8	Tamoxifen	8.76	3°	$C_{26}H_{29}N$–O	COCar		6.51	5.29	6.30
9	Fluoxetine	9.8	2°	$C_{17}H_{18}N$–F_3O	COCar, CF_3	0.12	4.28	4.10	4.05*
10	Duloxetine	9.7	2°	$C_{18}H_{19}N$–OS	COCar/thiophen/naph		3.93	3.93	4.20
11	Escitalopram	9.78	3°	$C_{20}H_{21}N$–FNO	COC/≡N/F		3.28	1.53	3.74
12	*N*-Benzylaminoacetal dehydediethylacetal	8.41	2°	$C_{13}H_{22}N$–O_2	COC × 2	0.13	1.85	2.16	2.48
13	Spiroxamine	9.34	3°	$C_{18}H_{35}N$–O_2	COC × 2		3.30	3.23	4.38
14	MDMA	10.14	2°	$C_{11}H_{15}N$–O_2	COC × 2(dioxalane)	0.10	1.74	1.63	2.15*
15	3-Dimethylamino-propiophenone	8.29	3°	$C_{11}H_{16}N$–O	=O	0.14	1.33	0.83	1.66
16	Bupropion	8.22	2°	$C_{13}H_{18}N$–ClO	=O/Cl		2.42	2.00	3.27
17	Ketamine	7.5*	2°	$C_{13}H_{16}N$–ClO	=O/Cl	0.11	1.88	1.23	2.18*
18	Ethyl-3-(benzylamino) propanoate	9.04	2°	$C_{11}H_{16}N$–O_2	C(=O)OC	0.12	1.27	1.60	1.75
19	Butyrylcholine	Q	4°	$C_9NH_{20}N$–O_2	C(=O)OC	0.11	0.66	0.41	Q
20	Butyrylthiocholine	Q	4°	$C_9NH_{20}N$–OS	C(=O)SC		1.13	1.23	Q
21	Valethamate	Q	4°	$C_{19}H_{32}N$–O_2	C(=O)OC	−0.01	3.06	2.65	Q
22	Cocaine	8.61*	3°	$C_{17}H_{21}N$–O_4	C(=O)OC × 2	0.09	2.10	1.77	2.30*
Simple dipolar amines									
23	*N*-Ethanolbenzylamine	9.18	2°	$C_9H_{14}N$–O	OH	0.13	0.97	1.42	0.79*
24	Methylephedrine	8.86	3°	$C_{11}H_{17}N$–O	OH	0.12	1.35	1.44	1.70*
25	Benzyldimethyl-hydroxyethyl-ammonium	Q	4°	$C_{11}H_{18}N$–O	OH	0.17	1.00	0.90	Q
26	2,2′-(Benzylimino)-diethanol	8.5	3°	$C_{11}H_{17}N$–O_2	OH × 2	0.11	1.01	0.99	0.53
27	Dopamine	8.93*	2°	$C_8H_{11}N$–O_2	OH × 2		0.96	2.64	−0.98*
28	Albuterol	9.4	2°	$C_{13}H_{21}N$–O3	OH × 3		1.23	1.83	0.34
29	Tramadol	9.23	3°	$C_{16}H_{25}N$–O_2	COCar/OH	0.09	2.01	1.76	3.01*
30	Alprenolol	9.67	2°	$C_{15}H_{23}N$–O_2	COCar/OH	0.05	2.79	3.00	3.10*
31	Nadolol	9.67*	2°	$C_{17}H_{27}N$–O_4	COCar/OH × 3		1.88	1.74	0.81*

Table 1 (*Contd.*)

A (1–60)

#	Name	pK_a[a]	IOC type	Molecular formula $C_xH_yN^+$ + rest	Rest moieties[b]	Δ_{pH5-3}[c]	log K_{MW}(IAM)[d]	log K_{DMPC-W} TUHH[e]	log K_{OW}[f]
32	Metoprolol	9.67	2°	$C_{15}H_{25}N–O_3$	COCar/COC/OH	0.11	2.06	2.57	1.88*
33	Propranolol	9.42*	2°	$C_{16}H_{21}N–O_2$	COCar/OH/naph	0.07	3.12	3.11	3.48*
34	Atropine	9.43*	3°	$C_{17}H_{23}N–O_3$	C(=O)OC/OH	0.07	1.95	1.14	1.83*
35	Scopolamine	7.75*	3°	$C_{17}H_{21}N–O_4$	C(=O)OC/OH/COC	0.11	1.42	1.39	0.98*
Simple amines with polar N groups									
36	Procaine	8.05*	3°	$C_{13}H_{20}N–NO_2$	Aniline/C(=O)OC	0.21	1.57	0.99	2.14*
37	4-Amino-2-methylquinoline	8.81	3°	$C_{10}H_{10}N–N$	Aniline/naph	0.16	2.04	0.67	1.43
38	Ropivacaine	7.82	3°	$C_{17}H_{26}N–NO$	C(=O)NC	0.17	2.15	1.74	2.90*
39	Bupivacaine	8.01*	3°	$C_{18}H_{28}N–NO$	C(=O)NC	0.17	2.45	2.14	3.41*
40	Lidocaine	8.01*	3°	$C_{14}H_{22}N–NO$	C(=O)NC	0.16	1.66	1.58	2.44*
41	Prilocaine	8.82	2°	$C_{13}H_{20}N–NO$	C(=O)NC	0.12	1.78	2.34	2.11*
42	1-Benzyl-3-acetamidopyrrolidine	8.38	3°	$C_{13}H_{19}N–NO$	C(=O)NC	0.16	1.27	0.23	0.88
43	Propamocarb	9.5*	3°	$C_9H_{20}N–NO_2$	Carbamate		0.95	0.70	1.12*
44	Neostigmine	Q	4°	$C_{12}H_{19}N–NO_2$	Carbamate	0.19	1.19	0.45	Q
45	Pyridostigmine	Q	4°	$C_9H_{13}N–NO_2$	Carbamate	0.26	0.59	−0.07	Q
46	4-Carbamoyl-1-pentylpyridinium	Q	4°	$C_{11}H_{17}N–NO$	Acetamide	0.11	1.12	0.75	Q
47	4-Carbamoyl-1-octylpyridinium	Q	4°	$C_{14}H_{23}N–NO$	Acetamide	0.09	2.52	2.40	Q
48	Atenolol	9.6*	2°	$C_{14}H_{22}N–NO_3$	Acetamide/COC/OH	0.12	1.26	0.70	0.16*
49	Tryptamine	9.73	1°	$C_{10}H_{12}N–N$	Indole	0.08	1.91	2.50	1.55*
50	Serotonin	9.97*	1°	$C_{10}H_{12}N–NO$	Indole/OH	0.10	1.59	2.42	0.21*
51	Benzimidazole	5.3*	3°	$C_{13}H_{11}N–N$	Imidazole		1.22	0.96	1.32*
52	Thiabendazole	4.64*	3°	$C_{10}H_7N–N_2S$	Imidazole/thiazol		2.28	0.94	2.47*
53	Pindolol	9.25*	2°	$C_{14}H_{20}N–NO_2$	Imidazole/COC/OH		2.11	2.11	1.75*
54	Nicotine	8.86	3°	$C_{10}H_{14}N–N$	Pyridine		0.96	−0.26	1.17*
55	Clonidine	8.05*	3°	$C_9H_9N–N_2Cl_2$	Imidazole/Cl × 2	0.19	1.71	1.36	1.59*
56	Clotrimazole	6.62	3°	$C_{22}H_{17}N–NCl$	Imidazole//Cl	0.12	4.10	3.33	6.26*
57	Timolol	9.21*	2°	$C_{12}H_{22}N–N_3O_3S$	Thiadiazol/2COC/OH		1.98	1.89	1.83*
58	1,3-Diphenylguanidine	10.1*	2°	$C_{13}H_{13}N–N_2$	Guanidine	0.13	2.05	2.17	3.13
59	Sotalol	9.43	2°	$C_{12}H_{20}N–NO_3S$	Sulfon/OH		1.33	1.22	0.24*
60	Sulpiride	9.12*	3°	$C_{15}H_{23}N–N_2O_4S$	Sulfon/C(=O)NC/COC		1.64	0.26	0.57*

Table 1 (*Contd.*)

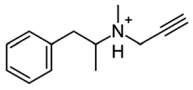
B (61–86)

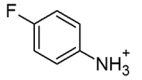
#	Name	pK_a[a]	IOC type	Molecular formula $C_xH_yN^+$ + rest	Rest moieties[b]	Δ_{pH5-3}[c]	log K_{MW}(IAM)[d]	log K_{DMPC-W} TUHH[e]	log K_{OW}[f]
Polyaromatic amines									
61	Morphine	8.21*	3°	$C_{17}H_{19}N–O_3$	Polyarom/OH × 2/COCar	0.14	1.29	0.69	0.89*
62	Naloxone	7.0*	3°	$C_{19}H_{21}N–O_4$	Polyarom/OH × 2/=O/COCar		1.63	1.22	2.09*
63	Strychnine	8.26*	3°	$C_{21}H_{22}N–NO_2$	Polyarom/C(=O)NC/COC		2.28	0.51	1.93*
64	Harmine	6.15	3°	$C_{13}H_{12}N–NO$	Azaarene/COCar/cyclic N	0.10	3.08	1.48	3.56*
65	Harmaline	7.96	3°	$C_{13}H_{14}N–NO$	Azaarene/COCar/cyclic N		2.78	1.21	1.67
66	Promazine	9.36*	3°	$C_{17}H_{20}N–NS$	Tricyclic-N + S	0.12	3.87	2.77	4.55*
67	Chlorpromazine	9.3*	3°	$C_{17}H_{19}N–Cl_2S$	Tricyclic-N + S/Cl	0.12	4.34	3.21	5.41*
68	Triflupromazine	9.2	3°	$C_{18}H_{19}N–NF_3S$	Tricyclic-N + S/CF_3	0.12	4.58	3.45	5.54*
69	Doxepin	9.76	3°	$C_{19}H_{21}N–O$	Tricyclic-O		3.31	2.73	4.36*
70	Difenzoquat	Q	4°	$C_{17}H_{17}N–N$	Tricyclic-N	0.13	2.56	1.67	Q
Complex amines (≥4 polar groups, and/or M_w > 400, and/or large rings)									
71	Verapamil	8.92*	3°	$C_{27}H_{38}N–NO_4$	COCar × 4/≡N	0.12	4.08	1.08	3.79*
72	Trimethoprim	7.12*	3°	$C_{14}H_{18}N–N_3O_3$	COCar × 3/pyrimidine-2,4-diamine	0.14	2.14	1.51	0.91*
73	Loperamide	9.41	3°	$C_{29}H_{33}N–NClO_2$	C(=O)NC/OH/Cl	0.12	4.44	4.23	4.77
74	Ketanserin	7.29	3°	$C_{22}H_{22}N–N_2FO_3$	C(=O)NC × 2/=O/F	0.05	3.19	1.92	3.29*
75	Amlodipine	9.45	1°	$C_{20}H_{25}N–NClO_5$	Pyridine/C(=O)OC × 2/COC/Cl/phenyl		4.15	4.44	3.00*
76	Nicardipine	8.18	3°	$C_{26}H_{27}N–N_2O_6$	Nitro/C(=O)OC × 2		3.89	3.00	3.82*
77	Acebutolol	9.57	2°	$C_{18}H_{26}N–NO_4$	C(=O)NC/=O/COC/OH		2.40	2.62	1.71*
78	Lincomycin	7.8*	3°	$C_{18}H_{34}N–NO_6S$	C(=O)NC/OH × 4/COC/CSC	0.09	1.28	2.06	0.20*
79	Diltiazem	8.18	3°	$C_{22}H_{26}N–NO_4S$	C(=O)NC/C(=O)OC/COCar/S		3.26	2.24	2.70*
80	Amiodarone	6.56*	3°	$C_{25}H_{29}N–I_2O_3$	Furan/=O/COCar/I × 2		6.50	5.68	7.57*
81	Ergocornine	7.78*	3°	$C_{31}H_{39}N–N_4O_5$	C(=O)NC × 3/COC/OH/imidazole		4.30	2.03	2.64
82	Nefopam	8.01	3°	$C_{17}H_{19}N–O$	C_8 ring/COC		2.52	1.91	3.05*
83	Tiamulin	9.51	3°	$C_{28}H_{47}N–O_4S$	C_8 ring/C(=O)OC/=O/OH/S	0.12	0.99	2.01	4.50
84	Clarithromycin	8.99*	3°	$C_{38}H_{69}N–O_{13}$	C_{14} ring/C(=O)OC/=O/COC × 6/OH × 4		3.70	3.30	3.16*
85	Erythromycin	8.88*	3°	$C_{37}H_{67}N–O_{13}$	C_{14} ring/C(=O)OC/=O/COC × 5/OH × 5		3.38	3.86	3.06*
86	Tylosine	7.73*	3°	$C_{46}H_{77}N–O_{17}$	C_{16} ring/C(=O)OC/=O × 2/COC × 8/OH × 5		3.40	2.48	1.63*

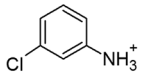
[a] pK_a values are recommended experimental data (*) from SRC/EPISuite/Pubchem, or predictions by ChemAxon (March 2016). [b] COC = ether, COCar = ether next to the aromatic ring, =O = ketone, C(=O)O = ester, C(=O)NC = amide, sulfon = sulfonamide. Full 2D molecular structures are in Table 2 of the manuscript. [c] For 51 amines measurements of log K_{IAM} were made in both pH 5.0 and pH 3.0 eluents, Δ_{pH5-3} expresses the difference (average 0.12 ± 0.04). [d] log K_{MW}(IAM) values are obtained with IAM-HPLC at pH 5 for bases with pK_a > 7, or pH 3 for weaker bases (including Δ_{pH5-3}), and all are corrected for *N*-methylation contribution corrections *via* $\delta_{IAM\text{-}SSLM}$. [e] log K_{DMPC-W} values were calculated with the TUHH-system files and membrane potential settings, no offset or structural correction factors included. [f] log K_{OW} values are recommended experimental data (*) from SRC/EPISuite/Pubchem, or predictions by EPISuite's KOWwin. Q represents quaternary ammonium salts, for which no log K_{OW} can be defined.

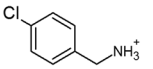
Table 2 Chemical structure of tested organic cations

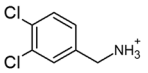
Nonpolar/halogenated amines

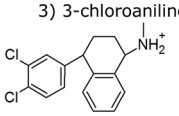

1) deprenyl


2) 4-fluoroaniline

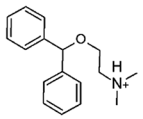

3) 3-chloroaniline

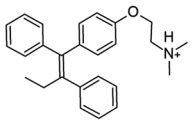

4) 4-chlorobenzylamine

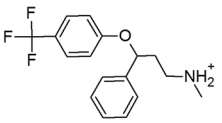

5) 3,4-dichlorobenzylamine

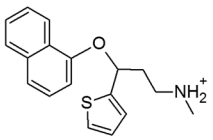

6) sertraline

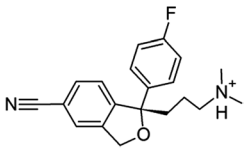
Simple monopolar amines

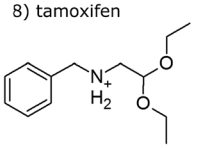

7) diphenhydramine

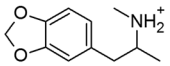

8) tamoxifen

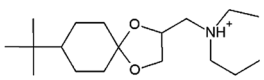

9) fluoxetine

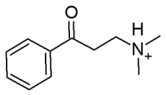

10) duloxetine

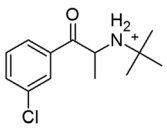

11) escitalopram

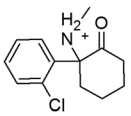

12) N-Benzylaminoacet-aldehydediethylacetal

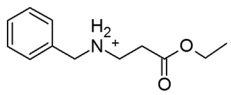

14) MDMA

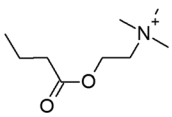

13) spiroxamine

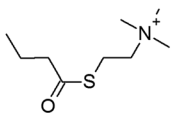

15) 3-dimethylamino-propiophenone

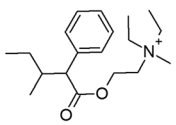

16) bupropion

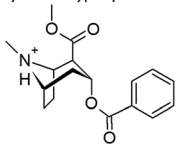

17) ketamine


18) ethyl 3-(benzylamino)propanoate

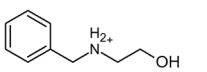

19) butyrylcholine

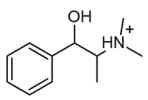

20) butyrylthiocholine

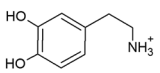

21) valethamate

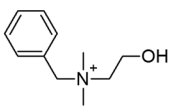

22) cocaine

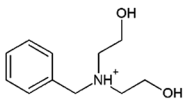
Simple dipolar amines

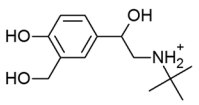

23) N-ethanolbenzylamine

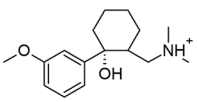

24) methylephedrine

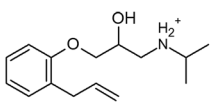

25) dopamine

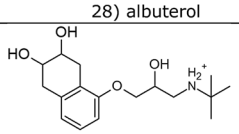

26) benzyldimethylhydroxyethylamm

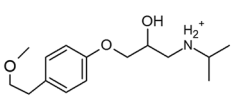

27) 2,2'-(benzylimino)-diethanol

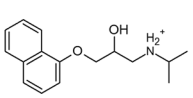

28) albuterol

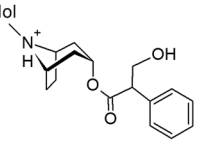

29) tramadol


30) alprenolol


31) nadolol


32) metoprolol


33) propranolol


34) atropine

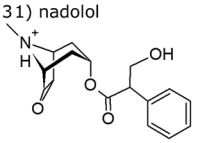
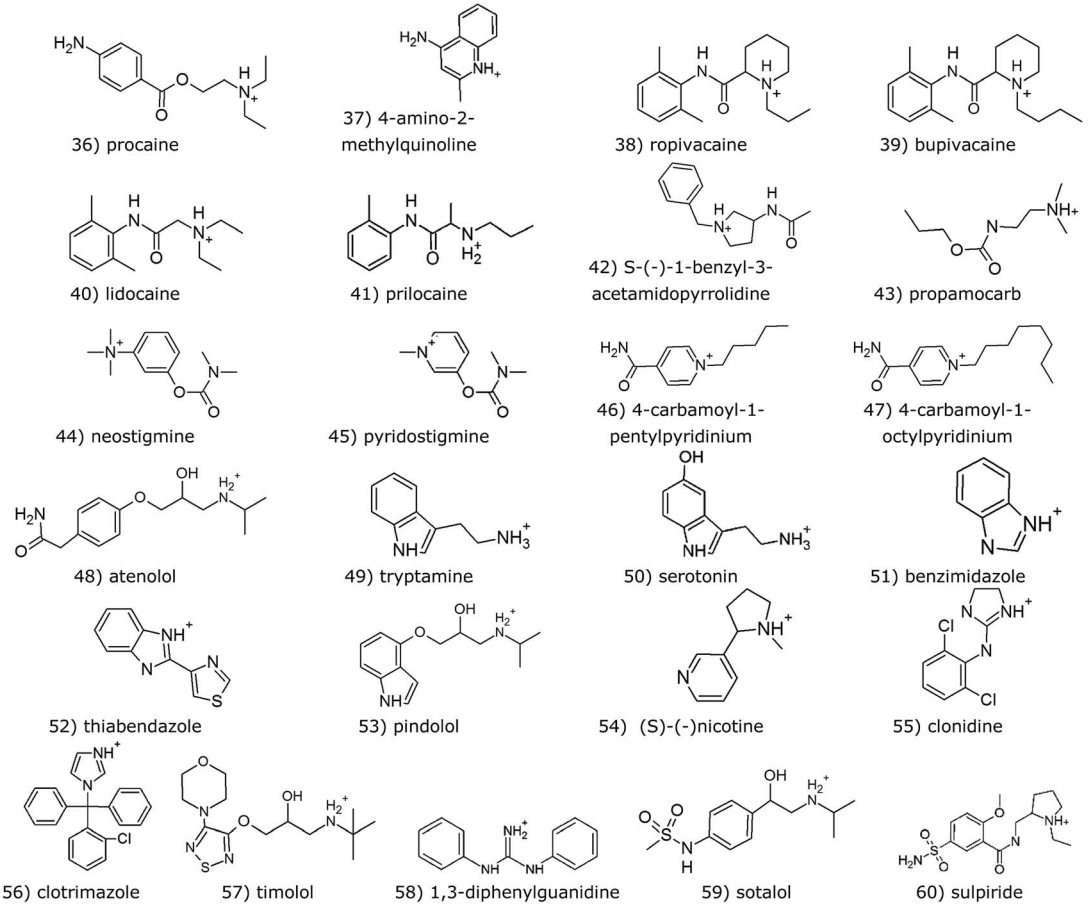

35) scopolamine

Table 2 (*Contd.*)

Simple amines with polar N groups



Polyaromatic amines

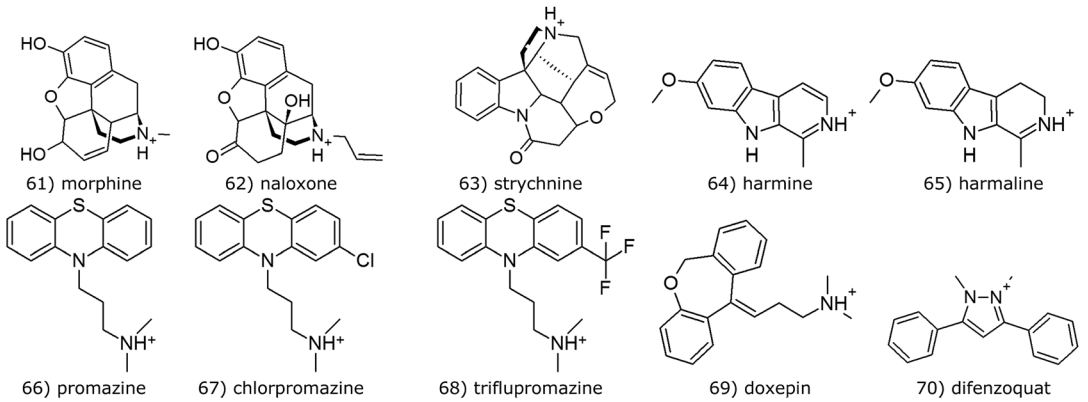
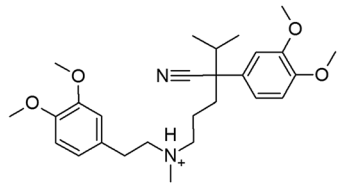
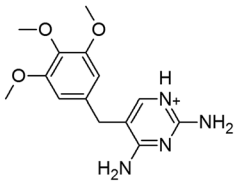
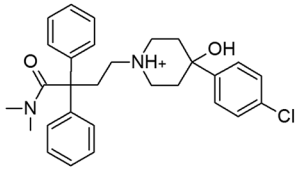
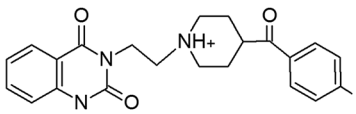
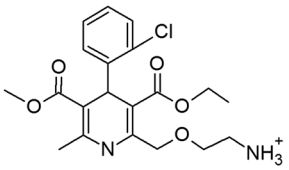
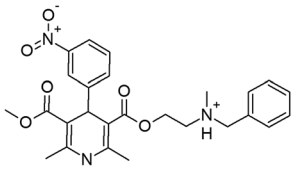
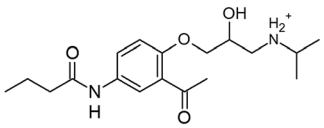
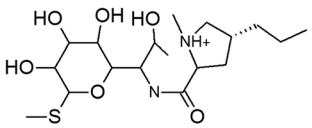
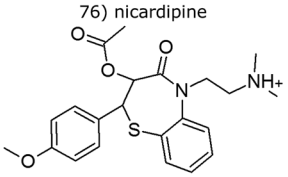
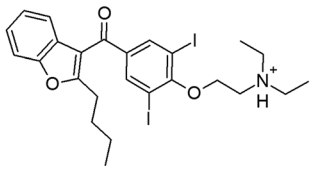
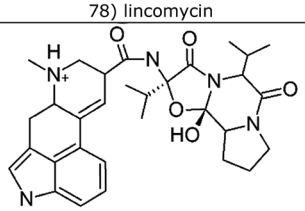
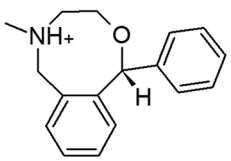
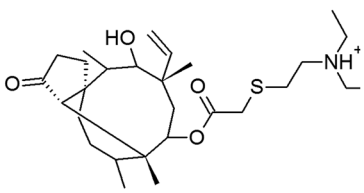
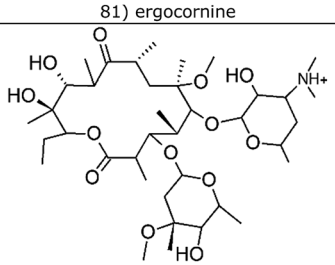
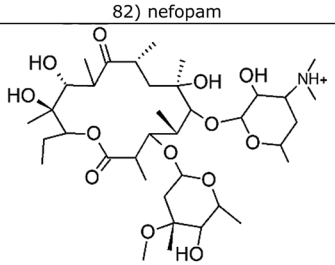
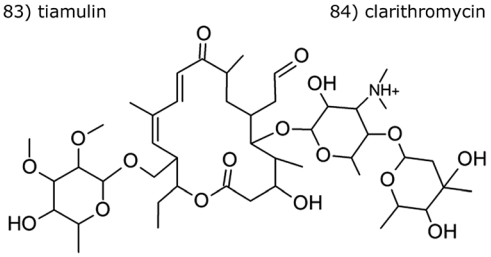


Table 2 (*Contd.*)

Complex amines
 71) verapamil
 72) trimethoprim
 73) loperamide
 74) ketanserin
 75) amlodipine
 76) nicardipine
 77) acebutolol
 78) lincomycin
 79) diltiazem
 80) amiodarone
 81) ergocornine
 82) nefopam
 83) tiamulin
 84) clarithromycin
 85) erythromycin
 86) tylosine

This pH 3 value being representative of the cationic species was confirmed by Henderson–Hasselbalch pH model[18] fitting of the measurements at three tested pHs.

To convert the intrinsic sorption affinity to the IAM monolayer to the membrane–water sorption affinity for a phospholipid bilayer (K_{MW}), we applied corrective increments ($\delta_{IAM\text{-}SSLM}$) to all $K_{IAM,intr}$ for the differential *N*-methylation contributions to K_{IAM} and K_{MW}, as defined in ref. 23, with $\delta_{IAM\text{-}SSLM}$ of +0.78 log units for 1° amines, +0.47 log units for 2° amines, −0.03 log units for 3° amines, and −0.11 log units for 4° ammonium compounds:

$$\log K_{MW}(IAM) = \log K_{IAM,intr}\ (pH\ 5) + \delta_{IAM\text{-}SSLM} \quad (8)$$

2.3 Molecular simulations

COSMOmic was run within COSMOthermX Version C30_1501, using the same time-averaged atomic distribution micelle file for a hydrated system of 1,2-dimyristoyl-*sn-glycero*-3-phosphocholine (DMPC), and a TZVP-optimized representative DMPC phospholipid molecule, optimized at the Hamburg University of Technology (TUHH).[32] The same DMPC files used in the study of Bittermann *et al.*[16] that fitted the COSMOmic default "self-defined" electrostatic potential settings were used (position 17.0796, width 8.86629, depth 7.51583 kcal (326 mV)). Bittermann *et al.*[16] observed that the COSMOmic calculated $K_{DMPC\text{–}W}$ values overestimated the sorption affinity to liposomes for all types of compounds (charged and neutral), and recommended to apply generic offset corrections of −0.32 log units to the calculated $K_{DMPC\text{–}W}$ values with their system files and membrane potential settings. In more detail, they reported −0.30 log units for neutral compounds and −0.40 log units for organic cations. In this study we evaluated COSMOmic $K_{DMPC\text{–}W}$ values with and without an offset for organic cations of −0.4 log units. The example bilayer dmpc.mic provided with COSMOmic V15 ("COSMOtherm15-DMPC"), which we applied in our previous IAM-HPLC/COSMOmic comparison study with $C_xH_yN^+$ amines,[25] uses a slightly different set of input structures, and is here compared to the "TUHH-DMPC" bilayer files. COSMOmic divides one half of the hydrated DMPC bilayer (composed of a "time-average" atomic distribution obtained *via* molecular dynamics equilibration) into 31 horizontal layers along the axis from the bulk water to the membrane core, and calculates at each layer the free energy (relative to that in bulk water) for each chemical input structure in 162 orientations. COSMOconf Version 3.0.0.0 was used to create TZVP based input structures for up to 6 of the most relevant conformers for each of the charged amines, as discussed in the $C_xH_yN^+$ amine study.[25]

3. Results and discussion

3.1 K_{MW} values obtained by IAM-HPLC: extrapolating solvent mixture series, pH corrections, amine type corrections

IAM-HPLC retention capacity factors were obtained in buffered aqueous eluents ($K_{IAM,aq}$) for 62 amines. For 24 strongly sorbing amines varying mixtures of acetonitrile/water were used to extrapolate the $K_{IAM,apparent}$ to $K_{IAM,aq}$ (see ESI Fig. S1,† examples of linear extrapolation curves to $K_{IAM,aq}$ shown for amines #72 and #74). The triplicate $K_{IAM,aq}$ measurements in fully aqueous eluent differed by less than 0.1 log units. The extrapolated $K_{IAM,aq}$ values in the solvent have higher uncertainty margins; for 18 compounds more than 4 measurements were established in the solvent series, resulting in average 95% c.i. for $K_{IAM,aq}$ of 0.25 log units. For 6 compounds only four single measurements were performed in 15%, 20%, 25%, and 30% solvent, and corresponding 95% confidence intervals (c.i.) ranged over 0.22–0.91 log units (see Fig. S1†). The $K_{IAM,aq}$ values for pH 5.0 aqueous eluent were transformed to $K_{IAM,intr}$ values (listed in Table S2†). The tested range of log $K_{IAM,intr}$ values obtained in this study was 0.59 (#45, pyridostigmine) to 6.50 (#80, amiodarone). For 51 amines, measurements were made in both pH 5.0 and pH 3.0 eluents (both with physiological salinity of 0.15 M). The difference between the apparent log K_{IAM} in both media (Δ_{pH5-3}) was on average 0.12 ± 0.04 (see Table 1), which corresponds to the recently reported value for $C_xH_yN^+$ amines.[25] As demonstrated in recent systematic studies on the effect of salinity on the retention of anions and cations,[24] the IAM material is net positively charged at pH 3. Presumably, this is because most residual silanol acidic sites at the base of the IAM silica surface are protonated at pH 3, while a larger number of residual positively charged propylamines exist as well as the base of the IAM coating as unreacted groups during the anchoring of the phospholipids. At pH 5 the amount of deprotonated silanol groups balances the amount of the residual propylamine groups, resulting in net neutral surface charge and negligible effects of salinity on retention of ionic species. The Δ_{pH5-3} correction factor is thus due to moderate electrostatic repulsion of solute cations at highly saline pH 3 buffered eluents, and a negligible electrostatic repulsion/attraction at pH 5.[24] The Δ_{pH5-3} correction factor appears to be applicable to all monoprotic organic cations, but may be different for multiprotic compounds.

The IAM-HPLC study on charged $C_xH_yN^+$ amines showed that for different amine type analogues, the influence of methyl groups attached to the charged N (*N*-methylation) on $K_{IAM,intr}$ was minimal for 1–3° amines while the 4° ammonium analogues had $K_{PLIP\text{–}W}$ values lower by ~0.3 log units.[25] Timmer and Droge[23] recently demonstrated that the influence of *N*-methylation for charged $C_xH_yN^+$ amines was more apparent for phospholipid bilayers in solid supported lipid membrane (SSLM) assays than for IAM-HPLC. For a series of alkylamine analogues, with C_8-, C_{10}- and C_{12}-chains, they found that the order of K_{MW} values for the analogue series was 1° > 2° > 3° > 4°, with respective incremental differences of 0.28, 0.56 and 0.53 log units.[23] Protonated primary alkylamines thus sorbed ~1.3 log units stronger to phospholipid bilayers than analogue quaternary alkylammonium cations. We considered that this is solely an effect of the charged nitrogen moiety that will apply to all charged amines, either $C_xH_yN^+$ amines or polar amines, just like the Δ_{pH5-3} correction factor. By applying $\delta_{IAM\text{-}SSLM}$ corrective increments to our IAM-HPLC data for organic cations, *i.e.* +0.78 for 1° amines, +0.47 for 2° amines, −0.03 for 3° amines and −0.11 for 4° ammonium compounds, we corrected for the different *N*-methylation contributions in IAM monolayers and SSLM bilayers. Doing so, we suggest to have derived IAM-based

sorption affinities to phospholipid bilayers (K_{MW}(IAM)), as the closest proxy values of sorption affinities to cell membranes. These $\delta_{IAM\text{-}SSLM}$ corrected K_{MW}(IAM) data are presented in Table 1 and used in the following evaluations, unless specified otherwise.

3.2 Using IAM-HPLC and K_{OW} to define Δ_{MW} scaling factors

The bioaccumulation model for ionogenic compounds "BIONIC" currently applies the Δ_{MW}-approach in eqn (2) to predict K_{MW} values for ionic species.[27] The Δ_{MW}-values represent the difference between log K_{MW} values for the neutral and corresponding ionic species, which have been defined for several types of IOCs. In the absence of experimental K_{MW} values for the neutral species, they are based on K_{OW}-QSAR. Using the IAM-HPLC data, we can now evaluate the Δ_{MW}-values for 1°, 2°, and 3° amines for the 133 amines (86 polar amines and 69 $C_xH_yN^+$ amines, minus 22 QACs) collected in our current and previous study.[25] Using only experimental K_{OW} values for 87 compounds (ESI-Tables S3A and B†), uncertainty related to predictions of K_{OW} is partially avoided while maintaining a large chemical diversity. Using the $\delta_{IAM\text{-}SSLM}$ corrected IAM-HPLC data as $K_{MW,ion}$ and eqn (2) to calculate the $K_{MW,N}$, Fig. 1 shows the Δ_{MW}-values separated out for the three amine types, with each amine type split into $C_xH_yN^+$ amines and polar amines. For primary (1°), secondary (2°), and tertiary (3°) $C_xH_yN^+$ amines, log unit Δ_{MW}-values (±s.d. (*N*)) are −0.26 ± 0.32 (12), 0.30 ± 0.32 (7), and 1.29 ± 0.35 (11), respectively.

Using liposomal partitioning data for both ionic and corresponding neutral species of a small set of basic compounds, Δ_{MW}-values have been set to 0.3, 0.5 and 1.25 log units, for 1°, 2° and 3° amines, respectively, and used as such in the BIONIC model. The SSLM sorption study with a small but consistent data set of single chain cationic surfactants determined respective Δ_{MW}-values of −0.05, 0.44, and 1.25 using the K_{OW}-QSAR for neutral species.[23] Especially for 2° and 3° amines, these Δ_{MW}-values were very close to those observed with IAM-HPLC data, and in a similar trend to the Δ_{MW}-values used in the BIONIC model.

However, IAM-HPLC data for ($C_xH_yN^+$) amines do not always closely agree with the Δ_{MW}-values used in the BIONIC model. For example, the Δ_{MW}-values for 2° $C_xH_yN^+$ dioctylamine are two orders of magnitude higher (2.55) compared to the recommended 0.5 log units. The IAM-HPLC data generated in the current study provided a much more extensive set of polar amines than previously available for liposomal and SSLM partitioning data. For polar amines, Fig. 1 shows that the Δ_{MW}-values are lower and more variable per amine type than for $C_xH_yN^+$ amines, and therefore also in comparison to the Δ_{MW}-values set in the BIONIC model, −1.4 ± 0.73 (6 compounds) for 1°, −0.57 ± 0.57 (14 compounds) for 2°, and 0.35 ± 0.78 (37 compounds) for 3°. Polar moieties appear to result, on average, in relatively higher K_{MW} values for ionic species compared to corresponding neutral species, resulting in lower Δ_{MW}-values compared to $C_xH_yN^+$ amines. It should be kept in mind that the liposomal Δ_{MW}-values are based on measured K_{MW} values for neutral species, whereas our assessment still applied the K_{OW}-QSAR. Of course, predictions of K_{OW} are integral in any large chemical risk assessment screening. Fig. 1 therefore also shows predictions of K_{OW} by EPISuite's KOWwin and ACD/Labs (listed in ESI-Tables 3A and B†). The same trends in Δ_{MW}-values per amine type were observed as when using experimental K_{OW} values. The scatter in Δ_{MW}-values per amine type was somewhat increased, likely by including more complex amine structures, although for most of the very complex amines (#71–86) experimental K_{OW} values were available (Table 1). The Δ_{MW}-value of 5.14 for tiamulin (#83) for the ACD labs predicted K_{OW} seems unrealistically high. For most simple nonpolar alkylamines this Δ_{MW}-approach based on K_{OW} includes uncertainty margins of

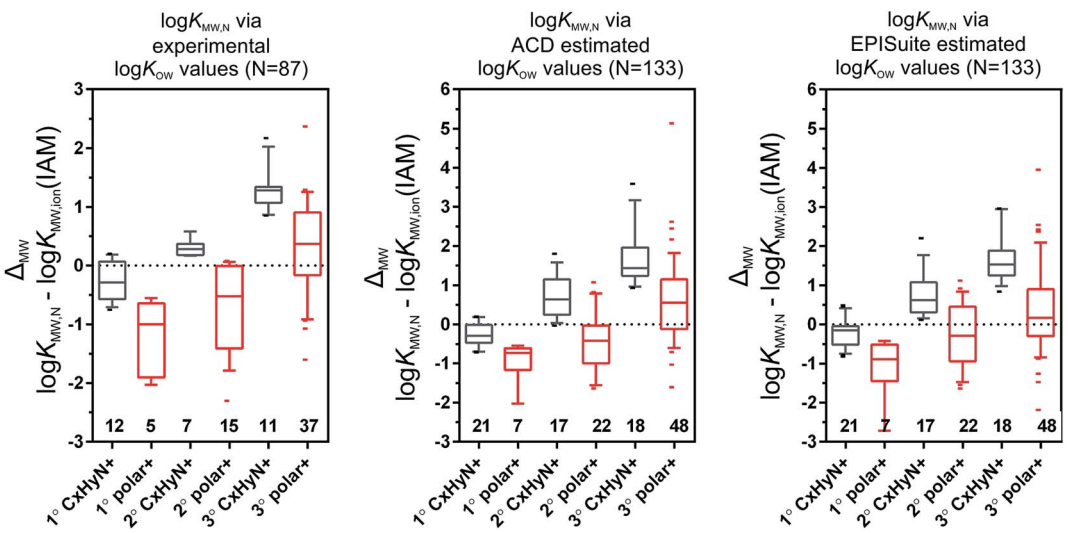


Fig. 1 Box-whisker plot evaluation of the scaling factor Δ_{MW} (as log $K_{MW,neutral}$−log $K_{MW,cation}$) for different types of amines, using $\delta_{IAM\text{-}SSLM}$ corrected IAM-HPLC based K_{MW} values for organic cations (log K_{MW}(IAM)), and K_{OW} based predictions of K_{MW} values for corresponding neutral species. $C_xH_yN^+$ amine data are from ref. 23, the polar amine data are from this study. Numbers on the *x*-axis refer to the number of amines included.

a factor of ±2 (~0.3 log units), which seems suitable for first screening purposes in risk assessment (except *e.g.* dialkylamines[25]), as was also concluded with the SSLM data for cationic surfactants.[23] However, provided that the (amine type corrected) IAM-HPLC data for polar amines are representative of partitioning affinity into phospholipid bilayers, the above analysis indicates that separate, lower, Δ_{MW}-values should be used for polar amine structures (*e.g.*, Δ_{MW} -1.4 ± 0.73 for polar 1° amines compared to -0.26 ± 0.32 for $C_xH_yN^+$ 1° amines), accompanied by wider uncertainty margins, particularly for more complex structures. As discussed below, part of this uncertainty in structural predictions of K_{MW} for ionogenic compounds with the Δ_{MW}-approach may be reduced by using COSMOmic software to directly calculate K_{MW} values for the ionic species.

3.3 K_{DMPC-W} values predicted by COSMOmic simulations with different DMPC bilayers

COSMOtherm V.15 provides an example DMPC bilayer file and a TZVP optimized 'representative' DMPC phospholipid structure. The atomic distributions of the hydrated DMPC bilayer input structure originate from a different molecular dynamics (MD) run compared to the TUHH-files used by Bittermann *et al.*[16] The TZVP-optimized DMPC structures are also quite different (Fig. S2†), with the TUHH-DMPC structure having two stretched alkyl chains, whereas the COSMOtherm15-DMPC structure has 1 chain curled such that the chains differ in their position relative to the charged headgroup. As a result, for all of the 86 tested cations the COSMOtherm15-DMPC files calculated K_{DMPC-W} values that are on average 0.40 log units lower than when using the TUHH-DMPC structure, ranging between 0.02 (#73, loperamide) and 0.66 (#50, serotonin) (Table S3†). There is no directly obvious link between the molecular structure of the amines and the difference in the K_{DMPC-W} values. Even the more complex cationic macrolides (#84–86) differed only by 0.4–0.49 log units. For anionic compounds the difference is surprisingly much larger; the COSMOtherm15-DMPC system predicts 1.2 log units higher values than the TUHH-DMPC system[16] (see example anions listed below Fig. S2†). This suggests that refitting the membrane potential parameters for the COSMOtherm15-DMPC system to align the K_{MW} data set used by Bittermann *et al.*[16] may result in closer alignment of cations and anions. We therefore currently prefer to apply the TUHH-system K_{DMPC-W} values, as listed in Table 1.

3.4 Comparison between K_{MW}(IAM) measurements (IAM-HPLC) and K_{DMPC-W} calculations (COSMOmic)

As discussed in Section 3.1, the influence of the *N*-methyl groups on charged amine structures differs between IAM-HPLC retention data and SSLM partitioning data, and $\delta_{IAM\text{-}SSLM}$ corrective increments were proposed recently.[23] Fig. 2A, C and E show the plots of three sets of COSMOmic predicted K_{DMPC-W} values and uncorrected $K_{IAM,intr}$ values, using the COSMOtherm15-DMPC system, the TUHH-DMPC system, and the TUHH-DMPC predictions including an offset value of −0.4 log units for organic cations recommended by Bittermann *et al.*,[16] respectively. Using the $\delta_{IAM\text{-}SSLM}$ corrective increments for amine types as set in Section 3.1, corresponding graphs are presented in 2B, D and F. Comparing the TUHH-DMPC predicted log K_{MW} values with the full IAM-HPLC data set for 155 amines without $\delta_{IAM\text{-}SSLM}$ correction gives an RMSE of 1.08, which is reduced to 0.80 when including the $\delta_{IAM\text{-}SSLM}$ correction. The $\delta_{IAM\text{-}SSLM}$ correction thus considerably reduced the scatter between $K_{IAM,intr}$ values and COSMOmic predicted K_{DMPC-W} values, both for COSMOmic's example DMPC files, and the TUHH-DMPC files.

The data for $C_xH_yN^+$ amines from our previous IAM-HPLC study are also plotted in all Fig. 2 graphs as black dots.[25] In that study, only the COSMOtherm15 example DMPC files were used, and corrective increments for *N*-methylation differences between IAM and COSMOmic were fitted to the K_{DMPC-W} values, in order to best align with the IAM-K_{PLIP-W} values.[25] In the current study, however, we optimize the $K_{IAM,intr}$ values with $\delta_{IAM\text{-}SSLM}$ corrective increments, using sorption data for cationic surfactants to SSLM phospholipid bilayers.[23] Thus, we now account for the underestimation of *N*-methylation effects by IAM-HPLC, and compare the corrected IAM-HPLC data directly to COSMOmic predictions. This $\delta_{IAM\text{-}SSLM}$ correction not only aligns IAM-HPLC data with phospholipid bilayer data, but it also reduces the scatter between IAM-HPLC data and COSMOmic predictions, as shown by comparing the top row graphs of Fig. 2 with the bottom row graphs of Fig. 2.

As shown in Fig. 2B by the position of the black regression line for $C_xH_yN^+$ amines relative to the 1 : 1 line, the COSMOtherm15 DMPC files systematically underestimate the $\delta_{IAM\text{-}SSLM}$-corrected K_{MW}(IAM) values, on average by −0.6 log units, with an RMSE of 0.71 log units. Using TUHH-DMPC files (Fig. 2D), COSMOmic underestimated $C_xH_yN^+$ amines on average by only 0.22 log units (see Table 3), with an RMSE of 0.48. As shown by the linear regression line for the $C_xH_yN^+$ amines, the TUHH-DMPC predictions are almost 1 : 1 with the IAM-HPLC data (Fig. 2D). For the 86 polar amines from the current study compared to $\delta_{IAM\text{-}SSLM}$-corrected K_{MW}(IAM) values, the RMSE for COSMOtherm15-DMPC (Fig. 2B) is 1.12 log units, while the RMSE for TUHH-DMPC (Fig. 2D) is 0.80.

Bittermann *et al.*,[16] however, actually found that K_{DMPC-W} values overestimated K_{MW} values for organic cations, and recommended an offset of −0.4 log units. They observed that this offset aligned predictions with the experimental K_{MW} data for 24 organic cations obtained using dispersed liposomes. As shown in Fig. 2E and F, doing so actually shifts the K_{DMPC-W} values the wrong way compared to the corrected 86 IAM-HPLC values, and this increases the RMSE for polar amines from 0.8 (without offset, 2D) to 1.12 log units (with offset, 2F). Similarly, the 0.4 log unit offset reduced the predictive accuracy of COSMOmic compared to the SSLM bilayer data for 19 cationic surfactants.[23] Regarding the consistency of the IAM-HPLC and SSLM data sets, and the diversity of liposomal K_{MW} references used in the data set evaluated by Bittermann *et al.*, we currently continue the comparisons between IAM-HPLC and COSMOmic without using the offset for TUHH-K_{DMPC-W} values. Still, it should be kept in mind that a thorough validation of IAM-HPLC and SSLM K_{MW} data with liposomal K_{MW} data is currently still lacking.

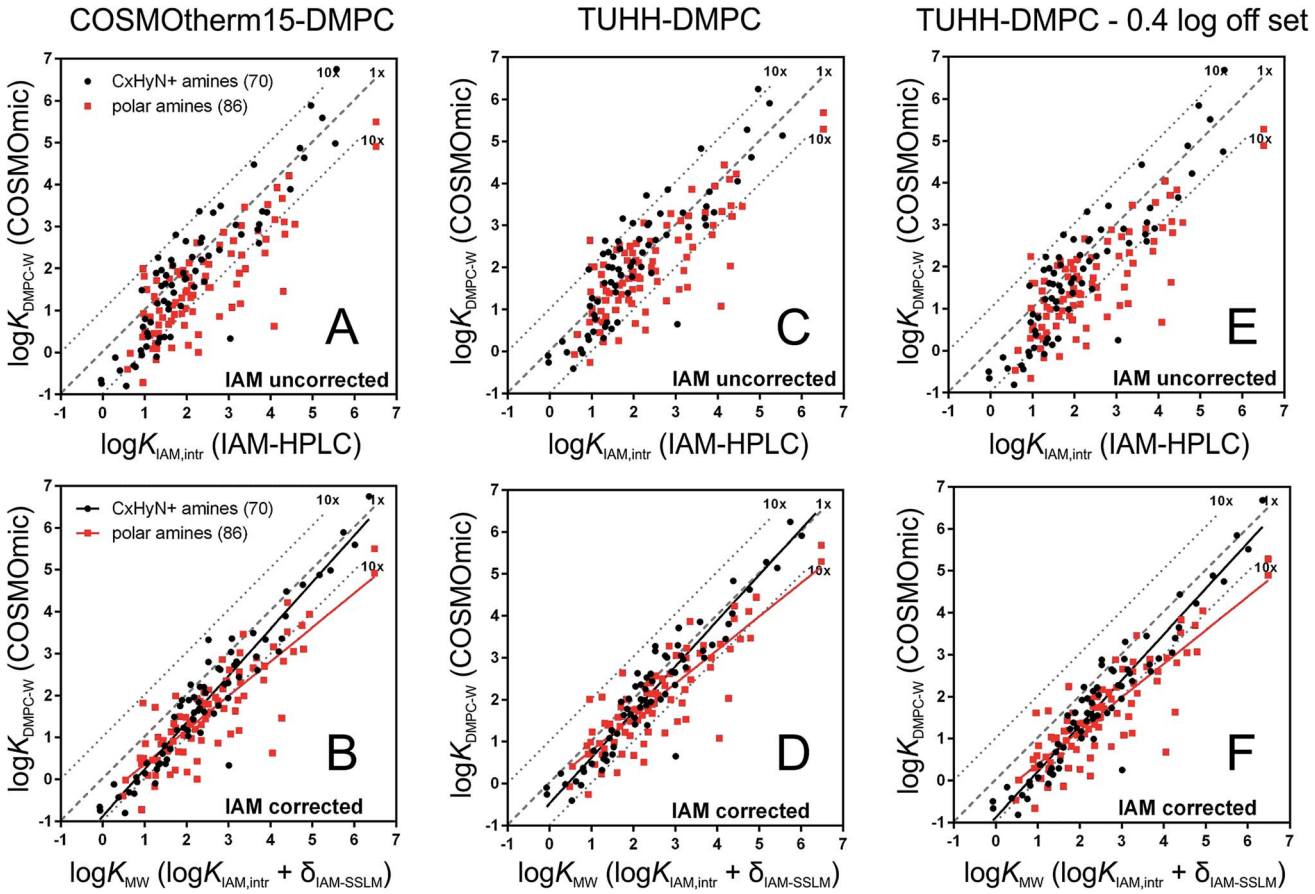


Fig. 2 Comparisons between IAM-HPLC values and COSMOmic predictions of $K_{DMPC-W,ion}$. The top row compares with uncorrected instrinsic sorption affinities to IAM monolayers K_{IAM}, while the bottom row applied $\delta_{IAM-SSLM}$ corrective increments for *N*-methylation differences in bilayers $K_{MW,ion}$(IAM). Graphs A and B represent predictions with COSMOmic example DMPC files. Graphs C and D represent predictions with TUHH-DMPC files used by Bittermann *et al.*[16] Graphs E and F include the recommended offset of −0.4 log units for cations by Bittermann *et al.*[16]

Table 3 RMSE values for different groups of polar amines, and suggested correction factors for log K_{DMPC-W} obtained with the DMPC files from Bittermann *et al.* (ref. 16), without using their recommended offset correction

	N	RMSE	Suggested correction factor	RMSE after correction factor
$C_xH_yN^+$ amines (without acridine)	70	0.48	+0.22	0.43
Halogenated/mono polar amines	22	0.68	+0.43	0.53
Simple dipolar amines (<4 polar moieties)	13	0.41	0	0.41
Amines with additional polar N moieties	25	0.69	+0.54	0.43
Polyaromatic amines	10	1.13	+0.79	0.51
Complex amines (≥4 polar groups/M_w > 400/large rings)	16	1.07	+0.60	0.89

We split the 86 polar amines into five different groups, according to polarity and complexity, to determine to what extent the predictive accuracy of COSMOmic for IAM-HPLC data is better for certain groups of structural features of the polar amines, compared to the RMSE of 0.82 for the 86 polar amines. Grouping the halogenated/non-polar amines with the simple monopolar amines with ether, ester and ketone moieties (compounds 1–22, Fig. 3A), an RMSE of 0.68 log units is obtained for 22 amines. On average these amines are underestimated by COSMOmic by −0.43 log units (±standard deviation (s.d.) of 0.55 log units), which may be considered as a representative K_{DMPC-W} scaling factor for this group of compounds, as listed in Table 3. As shown in Fig. 3A, the larger (polycyclic) compounds sertraline (#6) and escitalopram (#11) are the most underestimated compounds in this group.

The group of 13 "dipolar amines" with fewer than 4 functional groups, with at least 1 dipolar moiety (Fig. 3B, compounds 23–35), has an RMSE of 0.41. On average, these amines are underestimated in COSMOmic by only −0.09 ± 0.41

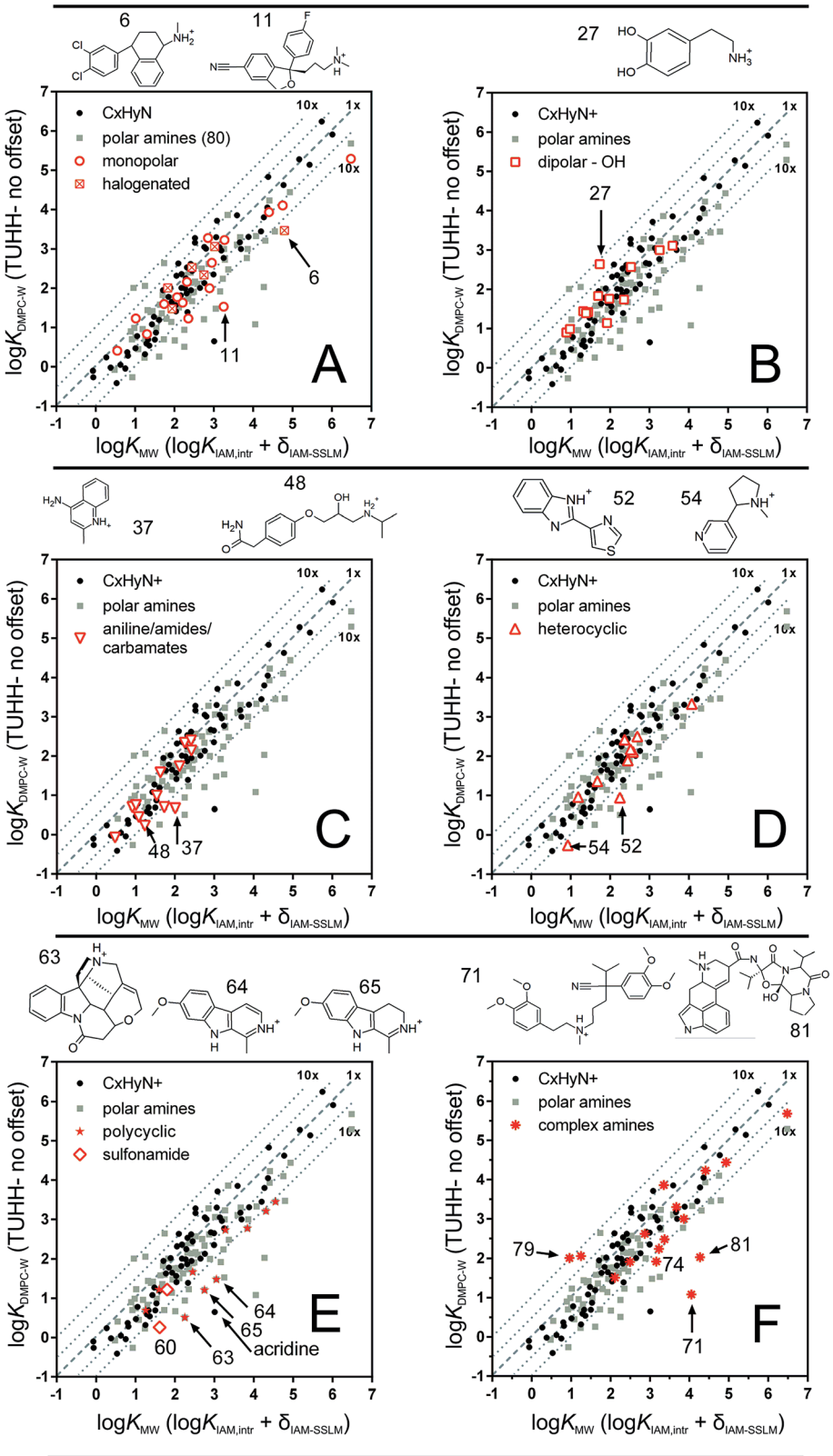


Fig. 3 COSMOmic predictions compared to $\delta_{IAM\text{-}SSLM}$ corrected IAM-HPLC data, highlighted for amines with various types of polar groups and structural complexity. Outliers are numbered and presented with cation structures. The highlighted types of amine structures are presented as red symbols.

log units. Compared to the monopolar amine structures, it thus seems that dipolar groups reduce the difference between IAM-HPLC and COSMOmic. The K_{DMPC-W} value for the small catechol amine dopamine (#27) is even overestimated by 0.8 log units (Fig. 3).

The group of 25 "amines with additional polar N moieties" that include anilines, amides, carbamates and N-heterocyclic structures (Fig. 3C and D, compounds 36–60) has an RMSE of 0.69. On average, these amines are underestimated in COSMOmic by -0.54 ± 0.44 log units. The cationic form of a polycyclic aminoquinoline (#37), the beta-blocker atenolol (48), the pesticide thiabendazole (#52), and nicotine (#54) have the most underestimated K_{DMPC-W} values.

The 10 polyaromatic amines (Fig. 3E, compounds 61–70) appear to be underestimated more by TUHH-DMPC than the previous groups, with an RMSE of 1.13 between IAM-HPLC and COSMOmic, and an average difference of -0.79 ± 0.47 log units. The prediction for the alkaloid pesticide strychnine (#63) deviates by 1.75 log units. This corresponds to the -2.4 log units underestimated K_{DMPC-W} relative to IAM-HPLC data for acridine, a tricyclic $C_xH_yN^+$ amine.[25] Also many of the most underestimated compounds mentioned above for the other groups have multiple rings. As noted elsewhere,[11,16] this may be a common problem of COSMOtherm with correctly dealing with polyaromatic compounds.

The RMSE for the group of 16 complex amine structures (Fig. 3F, compounds 71–86) is the largest of these groupings, 1.16 log units (average offset -0.63 ± 1.01). Still, the predictions for the three large macrolide antibiotics clarithromycin, erythromycin, and tylosin ($M_w > 730$) were well within a factor of 10 (overestimated by 0.14, 0.26 and 0.79 log units, respectively), just as for lincomycin (4 hydroxyl moieties, 0.80 log units). The largest deviating compounds were ergocornine (#81), a large (M_w of 562) polyaromatic compound underestimated by 2.24 log units, and verapamil (#71). Verapamil was underestimated by almost 3 orders of magnitude. However, closer inspection of the TZVP input structures of the six selected conformers showed that these were all structures where both polar ends, made up of two ethers linked to a phenyl ring, curled up around the central charged nitrogen (Fig. S3†). The weighted log K_{DMPC-W} for these structures was 1.08, ranging from 0.41 to 1.24. Using a separately created stretched conformer of verapamil (#71) predicts a log K_{DMPC-W} of 1.7, but this hardly improved the overall RMSE of this group (1.07 log units). Also, this is still 2.35 log units lower than the IAM-HPLC result, which was obtained with high confidence using the solvent mixture series at five different compositions (10–30% acetonitrile, Fig. S1†). Trimethoprim, a complex amine structure with three ethers linked to a phenyl ring, was underestimated by only 0.6 log units with TUHH-DMPC compared to the IAM-HPLC value. Apparently, it is not so much the ether linkages, but more selecting representative conformations for complex organic cations that remains a challenging feature in the COSMOmic approach, and part of the residual uncertainties in predicting K_{MW}.

3.5 Conclusions on establishing K_{MW} values for cationic compounds

The Δ_{MW}-approach in eqn (2) is a simple estimation method to derive the K_{MW} of ionic species based on K_{OW} predictions that can be readily incorporated in chemical risk assessment models. Based on reviewed liposomal partitioning data,[16,27,33] SSLM data for cationic surfactants,[23] and the current evaluation of IAM-HPLC, the Δ_{MW}-approach seems to be valuable for first chemical screening of basic compounds if the Δ_{MW}-values are not only specific for amine type, but also for the presence of polar functional groups. The Δ_{MW}-approach does not apply to quaternary ammonium compounds (QACs), and has higher uncertainty margins for more complex amines, especially when experimental K_{OW} values are not available. For such chemicals, measured IAM-HPLC values provide direct measures of the K_{MW} of ionic species, with the inclusion of $\delta_{IAM-SSLM}$ corrective increments for each amine type. There is an upper limit at log K_{MW} of ~6 to the experimentally feasible affinity to use IAM-HPLC, but experimental data on smaller analogues can be used to extrapolate for some organic cations.[25] Alternatively, the current study shows that the $K_{MW,ion}$ of a wide variety of amine structures can be predicted with relatively high confidence with COSMOmic, including QACs, using the TUHH-DMPC system files. We propose six empirically defined $\delta_{DMPC-IAM}$ corrective increments to the calculated K_{DMPC} values (Table 3), to improve the accuracy of COSMOmic for certain types of organic cation structures.

Obviously, there is a disadvantage in applying corrective increments to IAM-HPLC, offset values and somewhat arbitrarily selecting the input files in COSMOmic, in attempts to align the observed data. It is clear that the level of confidence in the experimental output of liposomal partition studies is higher than those of SSLM bilayers, while the latter are still more representative of cell membranes than the IAM-HPLC monolayer coating. Regarding the current data analysis on 156 amines, we would rate the confidence in the COSMOmic predictions, for organic cations, even on a relative scale, to be at a comparable level to the IAM-HPLC data. COSMOmic performs calculations with a bilayer while IAM consists of a monolayer phospholipid coating, and the effect of *N*-methylation is better predicted by COSMOmic than by IAM. Still, even the use of 3-dimensional input structures is not likely to be as representative as the retention behaviour of fully dissolved molecules on an actual phospholipid membrane, as shown by the examples of verapamil (#71) and tricyclic $C_xH_yN^+$ acridine, and very large alkaloids such as ergocornine (#81). Confidence in the COSMOmic predictions, and suggested offset values and corrective increments, increases with larger and more consistent data sets. The liposomal data on 25 organic cations reviewed and applied by Bittermann *et al.* to validate COSMOmic calculations with a DMPC bilayer were collected from a variety of studies, and were of limited structural variability. The SSLM data were obtained with a consistent methodology in a single study on a set of nineteen linear alkylamine-based surfactants, very simple cationic structures that covered all four types of *N*-methylation. IAM-HPLC chromatography allows for

even more methodological consistency than SSLM. Together with the set of 70 $C_xH_yN^+$ amines from a closely related study using similar test conditions,[25] we now have experimental data on 156 different structures tested on a single IAM-HPLC column to compare with COSMOmic. It thus seems fair to use simple correction terms to optimize the IAM-HPLC data set, and to use the extensive IAM-HPLC data set to evaluate the appropriateness of certain different COSMOmic input files and parameters. Dividing the chemical domain into simple amine structures (<4 polar groups) and complex amine structures (≥4 polar groups), we obtain an RMSE of 0.45 log units for the 138 simple amines, and 0.89 log units for 16 complex amine structures. As shown in Fig. 4, all simple amines except acridine are predicted within a factor of ±10, and the majority within a factor of ±3. The complex amines are predicted with reasonable success, with all tested compounds within a factor ±50.

Based on the findings of the current and closely related other studies, our suggestions for deriving K_{MW} values for (predominantly) cationic compounds experimentally *via* IAM-HPLC, or prediction *via* COSMOmic or the Δ_{MW}-approach can thus be summarized as:

• IAM-HPLC measurements (for quaternary ammonium compounds or basic compounds with $pK_a > 7$) should be performed with an IAM.PC.DD2 column using eluent buffered at pH 5 and a salinity of ≥10 mM; weaker bases with a pK_a between 5 and 7 should be tested at pH 3, with a salinity of 150 mM. For bases with a $pK_a < 5$, there is little need for measurements of the K_{MW} of the ionic species. From these IAM-HPLC measurements, representative K_{MW}(IAM) values can be obtained following eqn (5)–(8), with a Δ_{pH5-3} corrective increment of +0.12 log units, and $\delta_{IAM\text{-}SSLM}$ corrective increments of +0.78 log units for 1° amines, +0.47 log units for 2° amines, −0.03 log units for 3° amines and −0.11 log units for 4° quaternary ammonium compounds.

- For compounds with log $K_{MW} > 6$, the IAM-HPLC method is impracticable.

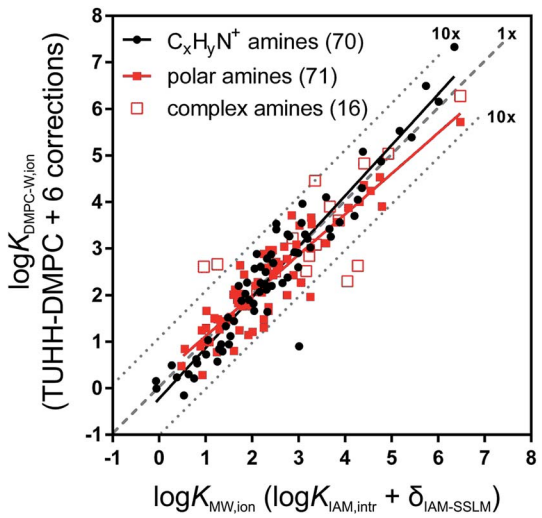


Fig. 4 Comparisons between corrected IAM-HPLC values $K_{MW,cation}$(IAM), and corrected COSMOmic predictions of $K_{DMPC-W,cation}$, following the correction terms listed in Table 3.

- Experimental K_{MW} values obtained with dissolved liposomes are considered most realistic, but standardization of a testing protocol would be valuable. SSLM measurements should ensure sufficient contact time between solutes and suspended particles, and ensure that phospholipids do not significantly detach from the solid particles. Although a consistent SSLM data set for $C_xH_yN^+$ amines is currently becoming available for comparison with IAM data (by which $\delta_{IAM\text{-}SSLM}$ has been defined), more experimental SSLM or liposome based K_{MW} values should be generated to further validate the IAM-HPLC data for polar amines.

• COSMOmic predictions should apply the TUHH-DMPC files, which are available upon request, rather than the COSMOtherm15-DMPC files. The IAM-HPLC data suggest that a previously recommended generic offset of −0.4 log units for the generated K_{DMPC-W} values for all organic cations will actually reduce the predictive accuracy. Instead, we propose the following corrective increments: +0.22 log units for $C_xH_yN^+$ amines, +0.43 log units for halogenated or simple monopolar amines, +0.09 log units for simple dipolar amines, +0.54 log units for simple amines with polar N groups and +0.79 for polyaromatic amines. We advise caution with COSMOmic predictions for large polyaromatic amines, as well as for complex amine structures (≥4 polar groups, and/or large rings, and/or $M_w > 400$; +0.59 log unit correction). COSMOmic may well be used to screen whether experimental K_{MW} determination for certain organic cations is feasible.

• For predictions of the log K_{MW} for bases by the Δ_{MW}-approach (log $K_{MW,N}$ − Δ_{MW} = log $K_{MW,ion}$), based on log K_{OW} input values to derive log $K_{MW,N}$, we suggest to use six distinct values to separate amine types and the presence of polar groups: −0.26 for 1° $C_xH_yN^+$ amines, −1.4 for 1° polar amines, +0.30 for 2° $C_xH_yN^+$ amines, −0.57 for 2° polar amines, +1.29 for 3° $C_xH_yN^+$ amines, and +0.35 for 3° polar amines.

• For multiprotic organic ions neither IAM-HPLC nor COSMOmic has been extensively validated.

These COSMOmic-predicted or IAM-HPLC based K_{MW} values of ionised species may be used, alongside K_{MW} values of corresponding neutral species, to predict partitioning of basic contaminants into biotic compartments. For example, this will improve the accuracy of predictions of the pH-dependent bioconcentration factors (BCF) of ionisable amines in fish,[27] and the interpretation of pH-dependent toxicity tests and analysis of the specificity of the toxic mode of action, as shown for toxicity tests with basic pharmaceuticals using algae.[26,33]

Acknowledgements

This study was funded by Unilever, Safety & Environmental Assurance Centre (SEAC), Colworth Science Park, Sharnbrook, United Kingdom.

References

1 J. Comer and K. Tam, in *Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical and Computational Strategies*, Wiley-VCH, Zurich, CH, 2001.
2 D. T. Manallack, *Persp. Med. Chem.*, 2007, **1**, 25.

3 A. Negrusz and G. A. A. Cooper, *Clarke's analytical forensic toxicology*, Pharmaceutical press, London, UK, 2nd edn, 2013.
4 S. Rayne and K. Forest, *J. Environ. Sci. Health, Part A: Toxic/Hazard. Subst. Environ. Eng.*, 2010, **45**, 1550.
5 W. C. Scott, B. Du, S. P. Haddad, C. S. Breed, G. N. Saari, M. Kelly, L. Broach, C. K. Chambliss and B. W. Brooks, *Environ. Toxicol. Chem.*, 2016, **35**, 983.
6 Y. Chen, J. L. M. Hermens and S. T. J. Droge, *Environ. Pollut.*, 2013, **179**, 153.
7 S. T. J. Droge and K.-U. Goss, *Environ. Sci. Technol.*, 2012, **46**, 5894.
8 S. T. J. Droge and K.-U. Goss, *Environ. Sci. Technol.*, 2013, **47**, 14224.
9 S. T. J. Droge and K.-U. Goss, *Environ. Sci. Technol.*, 2013, **47**, 14233.
10 J. G. Teeguarden, C. Tan, S. Edwards, J. A. Leonard, K. A. Anderson, R. A. Corley, A. Harding, M. Kile, S. L. Massey Simonich, D. Stone, K. M. Waters, S. Harper, D. E. Williams and R. L. Tanguay, *Environ. Sci. Technol.*, 2016, **50**, 4579.
11 S. Endo, B. I. Escher and K.-U. Goss, *Environ. Sci. Technol.*, 2011, **45**, 5912.
12 R. P. Austin, A. M. Davis and C. N. Manners, *J. Pharm. Sci.*, 1995, **84**, 1180.
13 B. I. Escher, R. P. Schwarzenbach and J. C. Westall, *Environ. Sci. Technol.*, 2000, **34**, 3954.
14 C. Ottiger and H. Wunderli-Allenspach, *Eur. J. Pharm. Sci.*, 1997, **5**, 223.
15 A. Avdeef, K. J. Box, J. E. A. Comer, C. Hibbert and K. Y. Tam, *Pharm. Res.*, 1998, **15**, 209.
16 K. Bittermann, S. Spycher, S. Endo, L. Pohler, U. Huniar, K.-U. Goss and A. Klamt, *J. Phys. Chem. B*, 2014, **118**, 14833.
17 A. Avdeef, *Absorption and Drug Development: Solubility, Permeability, and Charge State*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2003.
18 R. P. Austin, A. M. Davis and C. N. Manners, *J. Pharm. Sci.*, 1995, **84**, 1180.
19 K. Bittermann, S. Spycher and K.-U. Goss, *Chemosphere*, 2016, **144**, 382.
20 M. J. Hope, M. B. Bally, G. Webb and P. R. Cullis, *Biochim. Biophys. Acta, Biomembr.*, 1985, **812**, 55.
21 B. I. Escher, R. P. Schwarzenbach and J. C. Westall, *Environ. Sci. Technol.*, 2000, **34**, 3962.
22 A. Loidl-Stahlhofen, T. Hartmann, M. Schottner, C. Rohring, H. Brodowsky, J. Schmitt and J. Keldenich, *Pharm. Res.*, 2001, **18**, 1782.
23 N. Timmer and S. T. J. Droge, *Environ. Sci. Technol.*, 2017, DOI: 10.1021/acs.est.6b05662.
24 S. T. J. Droge, *Anal. Chem.*, 2016, **88**, 960.
25 S. T. J. Droge, J. L. M. Hermens, J. Rabone, S. Gutsell and G. Hodges, *Environ. Sci.: Processes Impacts*, 2016, **18**, 1011.
26 J. Neuwoehner and B. I. Escher, *Aquat. Toxicol.*, 2011, **101**, 266.
27 J. M. Armitage, J. A. Arnot, F. Wania and D. Mackay, *Environ. Toxicol. Chem.*, 2013, **32**, 115.
28 A. Klamt, U. Huniar, S. Spycher and J. Keldenich, *J. Phys. Chem. B*, 2008, **112**, 12148.
29 A. Klamt and F. Eckert, *Fluid Phase Equilib.*, 2000, **172**, 43.
30 F. Eckert and A. Klamt, *AIChE J.*, 2002, **48**, 369.
31 S. Ong and D. Pidgeon, *Anal. Chem.*, 1995, **67**, 2119.
32 S. Jakobtorweihen, T. Ingram and I. Smirnova, *J. Comput. Chem.*, 2013, **34**, 1332.
33 J. Neuwoehner, K. Fenner and B. I. Escher, *Environ. Sci. Technol.*, 2009, **43**, 6830.