

Equilibrium Partitioning of Ionic Organic  
Chemicals in Phospholipid Membranes:  
Experiments and Model Predictions  
Dissertation

zur Erlangung des  
Doktorgrades der Naturwissenschaften  
(Dr. rer. nat.)

der

Naturwissenschaftlichen Fakultät II – Chemie, Physik und Mathematik  
der Martin-Luther-Universität Halle-Wittenberg

vorgelegt

von Herrn Dipl. Chem. Kai Bittermann  
geboren am 29. Juli 1983 in Neustadt an der Aisch

Gutachter: 1. Prof. Dr. Kai-Uwe Goss  
2. Prof. Dr. Kathrin Fenner

Verteidigt am: 5. Dezember 2017

## Abstract

Ionizable or even permanently charged anthropogenic organic chemicals are pollutants of prevailing concern and a much debated issue. Given that ionic organic chemicals have different physico-chemical characteristics than neutral chemicals, the empirical literature models describing the environmental fate of neutral chemicals cannot be applied one to one. Hence, it is of pivotal interest to develop mechanistic models that describe relevant physico-chemical characteristics of ions such as the sorption behavior. This work focused on describing of the partition coefficient of ions between phospholipid membranes and water ( $K_{lipw}$ ), which is a crucial descriptor for environmentally relevant properties such as bioaccumulation and non-specific toxicity.

The main aims of this work were threefold: i) to assess the predictive power of the commercially available software *COSMOmic* (i.e., COSMO-RS for micelles) for the prediction of the partition behavior of organic ions. The model adapts the conductor-like screening model for real solvents (COSMO-RS) to anisotropic phases and has been shown to reliably predict  $K_{lipw}$  for neutral chemicals before. In case it might not be applicable for ions the goal was to detect the reason and implement potentially missing parameters in the model. ii) to compare the enhanced version of *COSMOmic* with two competing models to predict  $K_{lipw}$  of ions, namely an empirical approach based on the octanol-water partition coefficient ( $K_{ow}$ ) and the polyparameter linear free energy relationship (pp-LFER) approach. The prediction of  $K_{lipw}$  for neutral chemicals was included in this analysis to gain an exhaustive picture that incorporates the computational effort for the respective model and scrutinizes model consistency. iii) to investigate the baseline toxicity concept for ions with the help of the most reliable prediction model for  $K_{lipw}$ .

The  $K_{lipw}$  data gathered through an exhaustive literature research were complemented by own  $K_{lipw}$  measurements using equilibrium dialysis, in order to systematically increase data diversity. The resulting compilation of 51 experimental  $K_{lipw}$  values for anions (from which

five are own measurements in this work) and 24 experimental values for cations revealed that COSMOmic (version 1401) systematically overestimated the  $K_{lipw}$  of cations and underestimated the  $K_{lipw}$  of anions. To make the COSMOmic model applicable for ionic chemicals, the internal membrane dipole potential was implemented. We empirically optimized the potential with experimental  $K_{lipw}$  data of 161 neutral and 75 ionic chemicals, yielding shapes of the potentials that agree well with experimentally determined potentials from the literature. This model refinement had no negative effect on the prediction accuracy of neutral chemicals (RMSE = 0.62 log units), while it highly improved the prediction of ions (RMSE = 0.70 log units).

This enhanced version of COSMOmic (version 1501) was compared to two other models for the prediction of  $K_{lipw}$ , using a further extended data set of 56 anions, 36 cations, 2 divalent cations and 2 zwitterions (as well as 207 neutral chemicals for ensuring model consistency). The empirical correlation with  $K_{ow}$  of the corresponding neutral species yielded better results for the prediction of anions (RMSE=0.79) than for cations (RMSE=1.14). Though describing most anions reasonably well, the lack of mechanistic basis and the poor performance for cations constrain the usage of this model. The pp-LFER model performs worse for ions (RMSE=1.01/1.04 for anions/cations) than for neutral chemicals (RMSE=0.53) and also strongly depends on the fitting procedure. The differently charged species preferentially sorb to different membrane depths, according to the COSMOmic calculations, and are therefore encompassed by a different physicochemical environment. This cannot be described with a single pp-LFER model. COSMOmic has the widest applicability domain; it was the only model applicable for multiply charged chemicals and gave the best results for anions (RMSE=0.66) and cations (RMSE=0.71). In terms of  $K_{lipw}$  prediction of neutral chemicals, both the  $K_{ow}$  based model (RMSE=0.52) as well as the pp-LFER model (RMSE=0.53) are computationally less demanding than COSMOmic (RMSE=0.74). If any

mechanistic understanding of the partitioning process is desired, the pp-LFER approach will be the more suited one of both.

Finally, the successful modeling of  $K_{lipw}$  was applied to investigate the baseline toxicity concept. The two principal assumptions behind the baseline toxicity concept are a) that baseline toxicity can be described independently of the organism with only one partition coefficient (here the  $K_{lipw}$  is taken as a surrogate for the partitioning between real biological membranes and water) and b) that a critical toxic concentration in the membrane causing a toxic effect is fairly independent of the nature of the chemical. First, the range of organism- and chemical independent critical membrane concentrations causing 50% mortality ( $c_{mem}^{tox}$ ) was reevaluated based on a critical revision of a previously published toxicity dataset for neutral chemicals. In accordance to values reported in the literature a mean value for  $c_{mem}^{tox}$  of roughly 100 mmol/kg (membrane lipid) could be determined, based on pp-LFER predicted  $K_{lipw}$  values, for a broad variety of 42 aquatic organisms (333 different chemicals), albeit with a considerable scatter. Then this concept was applied to permanently charged ionic liquids (ILs). Using the enhanced COSMOmic,  $K_{lipw}$  of the anionic and cationic IL components was predicted. Doing so,  $c_{mem}^{tox}(total)$  for the ILs could be estimated assuming independent, concentration additive contributions of the cationic and its respective anionic species. The resulting values for some of the toxicity data for ILs were consistent with the expected range for baseline toxicity for neutral chemicals while other values were consistently greater or smaller. Based on the calculation of toxic ratios, ILs could be identified that exert a specific mode of toxic action and experimental data could be detected that are most likely due to experimental artefacts. It has to be kept in mind though, that the use of nominal concentrations instead of freely-dissolved concentrations in the published literature hampers definite conclusions.

The herein presented improvement of COSMOmic might help in future to not only further investigate the toxicity of charged chemicals but also their bioaccumulation potential.

Moreover, the presented work is a first step to subsequently extend the *COSMOmic* model to the calculation of membrane permeabilities of neutral and ionic chemicals, which will help to better understand toxicodynamic processes (such as ion trapping) as well as specific toxic modes of action (such as uncoupling).

## Zusammenfassung

Ionisierbare sowie permanent geladene anthropogene organische Chemikalien stellen eine zurzeit viel diskutierte Klasse von Schadstoffen dar. Da ionische organische Chemikalien andere physikochemische Eigenschaften als neutrale Chemikalien aufweisen, können weitläufig in der Literatur beschriebene empirische Modelle, die die Umweltverträglichkeit von Letzterem beschreiben, nicht eins-zu-eins übertragen werden. Aus diesem Grund ist es von zentraler Bedeutung mechanistisch fundierte Modelle zu entwickeln, die die relevanten physikochemischen Eigenschaften von Ionen erfassen. Dazu zählt insbesondere das Sorptionsverhalten. Die vorliegende Arbeit konzentrierte sich deshalb auf die Beschreibung der Verteilung von Ionen zwischen Phospholipid-Membranen und Wasser ( $K_{lipw}$ ), eines entscheidenden Deskriptors für umweltrelevante Eigenschaften wie Bioakkumulation und nichtspezifische Toxizität.

In dieser Arbeit wurden drei Hauptziele verfolgt: i) Die Eignung der kommerziell erhältlichen Software *COSMOmic* (also COSMO-RS für Mizellen) zu testen, das Verteilungsverhalten organischer Ionen vorherzusagen. Das besagte Modell adaptiert das „conductor-like screening model for real solvents“ (COSMO-RS) für anisotrope Phasen und kann, wie bereits gezeigt wurde,  $K_{lipw}$  neutraler Chemikalien zuverlässig beschreiben. Für den Fall, dass *COSMOmic* sich als für Ionen ungeeignet herausstellt, sollte der Grund dafür gefunden und potentiell fehlende Parameter im Modell implementiert werden. ii) Die verbesserte Version von *COSMOmic* mit zwei konkurrierenden Modellen zur Vorhersage des  $K_{lipw}$  von Ionen vergleichen: einerseits ein empirischer Ansatz, der auf dem Octanol-Wasser-Verteilungskoeffizienten ( $K_{ow}$ ) beruht, sowie andererseits ein Ansatz, der auf der „polyparameter linear free energy relationship“ (pp-LFER) aufbaut. Die Vorhersage von  $K_{lipw}$  neutraler Chemikalien wurde in der Analyse mit berücksichtigt, um ein ganzheitlicheres Bild zu bekommen, das auch den rechnerischen Aufwand und die Konsistenz der einzelnen

Modelle mit erfasst. iii) Mit Hilfe des verlässlichsten Vorhersagemodells für  $K_{lipw}$  das Konzept der minimal zu erwartenden Toxizität („baseline toxicity“ oder auch narkotische Toxizität) für Ionen untersuchen.

Die durch eine umfassende Literaturrecherche zusammengetragenen  $K_{lipw}$  Werte wurden durch eigene Messungen ergänzt, um die Diversität des Datensatzes systematisch zu erhöhen. Die durchgeführten Experimente beruhen auf der Gleichgewichtsdialyse. Die resultierende Zusammenstellung von 51 experimentellen  $K_{lipw}$  Werten für Anionen (von denen fünf selbst vermessen wurden) und 24 experimentellen  $K_{lipw}$  Werten für Kationen zeigte, dass COSMOmic (Version 1401) den  $K_{lipw}$  von Kationen systematisch überschätzt, während es den  $K_{lipw}$  von Anionen systematisch unterschätzt. Um COSMOmic auch für ionische Chemikalien nutzbar zu machen, wurde das interne Membrandipolpotential mit experimentellen  $K_{lipw}$  Werten von 161 neutralen und 75 ionischen Chemikalien empirisch optimiert und implementiert. Die so gewonnenen Potentialformen stimmen gut mit den experimentell bestimmten Potentialen aus der Literatur überein. Diese Weiterentwicklung des Modells hat keine negativen Effekte auf die Vorhersagegenauigkeit für neutrale Chemikalien (RMSE = 0.62 log Einheiten), verbessert die Vorhersage für Ionen jedoch deutlich (RMSE = 0.70 log Einheiten).

Die derart verbesserte Version von COSMOmic (Version 1501) wurde mit Hilfe eines nochmals erweiterten Datensatzes (56 Anionen, 36 Kationen, 2 divalente Kationen, 2 Zwitterionen sowie 207 neutrale Chemikalien zur Überprüfung der Modellkohärenz) mit zwei anderen Modellen zur Vorhersage von  $K_{lipw}$  Werten verglichen. Die empirische Korrelation mit dem  $K_{ow}$ -Werten der korrespondierenden neutralen Chemikalien erbrachte bessere Ergebnisse für die Vorhersage der Anionen (RMSE=0.79) als der Kationen (RMSE=1.14). Obwohl die meisten Anionen hinreichend gut beschrieben wurden, ist eine allgemeine Anwendbarkeit des Modells für Anionen fragwürdig aufgrund der fehlenden mechanistischen Basis des Modells. Das pp-LFER Modell schneidet bei der Vorhersage der Ionen schlechter

ab (RMSE=1.01/1.04 für Anionen/Kationen) als bei der Vorhersage der neutralen Chemikalien und ist stark von der Art und Weise abhängig, wie während der Modellkalibration gefittet wird. Die verschieden geladenen Spezies sorbieren, laut COSMOmic Berechnung, vorzugsweise unterschiedlich tief in die Membran und befinden sich infolgedessen in unterschiedlichen physikochemischen Umgebungen. Dies kann nicht in einer einzigen pp-LFER-Gleichung erfasst werden. COSMOmic hat von den drei untersuchten Modellen den größten Anwendungsbereich. Es war das einzige Modell, das auch auf mehrfach geladene Chemikalien anwendbar ist, und gab die besten Resultate für Anionen (RMSE=0.66) sowie für Kationen (RMSE=0.71). Hinsichtlich der  $K_{lipw}$  Vorhersage der neutralen Chemikalien sind sowohl das empirische, auf  $K_{ow}$  beruhende Modell (RMSE=0.52) als auch das pp-LFER Modell (RMSE=0.53) weniger rechenintensiv als COSMOmic (RMSE=0.74). Wenn man ein mechanistisches Verständnis für den Verteilungsprozess gewinnen möchte, dann ist der pp-LFER Ansatz der geeignetere von beiden.

Schließlich wurde die erfolgreiche Modellierung von  $K_{lipw}$  eingesetzt, um das Konzept der Basistoxizität näher zu untersuchen. Die zwei Grundannahmen hinter diesem Konzept sind a) die vom Organismus unabhängige Beschreibung der Basistoxizität mit nur einem Verteilungskoeffizienten (wobei der  $K_{lipw}$  hier als Ersatz für echte biologische Membranen benutzt wird) und b) die Induktion eines toxischen Effekts durch eine gewisse kritische, von der Art der Chemikalie weitestgehend unabhängige Membrankonzentration. Hierzu wurde vorab die Bandbreite der organismus- und chemikalienunabhängigen kritischen Membrankonzentration, die eine 50%ige Mortalität verursacht, ( $c_{mem}^{tox}$ ), auf Basis einer kritischen Revision eines bereits veröffentlichten Datensatzes zur Toxizität neutraler Chemikalien reevaluiert. Für eine Menge von 42 aquatischen Organismen konnte in Übereinstimmung mit den in der Literatur berichteten Werten ein Mittelwert für  $c_{mem}^{tox}$  von etwa 100 mmol/kg (Membranlipid) ermittelt werden (333 verschiedene Chemikalien), allerdings mit einer beträchtlichen Streuung. Anschließend wurde das Konzept auf permanent



geladene ionische Flüssigkeiten (ILs) angewandt. Die  $K_{lipw}$ -Werte der anionischen und kationischen IL-Komponenten wurden mit dem verbesserten COSMOmic berechnet. Durch die Annahme unabhängiger, konzentrationsadditiver Beiträge der kationischen und der zugehörigen anionischen IL-Komponenten konnte  $c_{mem}^{tox}(total)$  der ILs abgeschätzt werden. Die resultierenden Werte für einige der Toxizitätsdaten der ILs waren konsistent mit dem zuvor ermittelten Bereich, in dem narkotische Toxizität erwartet wird, während andere Werte durchweg höher oder auch niedriger waren als erwartet. Durch die Berechnung der „toxic ratios“ (d.h., dem Verhältnis aus vorhergesagter zu experimenteller Wasserkonzentration, die einen toxischen Effekt ausübt) konnten ILs identifiziert werden, die spezifisch toxisch sind. Darüber hinaus konnten experimentelle Datenpunkte herausgefiltert werden, die aller Wahrscheinlichkeit nach auf experimentellen Artefakten beruhen. Es muss hier allerdings darauf hingewiesen werden, dass die Verwendung nomineller wässriger Konzentrationen an Stelle von frei gelösten Konzentrationen in der Literatur die Möglichkeit einschränkt definitive Schlussfolgerungen zu treffen.

Die in dieser Arbeit präsentierten Verbesserungen von COSMOmic könnten in Zukunft nicht nur dabei behilflich sein, die Toxizität ionischer Chemikalien weitergehend zu verstehen, sondern auch ihr Bioakkumulationsvermögen zu untersuchen. Darüber hinaus stellt die vorliegende Arbeit einen ersten Schritt zur sukzessiven Erweiterung von COSMOmic für die Berechnung von Membranpermeabilitäten dar. Dies kann helfen sowohl, toxikodynamische Prozesse (wie „ion trapping“) als auch spezifische toxische Effekte (wie Entkopplung) besser zu ergründen.

## Preface

The present work was performed between November 2011 to May 2017 at the Helmholtz Centre for Environmental Research, Leipzig at the Department of Analytical Environmental Chemistry. The thesis was written in a cumulative form and is based on the following articles:

Bittermann, K., Spycher, S., Endo, S., Pohler, L., Huniar, U., Goss, K.-U., Klamt, A., 2014. Prediction of Phospholipid–Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic. *J. Phys. Chem. B* 118, 14833–42.

doi: 10.1021/jp509348a. (SI-1 available at

[http://pubs.acs.org/doi/suppl/10.1021/jp509348a/suppl\\_file/jp509348a\\_si\\_001.pdf](http://pubs.acs.org/doi/suppl/10.1021/jp509348a/suppl_file/jp509348a_si_001.pdf))

Bittermann, K., Spycher, S., Goss, K.-U., 2016. Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds. *Chemosphere* 144, 382–391. doi:10.1016/j.chemosphere.2015.08.065. (SI-2 available at

<http://www.sciencedirect.com/science/article/pii/S0045653515300655#MMCvFirst>)

Bittermann, K., Spycher, S., Goss, K.-U., 2017. Erratum to “Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds” [*Chemosphere* 144C (2016) 382–391]. *Chemosphere* 179, 405–406.

doi:10.1016/j.chemosphere.2017.03.132

Bittermann, K., Goss, K.-U., 2017. Assessing the toxicity of ionic liquids – Application of the Critical Membrane Concentration approach. *Chemosphere*.

doi:10.1016/j.chemosphere.2017.05.097. (SI-3 available at

<http://www.sciencedirect.com/science/article/pii/S0045653517308019>)

Note that text passages, tables and figures in this work are partly taken from the above listed original publications without further indication. Three additional Co-author publications are only cited in this work. The abstracts of all original publications were included at the end, the supporting information SI-1 to 3 can be found in the Appendix.



## Contents

<b>Abstract .....</b>	<b>II</b>
<b>Zusammenfassung .....</b>	<b>VI</b>
<b>Preface .....</b>	<b>X</b>
<b>Contents.....</b>	<b>XII</b>
<b>1 Summary: Equilibrium Partitioning of Ionic Organic Chemicals in Environmental Systems: Experiments and Model Predictions .....</b>	<b>1</b>
1.1 Introduction .....	1
1.1.1 What Makes Liposomes Special Compared to Bulk Solvents.....	2
1.1.2 Problems Arising When Octanol is Taken as a Surrogate for Anisotropic Membranes .....	3
1.1.3 Difficulties in the Modelling of Anisotropic Membranes .....	6
1.1.4 Environmentally Relevant Application of $K_{lipw}$ Predictions .....	7
1.2 Objective of this Work .....	9
1.2.1 Calibrating COSMOmic for the Use with Ions .....	9
1.2.2 Comparison with Two Other Models Predicting $K_{lipw}$ .....	9
1.2.3 Application of Predicted Partition Coefficients on Baseline Toxicity Concept for Ions .....	10
1.3 Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals using the Mechanistic Model COSMOmic .....	11
1.3.1 Materials and Methods .....	11
1.3.1.1 Determination of Membrane-Water Distribution Coefficients .....	11
1.3.1.2 Data Collection and Evaluation.....	12
1.3.2 Theoretical section .....	12
1.3.2.1 COSMO-RS and COSMOmic.....	13
1.3.2.2 Estimation of Membrane Potentials .....	13
1.3.2.3 Optimization of a Model Membrane Potential.....	16
1.3.3 Results and Discussion.....	18
1.3.3.1 Predicting $K_{lipw}$ Using COSMO-RS with Phosphatidylcholine Lipid as Bulk Solvent.....	18

---

1.3.3.2	Predicting $K_{lipw}$ using COSMOmic without considering the membrane potential $\Psi_d$ .....	19
1.3.3.3	Using COSMOmic with an Optimized Membrane Potential $\Psi_d$ .....	22
1.3.3.4	The influence of the Membrane Potential on the Prediction of Neutral Chemicals .....	29
1.4	Comparison of Different Models Predicting the Phospholipid-Membrane Water Partition Coefficients of Charged Chemicals.....	30
1.4.1	Materials and Methods .....	30
1.4.1.1	$K_{lipw}$ Data Compilation for Neutral and Ionic Chemicals .....	30
1.4.1.2	Empirical Correlation Approach with $\log K_{ow}$ .....	31
1.4.1.3	PP-LFER Extension for Ionic Compounds .....	31
1.4.1.4	COSMO-RS and COSMOmic.....	34
1.4.2	Results and Discussion.....	34
1.4.2.1	Empirical Correlation Approach with $\log K_{ow}$ .....	35
1.4.2.2	PP-LFER Extension for Ionic Compounds with Experimental Descriptors and ABSOLV .....	38
1.4.2.3	COSMO-RS and COSMOmic.....	41
1.5	Assessing the Toxicity of Ionic Liquids – Application of the Critical Membrane Concentration Approach .....	46
1.5.1	Materials and Methods .....	46
1.5.1.1	Basic Assumptions and Considerations .....	46
1.5.1.2	Compilation of Experimental Data .....	49
1.5.1.2.1	Toxicity Data for Neutral Chemicals.....	49
1.5.1.2.2	Toxicity Data for Ionic Liquids (ILs).....	50
1.5.1.3	Calculation Methods .....	51
1.5.1.3.1	pp-LFER for $K_{mw}$ (neutral).....	51
1.5.1.3.2	COSMOmic for $K_{mw}$ (ion).....	52
1.5.2	Results and Discussion.....	52
1.5.2.1	Reviewing the Toxicity of Neutral Chemicals.....	52
1.5.2.2	Toxicity of ILs.....	55
1.6	Conclusions and Outlook .....	64
1.6.1	Prediction of $K_{lipw}$ (ion) with COSMOmic.....	64
1.6.2	Assessment of Different Models Predicting $K_{lipw}$ (ion) .....	65

---

1.6.3	Applying the Baseline Toxicity Concept on Ions .....	65
1.7	Bibliography .....	67
1.8	Abbreviations .....	74
<b>2.</b>	<b>Abstracts of original publications .....</b>	<b>75</b>
2.1	Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals using the Mechanistic Model COSMOmic .....	75
2.2	Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds .....	76
2.3	Erratum - Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds .....	77
2.4	Assessing the toxicity of ionic liquids – Application of the Critical Membrane Concentration approach.....	78
2.5	Modeling Exposure in the Tox21 in Vitro Bioassays .....	79
2.6	General baseline toxicity QSAR for nonpolar, polar and ionisable chemicals and their mixtures in the bioluminescence inhibition assay with <i>Aliivibrio fischeri</i> .....	80
2.7	Baseline toxicity and ion-trapping models to describe the pH-dependence of bacterial toxicity of pharmaceuticals .....	81
	<b>Eidesstattliche Erklärung .....</b>	<b>83</b>
	<b>Angaben zur Person und zum Bildungsgang.....</b>	<b>84</b>
	<b>Publikationsliste .....</b>	<b>85</b>
	Veröffentlichungen.....	85
	Konferenzbeiträge .....	86
	<b>Danksagung.....</b>	<b>87</b>
	<b>Appendix .....</b>	<b>89</b>
	Supporting Information 1: Prediction of Phospholipid–Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic .....	89
1	Experimental Section.....	89
1.1	Chemicals details.....	89
1.2	Buffer Solution Preparation .....	89
1.3	Liposome Preparation .....	90
1.4	Dialysis cell experiments .....	91
1.5	Data collection anions .....	92

---

1.6 Data collection cations .....	97
2 Theory .....	100
2.1 Estimation of variance.....	100
2.2 Influence of the membrane potential on the $\Delta G$ profiles of the anions .....	101
2.3 Influence of the membrane potential on the $\Delta G$ profiles of the cations.....	110
2.4 Influence of the membrane potential on the relative distribution of the anions.....	114
2.5 Influence of the membrane potential on the relative distribution of the cations.....	123
2.6 Predictions with the different models.....	127
Supporting Information 2: Comparison of different models predicting the phospholipid- membrane water partition coefficients of neutral and charged compounds.....	132
1 Data selection .....	132
1.1 Data collection cations .....	133
1.2 Data collection anions .....	137
1.3 Data collection zwitterions.....	143
1.4 Omitted data .....	144
2 Log $K_{ow}$ based prediction .....	144
2.1 Differences in $K_{lipw}$ between neutral and corresponding ionic species ( $\Delta mw$ ) .....	144
2.2 Why does the log $K_{ow}$ approach predict $K_{lipw}$ for anions better than for cations.....	148
2.3 Ruling out artefacts from KowWIN estimation errors.....	151
2.4 Predicted data .....	153
2.4.1 Cations.....	153
2.4.2 Anions .....	155
3 pp-LFER based prediction.....	159
3.1 Derivation of the solute descriptors of ions from the solute descriptors of neutral compounds .....	159
3.2 Discussion of possible artefacts from Absolv predicted solute descriptors .....	160
3.3 pp-LFER solute descriptors.....	161
4 COSMOmic .....	171
4.1 Modelling details.....	171
4.2 Treatment of cetirizine .....	172
4.3 Calculated data using COSMOmic .....	173
Supporting Information 3: Assessing the toxicity of ionic liquids – Application of the Critical Membrane Concentration approach .....	180
1 Toxicity data for neutral compounds – sorted out data .....	180

---

2 Regression analyses of data from (Vaes et al., 1998).....	183
3 Comparison of pp-LFERs for $\log K_{\text{mem/w}}$ and $\log \text{TLM}$ .....	184
4 ‘Baseline toxicity-QSAR’ based on $c_{\text{mem}}^{\text{tox}}$ .....	184
5 Toxicity of ILs .....	189
Appendix Bibliography .....	201





# 1 Summary: Equilibrium Partitioning of Ionic Organic Chemicals in Environmental Systems: Experiments and Model Predictions

## 1.1 Introduction

There is an increasing interest in risk assessment for ionogenic organic chemicals from different stakeholders such as industry and authorities. Approximately 50% of the nearly 150,000 preregistered compounds under REACH (the registration evaluation authorization and restriction of chemicals regulation of the European Union) are ionogenic, i.e., they are acids, bases or zwitterions (Franco et al., 2010). It is well known that ionized compounds show different thermodynamic properties and environmental behavior than their neutral analogs (Schwarzenbach et al., 2003). Therefore, the knowledge gathered around the description of neutral compounds cannot be transferred one-to-one to the description of charged compounds and considering the scarcity of physicochemical data of ions mechanistic models are of high interest.

This work focuses on the partitioning of ions between phospholipid membrane and water ( $K_{mw}$ ), which is an essential process for various fields of science ranging from biophysics (Honig et al., 1986) to pharmaceuticals (Loidl-Stahlhofen et al., 2001) and environmental sciences (Endo et al., 2011). For the latter field, the  $K_{mw}$  can be considered as a key for understanding several major concerns: first, the  $K_{mw}$  is important for describing the bioaccumulation potential of charged compounds, because they are expected to accumulate mainly in anisotropic sorption phases like membrane lipids (Armitage et al., 2013) and proteins but not in storage lipids (Ng and Hungerbühler, 2013). Membrane lipids constitute e.g., around 70% (v/v) of red blood cells, liver, and kidneys on the dry volume basis and are therefore the major lipid component of several human tissues (Schmitt, 2008). Second, the  $K_{mw}$  is a key element for the description of toxicity: both baseline toxicity (narcosis) (Vaes et

al., 1998), as well as specific toxic modes of action like uncoupling (Escher et al., 1996; Spycher et al., 2008).

Because the value of  $K_{mw}$  is difficult to determine experimentally for real biological membranes, the liposome-water partition coefficient ( $K_{lipw}$ ) is used to approximate  $K_{mw}$  (Escher and Schwarzenbach, 1996). Liposomes are artificial lipid bilayer vesicles (made of phosphatidylcholine in this work) and can be considered the most realistic artificial model system to mimic anisotropic biological membranes (although lacking membrane proteins and cholesterol etc.). This applies for both the  $K_{lipw}$  of neutral and in particular for the  $K_{lipw}$  of ionic species (Escher and Sigg, 2004). As long as the liposomes are in the liquid crystalline state (i.e., above their transition temperature), the  $K_{lipw}$  of different phosphatidylcholines typically vary by only  $\pm 0.2$  log units, which is within the range of the typical experimental error (Endo et al., 2011).

### 1.1.1 What Makes Liposomes Special Compared to Bulk Solvents

Liposomes are artificial lipid bilayer vesicles of defined composition and size and have been used as an experimental system approximating cell membranes since around 1960 (Bangham et al., 1965). They are now well-established because they proved to be easy-to-handle and robust. Typically, liposomes consist of zwitterionic phospholipids with a negatively charged phosphate group and a positively charged choline structure; the former is esterified with two long-chain fatty acids. From their composition and structure, it appears obvious that liposomes are a much more realistic experimental approximation of cell membranes than any bulk organic solvent (Escher and Sigg, 2004; Krämer and Wunderli-Allenspach, 2001; Mouritsen et al., 2001).

There are major differences between liposomes and bulk organic solvents, and with regard to the partitioning of ions, two structural differences are noteworthy: the much larger surface-to-volume ratio of liposomes and the ordered structure which results in an internal dipole potential ( $\Psi_d$ ) of lipid bilayers. As a consequence of the high surface-to-volume ratio

(with a typical mean diameter of 0.27  $\mu\text{m}$  for liposomes) (Olson et al., 1979), sorption of charged species can be electrically neutralized by counterions from the electrolyte solution (diffuse double layer), while bulk media have to maintain electrical neutrality either by the partitioning of ion pairs or by the partitioning of free ions together with counterions. This explains the high sensitivity of measured octanol-water partition coefficient ( $K_{ow}$ ) of an ionic chemical to the ionic strength while  $K_{lipw}$  data show very little ionic-strength dependence (Escher and Sigg, 2004), as further discussed below. The internal dipole potential can be caused by several factors: charge separation in the head groups (this is not a necessary condition as also neutral glycerylmonooleate bilayers reveal a positive membrane dipole (Clarke, 2001)), alignment of dipolar residues of the lipids, and/or oriented water dipoles in the region between the aqueous phases and the hydrocarbon-like interior of the membrane (Clarke, 2001; Wang, 2012). The height of the hill-shaped  $\Psi_d$  in the center of zwitterionic phosphatidylcholine bilayers has been indirectly determined with several experimental approaches and ranges from 227 mV for DPPC bilayers (Wang, 2012) to 280 mV for egg phosphatidylcholine bilayers (Franklin and Cafiso, 1993), positive in the membrane interior.

### **1.1.2 Problems Arising When Octanol is Taken as a Surrogate for Anisotropic Membranes**

Traditionally, membrane affinity is approximated by the partitioning between a bulk organic solvent like octanol and water (Mouritsen et al., 2001), implying that the  $K_{ow}$  correlates well with that of the membrane-water partition coefficient of biological membranes. And indeed, in the case of neutral chemicals,  $\log K_{lipw}$ , and  $\log K_{ow}$  agree fairly well with each other. In the most comprehensive collection of publicly available experimental data,  $\log K_{lipw}$  values of 156 neutral compounds were compared with the respective experimental  $\log K_{ow}$  values, and a correlation coefficient  $R^2$  of 0.95 and a standard deviation of 0.43 log units were observed (Endo et al., 2011). The slope (1.01) and the intercept (0.12) of the regression indicate that the two partition coefficients are generally in agreement,

although mechanistically it is not fully clear why a homogenous solvent phase can emulate a heterogeneous structured lipid bilayer (Endo et al., 2011). One explanation might be that water-saturated octanol is not a completely isotropic phase, but comprises water clusters enclosed by about 16 octanol molecules (Franks et al., 1993). This situation seems to mimic most of the average interaction properties of real lipid bilayer membranes for neutral compounds (Endo et al., 2011).

For charged chemicals, however, the situation is completely different: first, the  $K_{ow}$  values of an ionic species strongly depend on the type and concentration of the counterions present (Escher and Schwarzenbach, 1996; Escher and Sigg, 2004), and second, the  $K_{lipw}$  values of ions are up to several orders of magnitude higher than the respective  $K_{ow}$  values. These differences in the partition behavior are due to the requirement for electroneutrality of both phases. When an ion changes from one bulk phase (e.g., water) to the other (e.g., octanol), it has to be accompanied by a counter ion either as an ion pair or through a separate ion. Depending on the salt concentration (i.e., the concentration of the counter ion), the  $K_{ow}$  of an ionized chemical can therefore differ by more than two orders of magnitude (Escher and Sigg, 2004). Hence, a singly reported  $D_{ow}$  (pH) value (the  $D_{ow}$  is often reported for ionizable chemicals and is the sum of the neutral fraction times the respective  $K_{ow}$  plus the ionized fraction times the respective  $K_{ow}$ ) is of very limited use for any further modelling, especially when no detailed experimental conditions are given (Jafvert et al., 1990; Johnson and Westall, 1990). In contrast,  $K_{lipw}$  values of ions are fairly independent of the salt concentration, because they do not necessarily partition as ion pairs between water and membrane, but are electronically neutralized by counterions located at the membrane water interface (Escher et al., 2000), as also outlined above. Additional limitations to the comparison between  $D_{ow}$  and  $K_{lipw}$  values of ions are imposed by surface active charged chemicals like surfactants (e.g. linear alkyl sulfonates), which accumulate at the octanol-water interface. Furthermore, surfactants can form co-micelles with octanol (even below their own critical micelle

concentration), which increases solubility reciprocally and again leads to a concentration dependent  $D_{ow}$  (Müller et al., 1999; Schwarzenbach et al., 2003).

Given the known discrepancies between  $D_{ow}$  and  $K_{lipw}$  outlined above it seems obvious that experimentally derived  $D_{ow}$  are rather operationally defined values than thermodynamic partition constants and should therefore not be used to model  $K_{lipw}$  values of ions. Nevertheless, a  $\log D_{ow}$  (pH 7) threshold of 4.5 is still used as a screening criterion for potential bioaccumulation for all chemicals in the REACH guidelines (ECHA, 2012), partly due to the lack of other suitable models. However, it has been repeatedly shown that the  $K_{lipw}$  is a more suitable descriptor for bioaccumulation than  $K_{ow}$  or  $D_{ow}$  (Endo et al., 2011; Müller et al., 1999; van der Heijden and Jonker, 2009) and that an experimental  $D_{ow}$  of an ionogenic chemical underestimates the partitioning into real membranes (Avdeef et al., 1998; Escher and Sigg, 2004). Potentially bioaccumulative compounds might therefore not be detected with the current thresholds and better  $K_{lipw}$  prediction models for ions are necessary. Moreover, an inappropriate characterization of bioaccumulation potential of ions has implications for risk assessment, too, when body burdens are estimated on the basis of external concentrations.

Due to problems arising with the operational nature of  $D_{ow}$ , a lot of work has been conducted in the literature that uses the  $K_{ow}$  of neutral chemicals (being a real thermodynamic property). In a first step the  $K_{lipw}$  of the neutral chemicals is modeled and in a second step the  $K_{lipw}$  of the corresponding charged species (Escher and Sigg, 2004). This approach has been used successfully in a number of toxicological studies (Escher et al., 2011; Tang et al., 2013), but it inherently is not feasible in the case of permanently charged organic compounds which do not have a neutral analog to compare with. It is also unclear how, or whether, this approach could be used for polyvalent ions or for zwitterions due to a lack of experimental data that could be used for validation.

### 1.1.3 Difficulties in the Modelling of Anisotropic Membranes

One approach to predict  $K_{lipw}$  based on a molecular description of the membrane is molecular dynamics (MD) simulation of lipid bilayers in the presence of solutes. MD simulations can reproduce a large number of effects and properties related to the membrane-solute interactions and can also yield an internal membrane dipole potential distribution (Ingram et al., 2013; Paloncýová et al., 2014a; Wang, 2012). However, no sufficient number of studies could be found predicting absolute values of  $K_{lipw}$  for lipophilic ions to evaluate the accuracy of predictions based on MD simulation. This may be due to the fact that the computational costs for MD simulations of membranes including a solute at a specific position are extremely high (Paloncýová et al., 2014b). A computationally much more efficient alternative to such MD simulations has been proposed by Klamt et al. in the form of the *COSMOmic* (i.e., COSMO-RS for MICelles) approach (Klamt et al., 2008). *COSMOmic* requires as input the structural composition of a micelle or membrane, usually derived from one or a series of snapshots from a MD simulation of the respective micellar system. The micelle, i.e., a phospholipid membrane in this work, is then virtually split into layers of approximately 1 Å thickness, and the probability to find each of the atoms of the phospholipid and of water in each of the layers is derived from analyzing the MD snapshots. DFT/COSMO calculations are performed in order to yield the surface polarities, i.e., the conductor surface polarization charge densities  $\sigma$  on the molecular and thus also on the atomic surfaces of the phospholipid and water molecules. Combining these with the atom distribution taken from MD simulations leads to a polarity profile, i.e., a  $\sigma$ -profile, for each layer. Then COSMO-RS (i.e., COnductor-like Screening Method for Real Solvents) in its *COSMOtherm* implementation is used in order to derive the affinity of each layer for a certain molecular surface polarity  $\sigma$ , shortly called the  $\sigma$ -potential of each layer. With this information, the free energy of any solute, which is also represented by its DFT/COSMO surface polarization charge densities, can be evaluated at each position and orientation in the membrane system.

An integration over all possible orientations for each position leads to a free energy profile of the solute throughout the membrane system and finally to predictions of the membrane-water partition coefficient. The COSMOmic approach has recently been demonstrated by two independent groups to yield results of comparable, if not slightly superior, quality, with respect to the distribution of neutral solutes in phospholipid membrane systems (Ingram et al., 2013; Jakobtorweihen et al., 2014; Paloncýová et al., 2014a), but at computational costs which are several orders of magnitude lower than for the respective MD simulations (using the CHARMM (Ingram et al., 2013; Jakobtorweihen et al., 2014) and Berger (Ingram et al., 2013; Paloncýová et al., 2014a) lipid force field). While, by 2014, the calculation of a free energy profile with COSMOmic takes a few minutes on a single core (given that all input files are ready to use), the same calculation conducted as MD simulation would take 15 to 48 h on supercomputers with more than 100 cores (Jakobtorweihen et al., 2014). COSMOmic has previously been used tentatively to predict partition coefficients and free energy profiles of anions (Spycher et al., 2008). For the studied 35 anions a reasonably low root-mean-square error (RMSE) was observed, but it was necessary to empirically fit the predicted values to experimental data, as apparently some relevant mechanism for the prediction of ions was not accounted for yet (Spycher et al., 2008).

#### **1.1.4 Environmentally Relevant Application of $K_{lipw}$ Predictions**

The high number of ionizable or even permanently charged organic chemicals potentially released into the environment is a challenge for ecotoxicology (Franco et al., 2010). For neutral chemicals the minimal level of nonspecific toxicity is referred to as narcosis or, in the field of environmental science, baseline toxicity (Escher and Schwarzenbach, 2002; Wezel and Opperhuizen, 1995). The baseline toxicity concept states that nonspecific toxicity occurs at a consistent range of membrane concentrations, independent of both the chemical as well as the (aquatic) organism, although the exact mechanisms is not yet fully clarified. Underlining the non-specificity, baseline toxicity was found to act via concentration addition for mixtures



(Deneer et al., 1988). It is likely that the chemicals sorbing to the membrane change its properties, e.g. its fluidity and permeability, to such a degree that its (biological) function is disturbed (Wezel and Opperhuizen, 1995). A different theory explains baseline toxicity via specific interactions of molecules with sensitive proteins in the central nervous system (Franks and Lieb, 1990). However, it was also demonstrated that baseline toxicants accelerate the decay of the membrane potential after a very short pulse of light that induced a certain membrane potential in an isolated photosynthetic membrane vesicle originating from a photosynthetic bacterium (Escher et al., 2002), which rather supports the explanation of baseline toxicity by non-specific disturbance of the membrane structure and functioning.

Vaes et al. showed that there is no difference in the baseline toxicity between polar and apolar neutral organic chemicals, when the  $K_{lipw}$  is used as a descriptor instead of the  $K_{ow}$  (Vaes et al., 1998). This finding has been corroborated later (Escher and Hermens, 2002; Escher and Schwarzenbach, 2002), albeit with a relatively limited set of chemicals. As outlined above, the  $K_{lipw}$  acts as a surrogate for the (biological) membrane-water partition coefficient  $K_{mw}$ . More recent studies substantiate these earlier findings (Endo, 2016; Escher et al., 2017; McCarty et al., 2013). Along the same line of thoughts are also earlier concepts like the critical body residue concept (Endo, 2016; McCarty and Mackay, 1993), or the target lipid model (TLM), based on a critical body burden (Kipka and Di Toro, 2009).

A different access to explain non-specific toxicity has been put forward recently with the activity approach (Thomas et al., 2015), which has been criticized (Goss and Endo, 2016) – partly because it is intrinsically not applicable to ionic chemicals. Moreover, the assumption of a critical membrane threshold concentration is principally not consistent with a critical membrane activity. Both concepts cannot be correct at the same time.

## 1.2 Objective of this Work

This work had three main goals as follows:

### 1.2.1 Calibrating COSMOmic for the Use with Ions

In order to calibrate COSMOmic for the use with ions an exhaustive compilation of published  $K_{lipw}$  data of organic anions and cations was aspired as well as an appraisal of the of the existing data's quality. In order to systematically increase data diversity, own measurements were conducted, leading to a more thorough validation of the modelling approaches. The goal was to apply the existing COSMOmic to the available  $K_{lipw}$  data, identify the areas where adaptations are needed, and refine the model in accordance. Accordingly, a membrane potential was newly implemented in COSMOmic to achieve an improved computation of interaction energy between phospholipid membrane and ions. Finally, the refined COSMOmic model was used for calculations of  $K_{lipw}$  of anionic, cationic and neutral species to evaluate the performance of the model.

### 1.2.2 Comparison with Two Other Models Predicting $K_{lipw}$

Next, the refined COSMOmic model was compared with two other models for the prediction of  $K_{lipw}$  of charged compounds, namely an empirical correlation approach based on  $K_{ow}$  and the pp-LFER approach (i.e., polyparameter linear free energy relationship). The focus is on the description of charged compounds, but in order to scrutinize also model consistency, the prediction quality of the different models for neutral chemicals is as well included in the discussion. In order to address the paucity of experimentally derived  $K_{lipw}$  values, it is not only assessed which model gives the best results, but it is also discussed which model complexity might be most suitable depending on the prediction accuracy needed.

### 1.2.3 Application of Predicted Partition Coefficients on Baseline Toxicity Concept for Ions

Finally, the  $K_{lipw}$  values predicted by COSMOmic were used to address the question whether the baseline toxicity concept is also applicable for ionic chemicals. In order to prevent additional complexity in the modelling of toxicity for ions such as ion trapping (further discussed below) the focus was on permanently charged ionic liquids (ILs). In the literature a multitude of quantitative structure property relationship (QSPR) models can be found that describe IL toxicity for different species (Thuy Pham et al., 2010), but after extensive literature search no QSPR could be found that is strictly examining whether the experimental toxicity data of ILs can be described as baseline toxicity. This might well be due to the fact that the partitioning of organic ions to membranes could not be reliably predicted before the model improvements presented in this work.

Further the baseline toxicity concept was used to shed light on those ILs that are most likely specifically acting toxicants and those that are prone to experimental artefacts. The critical membrane concentration of roughly 100 mmol/kg (membrane lipid), which is the fundament of this investigation, is well known in the literature but tested only for a limited number of organisms (e.g., (Escher and Schwarzenbach, 2002) and literature cited above). Therefore the investigation is started out by resuming the baseline toxicity for neutral chemicals, in order to ensure the broad applicability domain of the concept before applying it to charged chemicals. This exercise should help to assess the uncertainties within the baseline toxicity concept and to interpret the results when the concept is expanded to ILs.

### **1.3 Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals using the Mechanistic Model COSMOmic**

#### **1.3.1 Materials and Methods**

##### **1.3.1.1 Determination of Membrane-Water Distribution Coefficients**

All liposome-water partition coefficients were determined at 20 to 22°C via equilibrium dialysis experiments and HPLC analysis. The experimental details are described in the supplementary information, SI-1, and thus the description here is brief. Salt concentration (100 mM KCl) was constant for all experiments. Buffer (MOPS,  $pK_a=7.2$ , or CHES,  $pK_a=9.3$ ) was chosen so that the pH in the experiments was at least 3 pH units higher than the  $pK_a$  of the investigated chemicals (to be sure that only the anionic species is considered). POPC liposomes were prepared with a membrane extruder (Lipex Biomembranes, Vancouver, BC, Canada, with Whatman polycarbonate filter membrane, pore size 0.1  $\mu\text{m}$ ) as described elsewhere (Kaiser and Escher, 2006). Custom-made glass dialysis cells consist of two chambers that were separated by a dialysis membrane made of regenerated cellulose with a cutoff of 10,000 to 20,000 Da (Thomapor, Reichelt Chemie Technik, Heidelberg). One chamber was filled with buffer solution and the other with liposome suspension. The latter received the test anion. The liposome free side of every dialysis cell was sampled twice (i.e., on the fourth and sixth day), and the samples were subjected to the HPLC analysis. Mass recovery of every chemical was tested accordingly in control experiments without liposome where both dialysis chambers were filled with buffer solution (revealing that losses were less than 5%). Each dialysis cell experiment was conducted at least in triplicates. All experiments were conducted with a liposome load below 0.08 mol(substance)/mol(lipid), which has been shown to be within the linear part of the sorption isotherm (Escher et al., 2000). In order to cross-check the experimental setup for consistency,  $K_{lipw}$  of 2,3,4,6-tetrachlorophenol was measured. The anionic species of this chemical has a reported  $\log K_{lipw}$  of 3.46 (Escher et al., 2000), while in this work a  $\log K_{lipw}$  of 3.52 was determined.

### 1.3.1.2 Data Collection and Evaluation

All data collected from the literature were measured with phosphatidylcholine liposome. Overall neutrality of the phosphatidylcholine membrane as a sorption phase is important to note, since a charged membrane would have significant impact on the sorption of charged chemicals (Thomae et al., 2007). The experiments in the literature have been conducted at different ionic strengths, which should not be crucial for the modelling of ion partitioning since the ionic strength does not significantly influence the membrane partition coefficients of ionic compounds (Escher and Sigg, 2004). Only partition coefficients measured above the main phase transition temperature of the membrane were considered, ensuring that the membrane is in its natural condition, the liquid crystalline state. The state of the membrane has been shown to be an essential parameter for the partition coefficient of neutral chemicals (van Wezel et al., 1996).

All experiments considered here were conducted with unilamellar vesicles, preferably using the equilibrium dialysis method, but also other experimental methods are considered (see SI-1 for details). This results in a total of 51 experimental values for anions (from which five are own measurements of this work) and 24 experimental values for cations. When multiple experimental data for the same ion were found, the arithmetic mean of the  $\log K_{lipw}$  values was used (see SI-1). The difference of the single reported values from the corresponding mean value was between 0.02 and 0.21 log units for the anions (with a total of six repeatedly measured ions) and 0.03 to 0.28 log units for the cations (with a total of four repeatedly measured ions), with the exception of two reported values for atenolol that differ by 0.50 log units from their mean.

### 1.3.2 Theoretical section

Credit has to be given mainly to Simon Spycher for the idea that the unsatisfying  $K_{lipw}$  predictions (further discussed below) of ions by the originally published version of COSMOmic (Klamt et al., 2008) is due to the missing membrane potential. Andreas Klamt

and Uwe Huniar deduced how the membrane potential could best be integrated in the model and Larissa Pohler programmed the optimization algorithm deriving the adjustable parameters for the membrane potential and implemented this into the new version of COSMOmic as described in the following subsections.

### 1.3.2.1 COSMO-RS and COSMOmic

To run COSMOmic, as outlined in the Introduction, a detailed membrane structure is required which is taken out of MD simulations. Care was taken to use time averaged atomic distributions which are furthermore centered in the middle of the simulation box (Jakobtorweihen et al., 2013). The atom distributions were kindly simulated (CHARMM36 force field) and provided by Sven Jakobtorweihen (Jakobtorweihen et al., 2013). In addition, TZVP (Becke, 1988; Eichkorn et al., 1995; Perdew, 1986; Schäfer et al., 1994) cosmo files are needed for all involved relevant conformers of all the solutes in the partitioning process. Therefore, COSMOconfX13 (version 3.0, COSMOlogic) templates, based on Turbomole version 6.5 (Ahlich et al., 2015), have been used for full energy minimization and conformer generation (Vainio and Johnson, 2007). Each molecule has at least one and a maximum of 10 conformers (the investigated molecules in this work had on average 3.14 conformers).

### 1.3.2.2 Estimation of Membrane Potentials

In addition to the depiction of the membrane anisotropy, the membrane dipole potential may need to be described in the model, when the model is used for the prediction of ions. There is no direct experimental method to measure the membrane dipole potential, but several indirect approaches allow the quantification of  $\Psi_d$  for different types of bilayers. For egg phosphatidylcholine bilayer vesicles the values in the membrane interior range from 0.24 V (deduced with a combination of kinetic and binding data of lipophilic ions) (Flewelling and Hubbell, 1986a) to 0.28 V (electron paramagnetic resonance spectroscopy in combination

with nitroxide spin-labeled hydrophobic ions) (Franklin and Cafiso, 1993). For DPPC bilayers (dipalmitoylphosphatidylcholine), two values of 0.227 and 0.24 V are given in the review of Wang (Wang, 2012).

We initially tried to derive a profile of  $\Psi_d$  from the available MD simulations. Using the fundamental equations of electrostatics, the electrostatic potential of a planar membrane can either be derived from the charge density in each layer:

$$\Psi_d^c(z) = \frac{1}{4\pi\epsilon_0} \int_0^z du \int_0^u dv \rho(v) \quad (1)$$

where  $\rho(v)$  is the charge density (charge per area) in layer  $v$ , or from the dipole moment density:

$$\Psi_d^c(z) = \frac{1}{4\pi\epsilon_0} \int_0^z du D(u) \quad (2)$$

where  $D(u)$  is the  $z$ -component of the dipole moment density in membrane layer  $u$ . Here  $\Psi_d^c(z)$  is the value of the dipole potential with respect to the center of the membrane. The default definition of the dipole potential  $\Psi_d(z)$  with respect to the bulk water phase can easily be found from  $\Psi_d(z) = \Psi_d^c(z) - \Psi_d^c(\infty)$ . Using Eq. 1 together with the DMPC (1,2-dimyristoyl-sn-glycero-3-phosphocholine) snapshot from (Gurtovenko et al., 2004), which was used for predicting  $K_{lipw}$  of neutral chemicals previously (Klamt et al., 2008), and with the partial charges applied in the respective MD simulation results in a potential which is by 0.99 V higher at the center of the membrane than in water. This value is too high by more than a factor 3 compared to the experimentally expected value of  $\sim 0.3$  V. The value of 0.99 V originates from a contribution of  $-4.55$  V caused by the charges on the DMPC atoms and an overcompensation of  $5.55$  V by the water molecules. Thus, the MD-derived membrane potential is the difference of two large numbers and is highly sensitive to any inaccuracy in the potential contributions from both molecules. It is furthermore surprising and counterintuitive that the net potential is opposite to the potential generated by the phospholipid molecules themselves. If we calculate the membrane potential using partial

atomic charges from BP-TZVP-COSMO files instead of the charges used in the MD force field, then the water molecules produce a potential of 5.2 V, i.e., very similar to the result from the MD charges, but the DMPC contribution is only  $-0.6$  V, compared to  $-4.55$  V using the MD charges. This demonstrates that two plausible representations of the electrostatics of the phospholipid molecules can result in very large differences in the membrane potential, suggesting general difficulty to obtain a precise consensus for the potential distribution from MD simulations.

Wang reviewed the membrane potentials from 10 different MD simulations of phospholipid bilayers reported in the literature, all using different combinations of force-fields, partial charges, and electrostatic summation techniques (Wang, 2012). The three DMPC simulations had values of 0.9, 0.9, and 0.77 V, while all results (including diphytanoyl-, dipalmitoyl-, and diphytanylphosphatidylcholine) range from 0.3 to 1.0 V, with a mean value of 0.7 V. This means that, even if optimistically analyzed, the variability of the membrane potentials derived from MD simulations is at least 0.2 V, and they seem to have the tendency to be about 0.4 V higher than the experimental value of  $\sim 0.3$  V (Clarke, 2001). Wang mentions one promising MD simulation using polarizable force fields (Chowdhary et al., 2013; Harder et al., 2009), yielding dipole potentials closer to the experimental estimate, but it is at present not clear whether such force fields are generally more accurate.

At this point, it may be worth noting the theoretical maximum of the membrane potential that a DMPC double layer with a typical density of one DMPC molecule per  $33 \text{ \AA}^3$  would produce, if all zwitterionic dipoles pointed outward. Simple calculus yields a dipole moment of 25 D for the stretched zwitterion and that would yield a membrane potential of roughly  $-17$  V. In addition, each DMPC molecule has the dipole moments of the two ester groups, each being in the order of 2.5 D, which could add positively or negatively to the zwitterion dipoles. Obviously, such an arrangement would have a completely unrealistic high electrostatic energy, and thus the nature and the thermodynamic equilibration of the MD



simulation always care for a strong reduction of the average net dipole in DMPC membranes. As a result, the zwitterionic dipoles seem to be orientated essentially parallel to the surface. Given the enormous maximum value of 17 V, it is already a remarkable achievement that the different MD simulations seem to agree within 0.2 V in their calculations of the membrane potential, i.e., within  $\sim 1.2\%$  of the theoretical maximum value, and that they deviate from the experimental values only by about 0.4 to 0.7 V, i.e., by 2.5% to 4%.

Nevertheless, despite this remarkable success, an error of 0.4 V still causes a shift of the energy of a singly charged ion by almost 10 kcal/mol and thus a shift of  $K_{\text{lipw}}$  by almost 7 log units. This means that, for the prediction of  $K_{\text{lipw}}$  with a desired minimum accuracy of say 0.7 log units (which is a typical RMSE of COSMO-RS predictions for homogeneous phases) (Stenzel et al., 2014), 4 even 10 times higher accuracy in the description of the potential are needed; i.e., only tolerate errors of 0.04 V can be tolerated, which corresponds to 0.25% of the theoretical maximum of the DMPC potential.

### 1.3.2.3 Optimization of a Model Membrane Potential

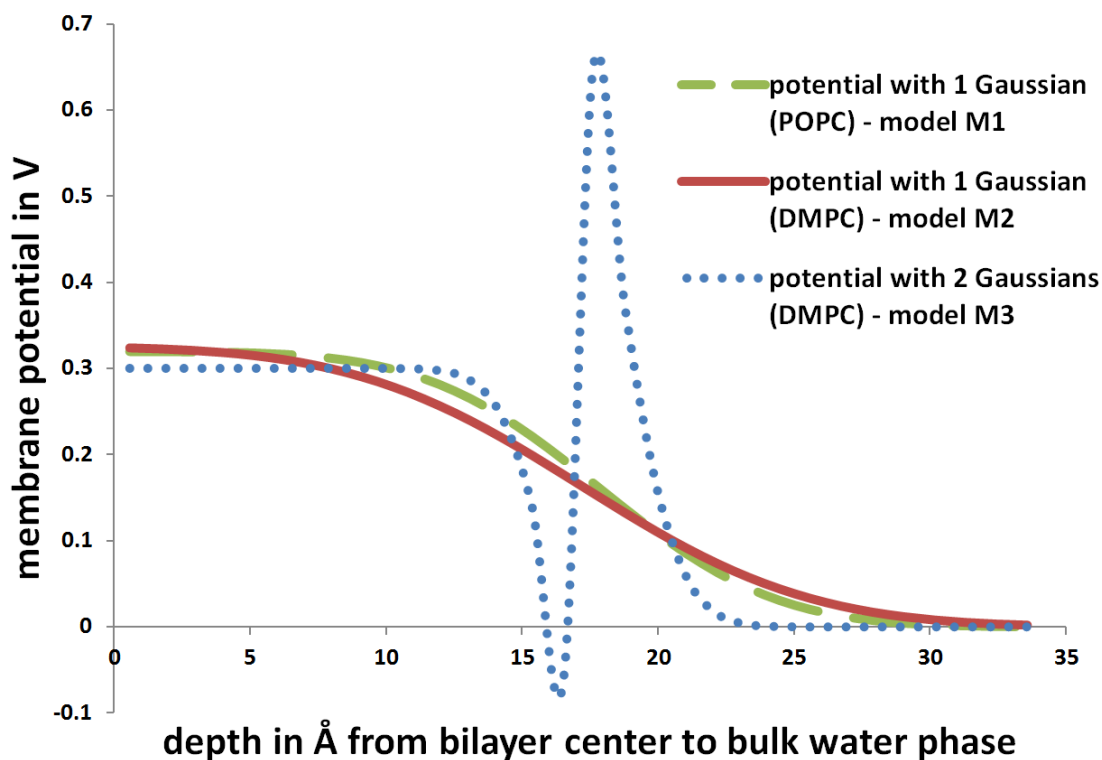
Given the large uncertainties involved in the derivation of the membrane dipole potential from MD simulations, we decided to use an empirical model potential with a small number of adjustable parameters. In order to achieve a physically most plausible shape of the membrane potential, we assume that the net dipole density of the membrane and of water can be represented by one or two Gauss distributions. As a result, the shape of the model potential (being the integral over dipole densities) is either a Gaussian-type error function or the sum of two such functions. Each Gaussian has three adjustable parameters: the height  $h$ , the center position  $p$ , and the width  $w$ .

$$D(u) = h \exp\left\{-\left(\frac{x-p}{w}\right)^2\right\} \quad (3)$$

In the optimization algorithm, we iteratively searched for the values of these three or six parameters ( $h$ ,  $p$  and  $w$ ) that minimize the error in the prediction of the 75 ionic and 161

neutral  $K_{lipw}$  values. As shown in Fig. 1, this approach leads to a plateau value in the membrane center when using one Gaussian distribution (red and green curves). With two Gaussian distributions this approach even allows for membrane potentials which may have a minimum or maximum in the head group region (dotted blue curve).

The robustness of the empirical fitting approach for adjusting the internal membrane potential  $\Psi_d$  with experimental  $K_{lipw}$  values was evaluated by dividing the data set into a training and a test set. Both ionic and neutral chemicals were ordered by their  $K_{lipw}$  within chemical classes and then distributed roughly in a ratio of 2:1. Ionic chemicals were categorized into a training set of 50 and a test set of 25 compounds (see SI-1), neutral compounds into a training set of 105, and a test set of 56 compounds (Klamt et al., 2008). The 161 neutral  $K_{lipw}$  values from the original publication (Klamt et al., 2008) were included in the process of adjusting  $\Psi_d$ , in order to check whether the adaptations of the model made with the addition of a membrane dipole affect the prediction accuracy for neutral compounds. The experimental values are averaged experimental  $K_{lipw}$  for temperatures up to 40°C and were taken from (Endo et al., 2011). They should therefore better represent the currently available data than the ones used in the original COSMOmic publication (Klamt et al., 2008).



**Figure 1.** Three different profiles of the adjustable membrane potential for two different bilayers (DMPC and POPC) resulting with different height  $h$ , position  $p$ , and width  $w$  parameters of one (solid red and dashed green) and two (dotted blue) Gaussian-type dipole moment distributions. The depth is given along the membrane normal, starting in the membrane interior.

### 1.3.3 Results and Discussion

#### 1.3.3.1 Predicting $K_{lipw}$ Using COSMO-RS with Phosphatidylcholine Lipid as Bulk Solvent

For a first comparison between isotropic and anisotropic solvents as depicted in COSMOmic, the partitioning from water to DMPC as bulk solvent were calculated. As shown before (Klamt et al., 2008), simple neutral compounds are surprisingly well predicted by considering phospholipid as bulk solvent (RMSE = 0.70 and  $n = 161$ ). Only for some bifunctional chemicals the consideration of the bilayered structure plays a decisive role (Endo et al., 2011). In contrast to these findings for neutral chemicals, the prediction of anions into bulk DMPC solvent in this work is 2.4 to 15.7 log units lower than the experimental  $K_{lipw}$  values (resulting in RMSE = 9.51 and  $n = 51$ ), while the cations are predicted 4.3 to 9.6 log

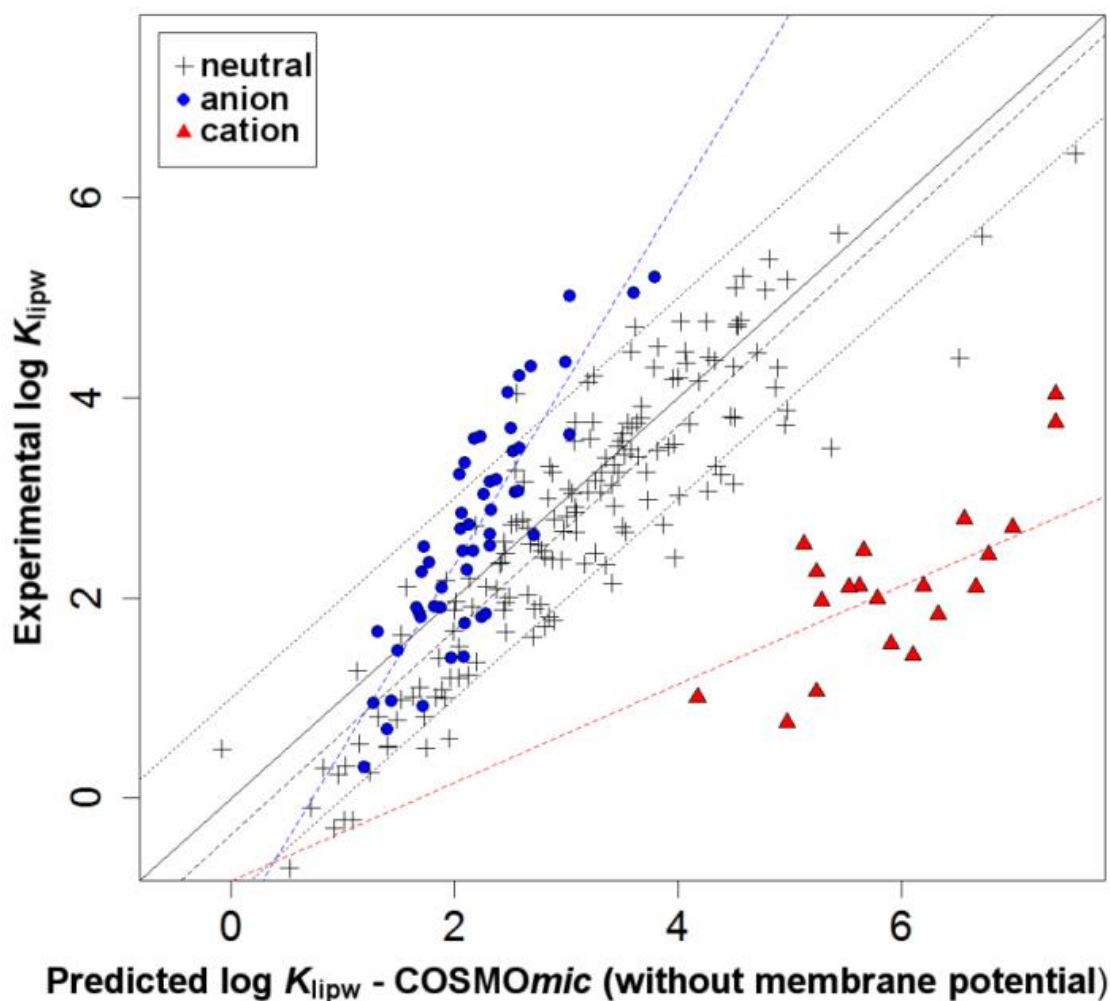
units too high (RMSE = 6.22,  $n = 24$ ). Using a bulk solvent of POPC lipid for the prediction gives the same picture.

As expected, bulk solvent lipids are not anywhere near to being an appropriate model for membranes when it comes to the prediction of ion partitioning. While anisotropy can be neglected for most neutral compounds – being one reason for the good correlation between their  $K_{ow}$  and  $K_{lipw}$  – this simplification is not suitable for ions. Here, it seems that the orientation and location in the membrane are of crucial importance when it comes to the description of the sorption behavior of ions.

### 1.3.3.2 Predicting $K_{lipw}$ using COSMOmic without considering the membrane potential

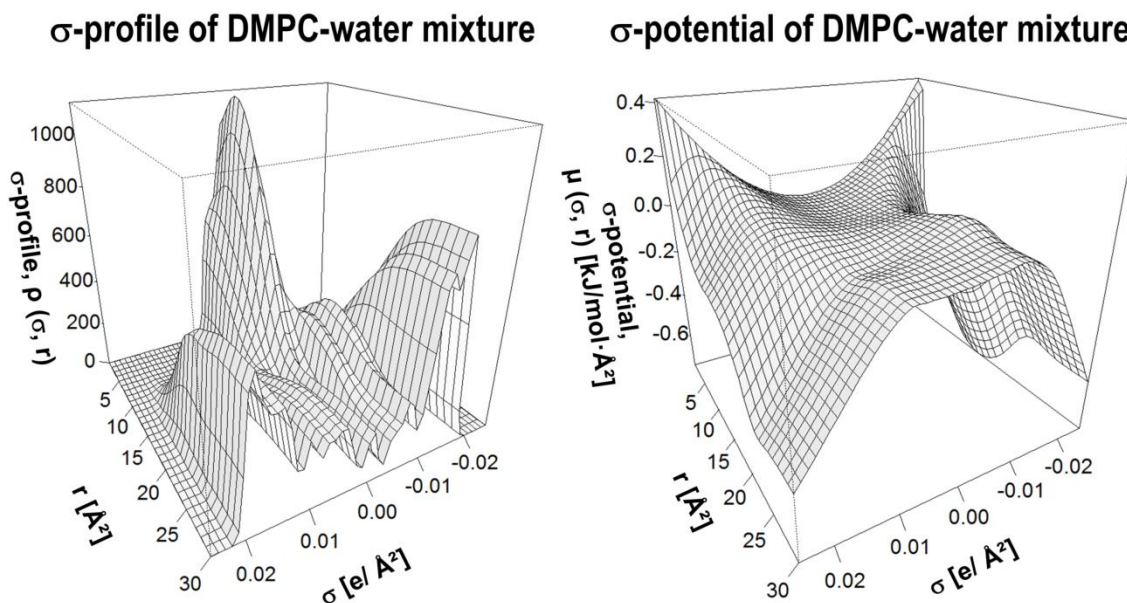
#### $\Psi_d$

In a next step, the partition coefficients were calculated using COSMOmic but without accounting for the internal dipole potential. Calculations were made using trajectory averaged membrane structures over 80 ns simulation time with 128 DMPC and 3919 water molecules (which corresponds to a mole fraction of 0.032 and 0.968, respectively) (Jakobtorweihen et al., 2013). The simulation box was split into 30 layers giving a resolution of 1.13 Å for each layer. For every chemical 162 different orientations in each layer were calculated (applying more orientations changes the calculated  $\log K_{lipw}$  values only insignificantly: with 1082 orientations the maximum deviation is 0.006%). Predicting  $K_{lipw}$  of the above introduced 51 anions, 24 cations, and 161 neutral compounds using COSMOmic as introduced previously (Klamt et al., 2008) – i.e., not considering the membrane potential – results in Fig. 2. The calculation without considering the membrane dipole potential leads to big systematic deviations for ionic chemicals but not for neutral chemicals.



**Figure 2.** Experimental  $K_{lipw}$  of 161 neutral (black crosses), 51 anionic (blue circles), and 24 cationic (red triangles) chemicals into a DMPC membrane against predicted values. The COSMOmic calculation here does not consider the membrane potential. The identity line (1:1 line) is indicated as solid line; deviations of  $\pm 1$  log units are shown as dotted lines. For the dashed regression lines for neutral compounds, anions, and cations a least-squares regression has been used.

Plot of the  $\sigma$ -profile and the  $\sigma$ -potential reveal the anisotropic nature of the DMPC bilayer used for the calculations as shown in Fig. 3. The probability distribution for the headgroup phosphorous and nitrogen atoms peaks at a distance of 18.7 and 19.8 Å, respectively, from the bilayer center, while the outermost bulk water layer is at 33.4 Å.



**Figure 3.** The  $\sigma$ -profile (left) and  $\sigma$ -potential (right) of the DMPC-water system used for models M2, 2a, and 3 as summarized in Table 2. These figures show the slicing of the membrane into consecutive liquids as done in COSMOmic (here no membrane dipole potential is additionally accounted for yet). It can be seen how the DMPC lipids span from the first layer (representing the membrane bilayer center) to layer 27 (at 30 Å), where the bulk water phase begins. Each layer has a thickness of 1.13 Å.

The neutral chemicals are as well predicted as expected. A linear equation of the regression line appears as follows:

$$\log K_{\text{lipw}}(\text{exp}) = 1.02 (\pm 0.04) * \log K_{\text{DMPC/w}}(\text{calc}) - 0.37 (\pm 0.13); \text{RMSE} = 0.70, n = 161$$

Note that assuming errors in both experimental and calculated values in the regression analysis would result in slightly different slopes and intercepts. The predictions of the ions give a more heterogeneous picture. While all of the  $K_{\text{lipw}}$  values for cations are 0.9 to 2.3 log units overestimated, most of the  $K_{\text{lipw}}$  values for anions are underestimated (up to 1.9 log units) for the DMPC membrane shown in Fig. 2. Using the POPC membrane yields the same result with marginally different numbers. Fitting a least-squares regression through both differently charged groups separately gives the following equations for the DMPC membrane:

$$\log K_{\text{lipw}}(\text{exp}) = 0.49 (\pm 0.12) * \log K_{\text{DMPC/w}}(\text{calc, cation}) - 0.83 (\pm 0.76); \text{RMSE} = 0.68, n = 24$$

$$\log K_{\text{lipw}}(\text{exp}) = 1.84 (\pm 0.15) * \log K_{\text{DMPC/w}}(\text{calc, anion}) - 1.34 (\pm 0.33); \text{RMSE} = 0.55, n = 51$$

Here, the RMSEs are given with respect to the regression lines. One could use the regression equations for semiempirical predictions as it has been done previously (Spycher et al., 2008). However, this would not be a satisfying approach, especially when considering the initial aspiration for a mechanistic model that is not limited to any kind of compound class or charge. Fortunately, the improvements presented in the next chapter render a purely empirical fit based on a simple regression equation unnecessary.

### 1.3.3.3 Using COSMOmic with an Optimized Membrane Potential $\Psi_d$

Empirical membrane potentials have been optimized as outlined above for different membrane types (DMPC and POPC) and different salt concentrations (0 and 0.1 M KCl). The center positions, heights, and widths as defined in Eq. 3 of the resulting Gauss curves are summarized in Table 1.

**Table 1.** Comparison of position, width and height of  $\Psi_d$  derived for different membrane structures based on time-averaged atom distributions. Center position  $p$ , height  $h$ , and width  $w$  are the three adjustable parameters in the Gauss-type error function as defined above. For each model given here, all 161  $K_{lipw}$  values for neutral and 75 values for ionic compounds have been used, except for model M2a, which has its potential optimized based on 56 neutral and 27 ionic  $K_{lipw}$  values.

number	model	center position $p$ [Å]	height $h$ [mV]	width $w$ [Å]
M1	POPC (1 Gauss)	17.891	320	7.138
M2	DMPC (1 Gauss)	17.080	326	8.866
M2a	DMPC training (1 Gauss)	15.948	357	9.332
M3	DMPC (2 Gauss curves)	Pos1: 17.131 Pos2: 17.663	Height1: -996 Height2: 1296	Width1: 0.198 Width2: 2.813
M4	DMPC 0.1 M KCl (1 Gauss)	16.258	340	10.796

There are only marginal differences in height and position of  $\Psi_d$  for all optimization runs with one Gaussian. The width differs slightly for the MD simulation including salt (0.1 M KCl), but this hardly has an influence on the predictive power as shown in Table 2. For the DMPC membrane, two different potentials have been optimized based on one and two

Gaussian distributions (i.e., three and six adjustable parameters, respectively). The double Gaussian model did perform only marginally better than the single Gaussian distribution. Furthermore, the extreme fluctuations of the dipole potential fitted based on two Gaussians appear to be rather unlikely (see Fig. 1). Hence, the single Gaussian model is taken as default in the following sections.

**Table 2.** Comparison of the calculation of  $\log K_{lipw}$  values for neutral and ionic compounds with different membranes, salt concentrations, forms of potential distribution (1 or 2 Gaussians) and different data sets underlying the potential optimization. All models are based on the optimization of one Gauss curve for the membrane potential, except model M3. Slope and intercept are given with the respective standard errors and are derived with a least-squares regression for neutral and ionic compounds together for the regression equation  $\log K_{lipw}(\text{experimental}) = \text{slope} * \log K_{lipw}(\text{calculated}) + \text{intercept}$ . The offset describes the nonweighted average of predicted minus experimental values for the calculated ionic and neutral species. The RMSE is obtained separately for n neutral and ionic chemicals after subtracting the offset from the calculated value.

num- ber	model	slope	intercept	offset	n	RMSE (ions)	RMSE (neur.)
M1	POPC	$0.94 \pm 0.04$	$-0.11 \pm 0.11$	$0.30 \pm 0.04$	236	0.71 (n=75)	0.63 (n=161)
M2	DMPC	$0.96 \pm 0.04$	$-0.21 \pm 0.12$	$0.32 \pm 0.04$	236	0.70 (n=75)	0.62 (n=161)
M2a	DMPC training set	$1.04 \pm 0.06$	$-0.35 \pm 0.19$	$0.25 \pm 0.07$	83	0.68 (n=27)	0.59 (n=56)
M3	DMPC (2 Gauss curves)	$0.97 \pm 0.03$	$-0.34 \pm 0.11$	$0.43 \pm 0.04$	236	0.66 (n=75)	0.60 (n=161)
M4	DMPC 0.1 M KCl	$0.96 \pm 0.04$	$-0.24 \pm 0.12$	$0.37 \pm 0.04$	236	0.71 (n=75)	0.63 (n=161)

To further evaluate the dependence of the potential optimization on the selection of chemicals, the 75 ionic and 161 neutral species were divided into a training and test set (see SI-1) as described above. The potential was optimized for the same DMPC membrane as in model M2, but for model M2a the optimized Gaussian potential is based only on the training set. The performance of the resulting model M2a has been tested with the chemicals of the test set, in order to evaluate how sensitive the predictions are in respect to the data set used for deriving the potential curves (see Table 2). Although there are slight differences in the model M2 and model M2a potentials, the predictions of  $K_{lipw}$  values differ less than 0.3 log units

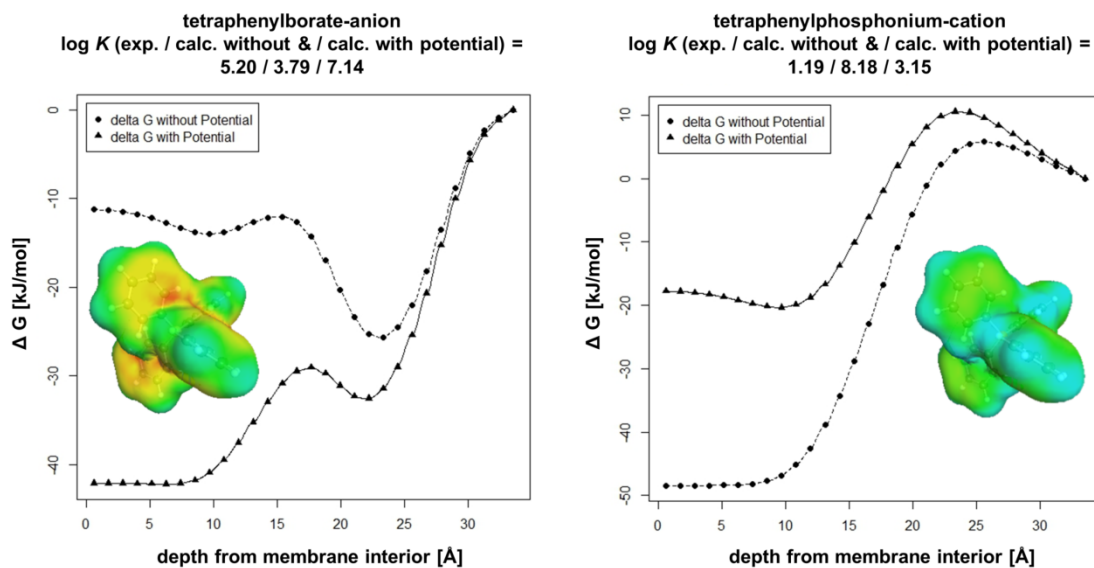


between the two models for the 27 test chemicals. The RMSEs and also the slopes and intercepts have very similar values within the range of error, which indicates that the chosen approach results in robust predictions despite slightly different shapes of potentials. However, the use of the model M2a potential is not recommended because a potential optimization based on all available experimental data should yield the most reliable potential shape.

On average, COSMOmic predicts the  $K_{lipw}$  values roughly 0.3 log units too high, as shown by the different offset values in Table 2. It is important to note that during the potential optimization procedure this systematic overprediction has not been minimized. The offset is significantly different from zero (two-tailed P values are not bigger than 0.0005 for the different models) and might be explained by remaining simplifications in the model like the assumption of structureless liquids for each of the membrane slices. Also, a possible contribution of the membrane deformation energy caused by the sorbing solutes is not considered. Accounting for such kind of ‘volume work’ would make the partitioning into the membrane less favorable and, therefore, reduce the absolute value of the offset. First attempts using an elastic term as introduced previously (Klamt et al., 2008) showed this trend at the cost of an increased scatter in the prediction, indicating that the empirical expression for the deformation energy should be reinvestigated in further refinement. Up to this end, it can be assessed that the offset is fairly constant for different membranes (see Table 2) as well as for differently charged species. For model M2, for example, the predictions of the neutral species have an offset of 0.30, the anions of 0.37, and the cations of 0.40 log units, resulting in an average of 0.32 log units. Thus, the RMSE in the predictions can be decreased by simply subtracting the average offset from the predicted  $K_{lipw}$  values as done in Table 2. The RMSEs of the ions were reduced by 0.09 to 0.13 log units by subtracting the offset values, except for model M3, having its RMSE reduced by 0.17 log units. Considering the remaining simplifications in COSMOmic as discussed above, the average overprediction of  $K_{lipw}$  as expressed in the offsets appears to be rather small.

The membrane potentials optimized for the DMPC and POPC membrane of course have a slightly different shape (Table 1) but lead to the same quality in the prediction (Table 2). This is in accordance with experimental results, which do not show significant differences in the sorption behavior of DMPC and POPC membranes either (Endo et al., 2011). Similarly, the inclusion of a 0.1 M KCl concentration in the DMPC-water system (model M4) does not result in a big difference of the derived membrane potential and partition coefficients. It has experimentally been demonstrated that different salt concentrations (0.001 – 0.1 M KCl) have only marginal influence on  $K_{lipw}$  of ions (Escher and Sigg, 2004).

A good example of the influence of the membrane potential on the Gibbs free energy profiles and resulting calculated  $K_{lipw}$  is given by the experimental and calculated sorption behavior of the two oppositely charged tetraphenyl analogs TPB and TPP (Fig. 4). Although the negatively charged TPB is structurally very similar to the positively charged TPP,  $K_{lipw}$  of TPB is 4 orders of magnitude higher than that of TPP (Demura et al., 1987; Flewelling and Hubbell, 1986b). Deviating almost exclusively by the sign of the surface charge (but not the charge density), this difference can only be explained by the influence of the membrane potential. The resulting attractive interactions between the positive inner potential of the membrane and negative TPB are reflected by a descending calculated  $\Delta G$  profile. In contrast, the inclusion of the repulsive interactions between membrane potential and the positively charged TPP elevate the calculated  $\Delta G$  profile.



**Figure 4.** Influence of the membrane potential on the  $\Delta G$  profiles of TPB and TPP in the DMPC membrane (model M2). Experimental data are from (Flewelling and Hubbell, 1986b) and (Demura et al., 1987).  $\Delta G$  equals zero in the bulk water phase.

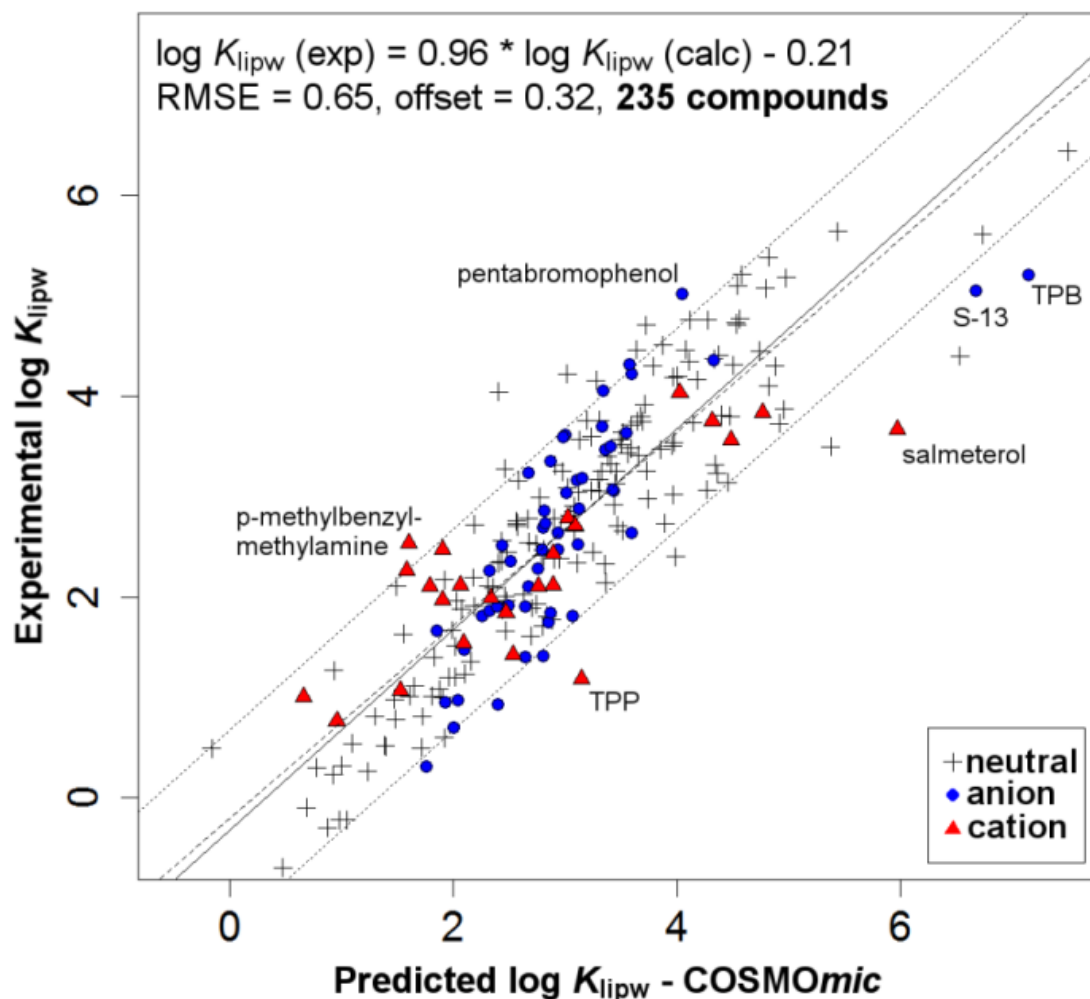
For model M2, on average, the  $\Delta G$  values at the membrane center are 30.63 kJ/mol more negative for the anions, while the values for the cations are 30.58 kJ/mol more positive in comparison to the values without considering the membrane potential (see SI-1 for  $\Delta G$  profiles of all ions). For most of the anions, a local  $\Delta G$  minimum can be found under the influence of the membrane potential in the area around 10 Å, while the global minima are mostly around the head group region at 22 Å from the bilayer center. In contrast, the  $\Delta G$  minima for the cations are located around 11 to 13 Å from the bilayer center, i.e., deeper in the membrane despite the repulsive forces of the potential (only TPP has its  $\Delta G$  minimum even deeper in the membrane). While 13 out of the 51 anions yield a  $\Delta G$  profile that is negative throughout the whole expansion of the membrane, all of the cations do have an energy barrier in the membrane-water interface that might be explained by unfavorable interactions with the positively charged choline.

Reflecting the  $\Delta G$  profiles, the peak maxima for the relative solute distributions are further away from the membrane bilayer center for the anions (mostly around 23 Å for M2) than for the cations (mostly around 13 Å for M2). Plots of the relative solute distribution of all ions are

shown in the SI-1. The membrane potential has only little influence on the maximum peaks of the relative distribution for most ions: for most anions in model M2 it gets shifted one layer further toward the membrane center, while for most cations it gets shifted one layer toward the head groups. Unexpectedly, almost all presented cations can be found closer to the membrane center than the anions, despite the positive membrane potential in the membrane interior. This could, however, be rather due to the present selection of ions than a generalizable finding because the cationic and anionic chemicals in the present data set have very different structures. In fact, structurally similar TPP and TPB show a contrary trend with regard to their relative distribution (i.e., the positively charged TPP tends to be a little further away from the membrane interior in comparison to the negatively charged TPB).

Fig. 5 shows the experimental values against the overall satisfying predictions of the model M2. Looking at cations and anions separately reveals that the predictions for anions (RMSE = 0.68) are slightly better than for cations (RMSE = 0.74). Note, however, that there are considerably less  $K_{lipw}$  data for cations ( $n = 24$ ) than for anions ( $n = 51$ ). In addition, some of the experimental  $K_{lipw}$  data from (Fruttero et al., 1998) for cationic secondary amines show an unusual sorption behavior; i.e.,  $K_{lipw}$  decreases with increasing chain length for relatively short-chain amines.

The strong outliers which are more than 1.2 log units off in the prediction are mainly big molecules with a molecular weight above 300 except for p-methylbenzylmethylamine cation, which is one of the secondary amines measured by (Fruttero et al., 1998). A reason for the inaccurate prediction of these big molecules might be changes in the membrane provoked by the sorbing molecules that are not accounted for in COSMOmic, like the possible membrane perturbation caused by salmeterol that may have an influence on the fluidity as proposed in (Lombardi et al., 2009).



**Figure 5.** Prediction of neutral (black cross), anionic (blue circles) and cationic (red triangles) chemicals with COSMOmic including the membrane potential using one Gaussian potential for a DMPC membrane (model M2). The identity line (solid) as well as the deviations of  $\pm$ one log unit (dotted) are shifted by 0.32 log units according to the offset of model M2. The linear equation describes the least-squares regression (dashed line). The RMSE is calculated for all 235 neutral and ionic compounds after subtracting the offset. All ions that are predicted more than 1.2 log units off are annotated.

The implementation of the membrane potential leads to a contrariwise shift of the calculated  $K_{lipw}$  values for anions and cations, as expected. The impact of the potential on the calculation of  $K_{lipw}$  is different for each ion but leads to an improved prediction for almost all ions. Not only the potentials absolute height is of importance but also the position and the width matter because most ions have their maximum probability of presence in the headgroup area, where the potential levels off to zero. Within the model M2, 5-chloro-3-tert-butyl-2'-chloro-4'-nitrosalicylanilide (S-13) and TPB show the biggest increase among the anions of

more than 3 log units, while 4-octylbenzene-1-sulfonate exhibits the lowest change with 0.52 log units. The changes for the cations are overall larger than those for anions, going from a decrease of 3.06 (amlodipine) to 5.03 log units (TPP).

#### **1.3.3.4 The influence of the Membrane Potential on the Prediction of Neutral Chemicals**

The prediction of  $K_{lipw}$  values for neutral chemicals is better than for the ionic chemicals according to the lower RMSE. If no constant offset was subtracted from the calculated values, the RMSEs would be 0.05 to 0.11 log units higher than given in Table 2. Note that the data set for neutral chemicals comprises disproportionately more values and spans over more orders of magnitude than the data set of ionic chemicals. The membrane potential has only a marginal influence on the calculation of  $K_{lipw}$  for neutral compounds – for model M2 (Fig. 2 and 5) the largest change due to the implementation of membrane potential is 0.23 log units for carbonyl cyanide p-methoxyphenylhydrazone. Note that in the case of PAHs, there seems to be a trend to underestimate their partition coefficients from water to both octanol and the membrane lipid phase, as has been observed previously (Endo et al., 2011). As the data set of 161 neutral compounds used in this work contains no PAHs, this effect does not show up in the present work.

The observed insensitivity of  $K_{lipw}$  predictions of neutral compounds with respect to the membrane potential is as expected and an affirmative result, confirming the presumed considerations of the model. If the membrane potential had a crucial impact on the sorption of neutral compounds, the bulk phase partitioning between octanol and water could not be expected to correlate so well with neutral  $K_{lipw}$  values.

## 1.4 Comparison of Different Models Predicting the Phospholipid-Membrane Water Partition Coefficients of Charged Chemicals

### 1.4.1 Materials and Methods

#### 1.4.1.1 $K_{lipw}$ Data Compilation for Neutral and Ionic Chemicals

The  $K_{lipw}$  data for neutral chemicals is comprised of 207 chemicals as published in the most thorough and recent data collection (Endo et al., 2011). The  $K_{lipw}$  data for ionic chemicals extended the above used data collection (i.e., 24 cationic and 51 anionic chemicals, which were used to calibrate COSMOmic, as well as 2 divalent cations) to 36 cationic and 56 anionic chemicals, including two perfluorinated anions (PFOS and PFOA). All  $K_{lipw}$  data were measured above the main phase transition temperature and refer to the liquid crystalline state; values and details on the experimental methods are given in Tables 1-3 in the SI-2. 59 % of the experimental data for ions were obtained with equilibrium dialysis experiments. In the case of ionizable, but not permanently charged chemicals, the pH during the experiments was adjusted with buffers to be at least 3 pH units higher (in the case of acids) or 3 pH units lower (in the case of bases) than the  $pK_a$  of the investigated compound. If more than one value could be found in the literature, the arithmetic means of the experimental  $\log K_{lipw}$  values were taken. This is the case for 9 cations (with their  $\log K_{lipw}$  [L/kg] differing by 0.01 to 0.99 log units) and 6 anions (with their  $\log K_{lipw}$  [L/kg] differing by 0.01 to 0.42 log units). A homologous series of four linear quaternary amines and three linear sulfates (Inoue et al., 1986) has been omitted, because the data seem to contradict unpublished LCMS measurements (personal communication with Steven Droge, see SI-2, section 1.4).

Additionally,  $K_{lipw}$  values of two zwitterionic compounds, cetirizine and acrivastine, were included in the evaluation (Plemper van Balen et al., 2001). Zwitterionic compounds are of special interest because they occur frequently in medicinal chemistry (e.g., as drugs) and biochemistry (e.g., as metabolites) (Pagliara et al., 1997).

### 1.4.1.2 Empirical Correlation Approach with $\log K_{ow}$

For the empirical correlation approach with  $\log K_{ow}$  the general workflow was reproduced to derive a  $K_{lipw}$  for neutral and ionic species out of a  $K_{ow}$  as given elsewhere (Escher et al., 2011; Tang et al., 2013). First, the  $\log K_{ow}$  was retrieved from a database like KowWIN (U.S.EPA, version 1.68)<sup>1</sup>. KowWIN only needs a structure input (like a smiles string or a CAS number); note, however, that exclusively  $\log K_{ow}$  values of neutral chemicals are reported, even when the smiles string of a charged chemical is entered (and even when the smiles string represents a permanently charged chemical that has no neutral analog). When no experimental value was found, the predicted values were taken (this was the case for 7 out of 36 cationic chemicals, 16 out of 56 anionic chemicals and 52 out of 207 neutral chemicals). Then the respective  $\log K_{lipw}$  of the neutral chemicals was calculated via a simple linear regression equation of the form  $\log K_{lipw} = a \cdot \log K_{ow} + b$ . The most recent and most thoroughly tested regression equation is based on 156 neutral chemicals (SD=0.426, R<sup>2</sup>=0.948), published in (Endo et al., 2011):

$$\log K_{lipw} = 1.01 * \log K_{ow} + 0.12 \quad (4)$$

The difference between the  $\log K_{lipw}$  of the neutral and the  $\log K_{lipw}$  of the charged species is commonly denominated as  $\Delta mw$  and in a rough approximation it is assumed to be a constant for different chemicals:

$$\Delta mw = \log K_{lipw}(\text{neutral species}) - \log K_{lipw}(\text{charged species}) \quad (5)$$

The empirical correlation approach is based on the observation that  $\Delta mw$  is approximately one log unit for most species investigated (Escher and Schwarzenbach, 1996; Escher and Sigg, 2004).

### 1.4.1.3 PP-LFER Extension for Ionic Compounds

The linear solvation energy relationship (LSER) equation, or generally polyparameter linear free energy relationship (pp-LFER) equation has been shown to account for the relevant

---

<sup>1</sup> EPISuite Exposure Assessment Tools and Models.  
<http://www.epa.gov/opptintr/exposure/pubs/episuite.htm>.



intermolecular interactions in various partition processes (Abraham, 1993; Abraham et al., 2010, 2004). The mandatory descriptors are divided into an energetic contribution from the solute (given in capital letters) and the partition system that is described (the so called system parameters in lower case letters). The following general equation was used in this work:

$$\log K_{lipw} = c + eE + sS + aA + bB + vV \quad (6)$$

where E is the excess molar refraction; S the polarizability/dipolarity parameter; A the solute H-bond acidity; B the solute H-bond basicity and V the McGowan molar volume (units of  $(\text{cm}^3 \text{ mol}^{-1})/100$ ). The complementary system parameters (including c as a constant) are obtained from a multiple linear regression analysis against the experimental partition coefficients.

The pp-LFER equation can be extended for the description of the partitioning of ions via the inclusion of a  $j^+J^+$  term for cations and a  $j^-J^-$  term for anions, leading to the equation:

$$\log K_{lipw} = c + eE + sS + aA + bB + vV + j^+J^+ + j^-J^- \quad (7)$$

As outlined elsewhere (Abraham and Acree, Jr, 2010a, 2010b, 2010c; Saifullah et al., 2011; Zhao and Abraham, 2005), Eq. 7 is derived in a two-step procedure: first, Eq. 6 is fitted for neutral compounds only; then all system parameters describing the neutral interactions (c, e, s, a, b and v) are fixed and only  $j^+$  and  $j^-$  are fitted with the  $\log K_{lipw}$  data of the anions and cations. The applicability of Eq. 7 has been demonstrated for various systems like the partitioning from water to wet octanol (Abraham and Acree, Jr, 2010a; Zhao and Abraham, 2005), to ethylene glycol and to propylene carbonate (Abraham and Acree, Jr, 2010b), to tetrahydrofuran (Saifullah et al., 2011) and others (Abraham and Acree, Jr, 2010c). All of these models are made for the description of single ion partitioning and not of ion pair partitioning or ion exchange. They are based on the broadly used and accepted extrathermodynamic reference electrode assumption, making it possible to derive partition coefficients for single ions between two bulk solvents (Hefter et al., 2002). It is assumed that a measurable thermodynamic property (e.g., a partition coefficient) of a well-selected salt can

be split into equal contributions of the anion and the cation. These two ions must be of similar (large) size and structure. Most commonly used are tetraphenylarsenate or tetraphenylphosphonium (TPP) and tetraphenylborate (TPB) for this assumption, i.e., they are assumed to have the same partition coefficient, because they have the same highly delocalized although still opposite charge on a quasi-spherical surface (Wachter et al., 2006). However, while this assumption is well-suited for bulk solvents, it does not hold in the case of an anisotropic lipid bilayer, where the negatively charged TPB ( $\log K_{lipw} = 5.20$ ) (Flewelling and Hubbell, 1986b) sorbs 4 orders of magnitude stronger than its positively charged analog TPP ( $\log K_{lipw} = 1.19$ ) (Demura et al., 1987; Flewelling and Hubbell, 1986b). This difference can only be explained by the different influence of the membrane potential on positively and negatively charged chemicals (Flewelling and Hubbell, 1986b; Wang, 2012), as outlined above and further discussed below.

Solute descriptors of neutral compounds can be obtained from the ‘UFZ-LSER database’, which contains several thousands of experimentally determined descriptors and is available free of charge (Ulrich S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U., 2017); or the estimation method Absolv (module in ADME Boxes version 5.0, ACD/Labs)<sup>2</sup> can be used. The latter is a group contribution method for the solute descriptors (Platts et al., 1999), including some additional but not reported optimizations. In combination with calibrated system parameters, the RMSE of Absolv based partition coefficients can be expected to be less than one log unit for the prediction of the partitioning of neutral compounds between various solvents (Endo and Goss, 2014).

All solute descriptors for ionic compounds have to be derived based on an empirical summation of certain fractions of the solute descriptors of the corresponding neutral compounds and in some cases additional information like the  $pK_a$  (for phenoxide anions) or

---

<sup>2</sup> Advanced Chemistry Development, Inc. (ACD/Labs). Absolv prediction module data sheet. Toronto, ON (Canada). <http://www.acdlabs.com/products/percepta/predictors/absolv/>

the number of hydrogen atoms attached to charged nitrogen (for amine cations). See SI-2, section 3.1 for all equations. To date, these derivations are only described for a limited number of chemical classes (for carboxylic acid anions and amine cations (Abraham and Acree, Jr, 2010c), for phenoxide anions (Abraham and Acree, Jr, 2010d) and pyridinium cations (Abraham and Acree, Jr, 2010b) ). It is intuitively comprehensible that all solute descriptors differ between a neutral and the corresponding ionic chemical (and not only  $J^+$  and  $J^-$  have to be added), because neutral and ionic chemical can undergo different interactions with the solvent; e.g. a neutral phenol can act as a hydrogen bond donor and acceptor, while a deprotonated phenol can only act as a hydrogen bond acceptor.

Due to the descriptor limitations to selected chemical classes as described above (Abraham and Acree, Jr, 2010b, 2010c, 2010d), only 32 out of 36 cations and 42 out of 56 anions could be taken into consideration. 11 out of the 32 descriptors for cations and 25 out of 42 descriptors for anions are based on experimentally determined solute descriptors of the corresponding neutral chemicals. Absolv predictions were taken for the remaining 21 cations and 17 anions, where no solute descriptors of the corresponding neutral compounds could be found. See all values and further discussion in the SI-2, section 3.

#### **1.4.1.4 COSMO-RS and COSMOmic**

The extension (available with COSMOtherm(X) release C30-1501) of the originally (Klamt et al., 2008) published COSMOmic was used, which also incorporates a membrane potential in the membrane interior of roughly +0.3 V in reference to the surrounding bulk water, as outlined above. See SI-2, sections 2.1 and 4 for further details.

#### **1.4.2 Results and Discussion**

The experimental  $\log K_{lipw}$  values of neutral chemicals span over 9 orders of magnitude (from -1.24 to 7.86), while the experimental  $\log K_{lipw}$  values of cations and anions only differ by 3.37 log units (from 0.66 to 4.03) and 4.89 log units (from 0.31 to 5.20), respectively.

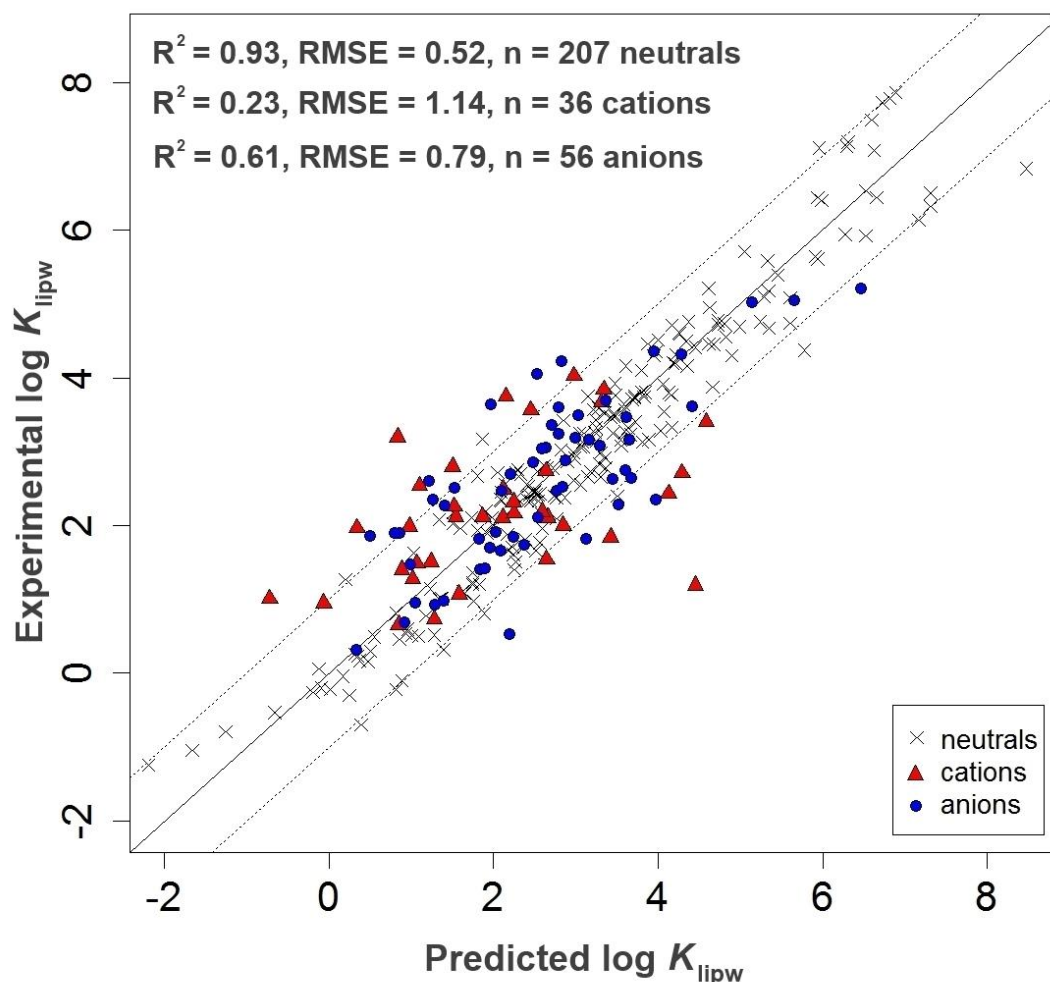
Given the larger range of values, it is very likely that the predictions of the neutral chemicals will generally yield a higher  $R^2$  than the predictions of the ionic chemicals. In order to compare predictions spanning over different orders of magnitudes and covering a different amount of data points ( $n=207$  neutrals, 36 cations, 56 anions), the RMSE is a more suited and meaningful measure. All RMSEs are reported in log units.

#### 1.4.2.1 Empirical Correlation Approach with $\log K_{ow}$

Fig. 6 shows that there is good agreement for the 207 neutral compounds between the experimental (Endo et al., 2011) and predicted  $\log K_{lipw}$  data ( $R^2=0.93$ ,  $RMSE=0.52$ ). This harmonizes well with the finding that, according to COSMOmic calculations, most neutral compounds sorb to a membrane depth that exhibits a similar chemical environment as wet octanol (see SI-2, section 2.2 for details). However, the statistical analysis has some restrictions, because 156 out of the predicted 207 neutral compounds have been within the training data set to derive the regression equation. As outlined previously (Endo et al., 2011), it appears practicable to use experimental  $K_{ow}$  values, if available, to derive  $K_{lipw}$  values of neutral chemicals. It must be noted though, that this is a purely empirical relationship that can accordingly only be used with some confidence for predictions within the chemical space of compounds used to derive the regression Eq. 4, which might not be easy to judge.

In contrast, the experimental  $\log K_{lipw}$  values of the 36 cationic compounds showed a poor correlation with the modeled predictions ( $R^2=0.23$ ,  $RMSE=1.14$ ). On the other hand the predictions for anions are better than could a priori be expected ( $R^2=0.61$ ,  $RMSE=0.79$ ). Although it is plausible that ions have a lower sorption to membranes than their corresponding neutral counterparts, it is not plausible that this difference should be constant for different chemicals: The generic value of 1 for  $\Delta mw$  is mainly based on data for phenolic acids (Escher and Schwarzenbach, 1996), but was also shown to be a useful descriptor for screening purposes for complex mixtures of compounds (Spycher et al., 2008; Tang et al.,

2013). However, it has also been discussed that  $\Delta mw$  for carboxylic acids is usually closer to 2, while it varies a lot for aliphatic amines (Escher and Sigg, 2004).



**Figure 6.** Comparison between the experimental  $\log K_{lipw}$  values of 207 neutral, 36 cationic and 56 anionic compounds and the predicted values according to the empirical correlation approach with  $\log K_{ow}$  using KowWIN, simple regression and  $\Delta mw$  as outlined above. Deviations of one log unit from the straight identity line are shown as dotted lines.

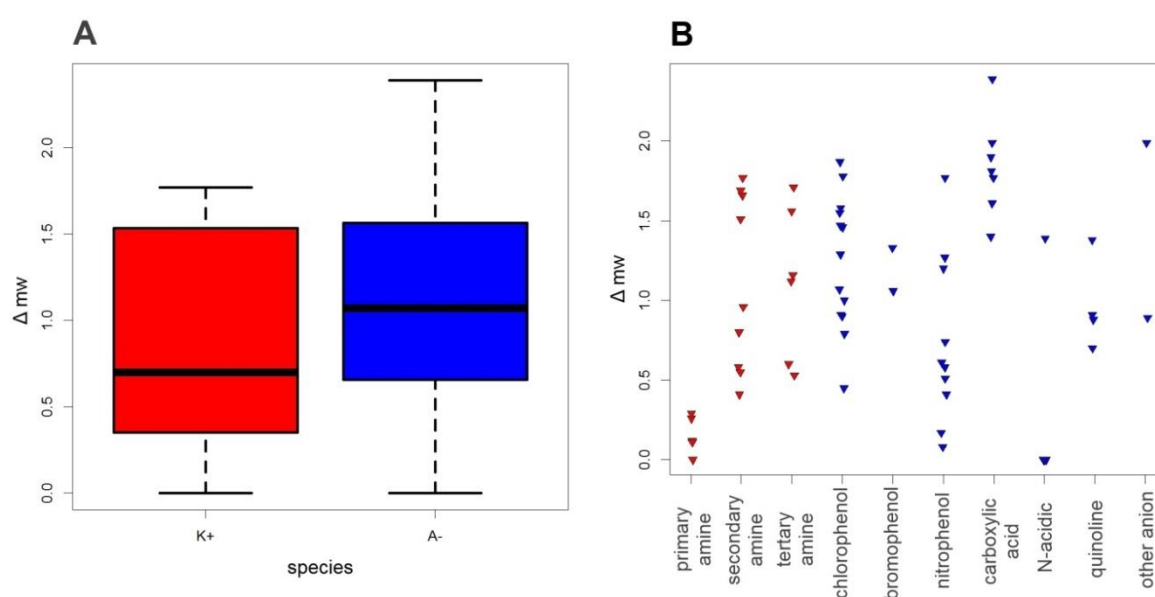
As already suggested by Fig. 6, Fig. 7 shows that the median of the 43 experimental  $\Delta mw$  values of the anionic chemicals is closer to the generic value of 1 than the median of the 20 experimental  $\Delta mw$  values of the cationic chemicals. Accordingly, most of the  $\log K_{lipw}$  predictions for cations in Fig. 6 scatter considerably more than the predictions for anions. This should be particularly relevant for pharmaceuticals and illicit drugs (Zuccato et al., 2008), industrial surfactants and biocides (Li and Brownawell, 2010), because they often have an

amine group and are therefore positively charged or partially charged in a physiological pH range.

The general limitations of the assumption that  $\Delta mw = 1$  are in agreement with previous evaluations with smaller data sets (Escher and Sigg, 2004; Neuwoehner et al., 2009). Comparing experimental  $\log K_{lipw}$  of ionic and corresponding neutral chemicals for 20 cations showed that  $\Delta mw$  varied between 0 (amlodipine) and 1.77 (p-methylbenzyl-hexylamine), while the mean  $\Delta mw$  was 0.87 ( $\pm 0.61$  standard deviation). In contrast, the  $\Delta mw$  for 43 anions varied between 0 (benzimidazoles and hydrazones) and 2.39 (octanoic acid), with a mean  $\Delta mw$  of 1.09 ( $\pm 0.63$ ). Looking at the difference between and the variability within chemical classes showed that  $\Delta mw$  for 25 phenols were on average 1.03 ( $\pm 0.51$ ), while 7 carboxylic acids had a  $\Delta mw$  of 1.84 ( $\pm 0.31$ ). Even with a more subtle classification into subclasses there is considerable variation: 13 chlorophenols had a  $\Delta mw$  of 1.24 ( $\pm 0.42$ ), while 10 nitrophenols had a  $\Delta mw$  of 0.73 ( $\pm 0.53$ ). See SI-2, sections 1 and 2 for all values and further discussion. While a class-specific fit of  $\Delta mw$  would improve the overall prediction quality, such a fit is not feasible with the limited available experimental data (Armitage et al., 2013) and would not solve the problem that  $\Delta mw$  is not applicable for multifunctional molecules or for permanently charged chemicals. There is a tendency, that  $\Delta mw$  increases with increasing charge density in the case of anions, while  $\Delta mw$  decreases with increasing charge density in the case of cations as discussed in detail in the SI-2, section 2.1. It is interesting that it was possible to confirm this experimental finding, but the correlation is rather qualitative than quantitative and cannot serve as a reliable predictor for  $\Delta mw$ .

Overall, the empirical correlation approach with  $\log K_{ow}$  seems to give a reasonable estimation of  $K_{lipw}$  for most monovalent ions presented here (Fig. 6) – however, one has to be alert that it is very difficult to judge when the approach is not applicable. It is not fully clear, why the approach works better for anions than for cations, but this is most likely an artefact due to the selection of chemicals (as suggested by the class-specific differences in  $\Delta mw$  as

shown above). According to the COSMOmic calculations, the cations mainly sorb to the membrane layers with the best H-bond acceptor properties, while the anions prefer membrane layers further away from the membrane center, which exhibit the best H-bond donor properties (see SI-2, section 2.2). Neither of the preferred sorption depths resembles the chemical properties of octanol well. Within the empirical correlation approach anions and cations are treated the same, as expressed by Eq. 5. Zwitterions, divalent ions and permanently charged chemicals are not accounted for at all.



**Figure 7.** Box-and-whisker plot of 20 cationic (red) and 43 anionic (blue) experimental  $\Delta mw$  values (A) and experimental  $\Delta mw$  values of the subclasses (B). The boxes in (A) outline the 25th to 75th percentiles, the lines through the centers represent the median and the whiskers extend to the most extreme data point. All  $\Delta mw$  values  $\pm$  standard deviation of the different species are summarized in the SI-2, Table 4.

#### 1.4.2.2 PP-LFER Extension for Ionic Compounds with Experimental Descriptors and ABSOLV

In order to have the most accurate pp-LFER model for the  $K_{lipw}$  prediction of neutral chemicals the following previously published equation (Endo et al., 2011) was used:

$$\log K_{lipw} = 0.26(\pm 0.08) + 0.85(\pm 0.05)E - 0.75(\pm 0.08)S + 0.29(\pm 0.09)A - 3.84(\pm 0.10)B + 3.35(\pm 0.09)V; SD = 0.279, n(\text{neutral}) = 131, R^2 = 0.979 \quad (8)$$

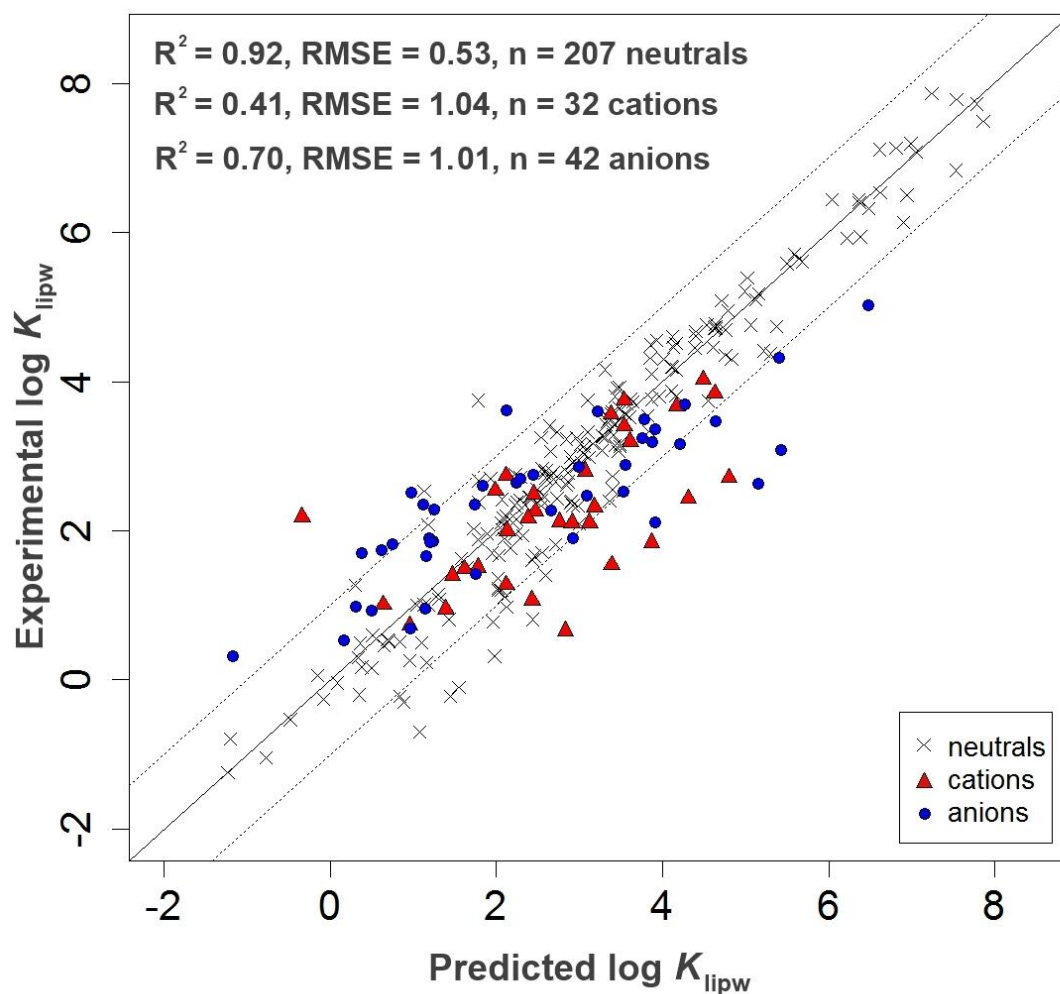
Eq. 8 is based on the 131 out of the 207 neutral chemicals, for which experimentally determined solute descriptors were available (Endo et al., 2011), in order to ensure the most

reliable  $K_{lipw}$  prediction of the neutral chemicals. Taking the system parameters of Eq. 8 and additionally fitting the system parameters  $j^+$  and  $j^-$  (that describe the solvent interactions for cationic and anionic chemicals) with solute descriptors for 74 ionic chemicals gives Eq. 9:

$$\log K_{lipw} = 0.26(\pm 0.08) + 0.85(\pm 0.05)E - 0.75(\pm 0.08)S + 0.29(\pm 0.09)A - 3.84(\pm 0.10)B + 3.35(\pm 0.09)V - 1.72(\pm 0.08)J^+ + 3.98(\pm 0.05)J^-; SD = 1.011, n(ion) = 74, R^2 = 0.988 \quad (9)$$

Fig. 8 shows the performance of Eq. 9 in predicting the  $\log K_{lipw}$ : neutral compounds are predicted equally well as in the empirical correlation approach with  $\log K_{ow}$  (given a similar restriction as above, that 131 out of 207 compounds were also used in the fitting procedure), with  $R^2 = 0.92$  and  $RMSE = 0.53$ . In contrast, the fit for cations ( $R^2=0.41$ ,  $RMSE=1.04$ ,  $n=32$ ) and anions ( $R^2=0.70$ ,  $RMSE=1.01$ ,  $n=42$ ) is not satisfying; particularly in view of the fact that the solute descriptors for all ions were used in the fitting procedure. A plausible explanation for this is, that both cations and anions sorb to different depths in the membrane and will therefore be exposed to a different physicochemical environment due to the heterogeneous structure of the membrane, which cannot be captured by only one pp-LFER equation. This consequence of the membrane anisotropy also explains the huge difference in  $K_{lipw}$  of TPP and TPB (as outlined above) and is further discussed below and Fig. 10. Limiting the fitting data set of ions to predictors that are only based on experimentally derived predictors for the corresponding neutral chemical does not alter this finding, as discussed in the SI-2, section 3.2.





**Figure 8.** Comparison between the experimental and predicted  $\log K_{lipw}$  values using the pp-LFER Eq. 6; deviations of one log unit from the straight identity line are shown as dotted lines.

Changing the fitting procedure for the pp-LFER equation seems to improve the model performance at first glance: when all system parameters of Eq. 7 are fitted together in only one step with one multi linear regression, the models seems to perform significantly better with respect to ions (32 cations: RMSE=0.69; 42 anions: RMSE=0.62), while being only slightly worse with respect to the prediction of neutral chemicals (RMSE=0.57, n=207). However, fitting all system parameters at once also results in a very different pp-LFER equation (Eq. 10):

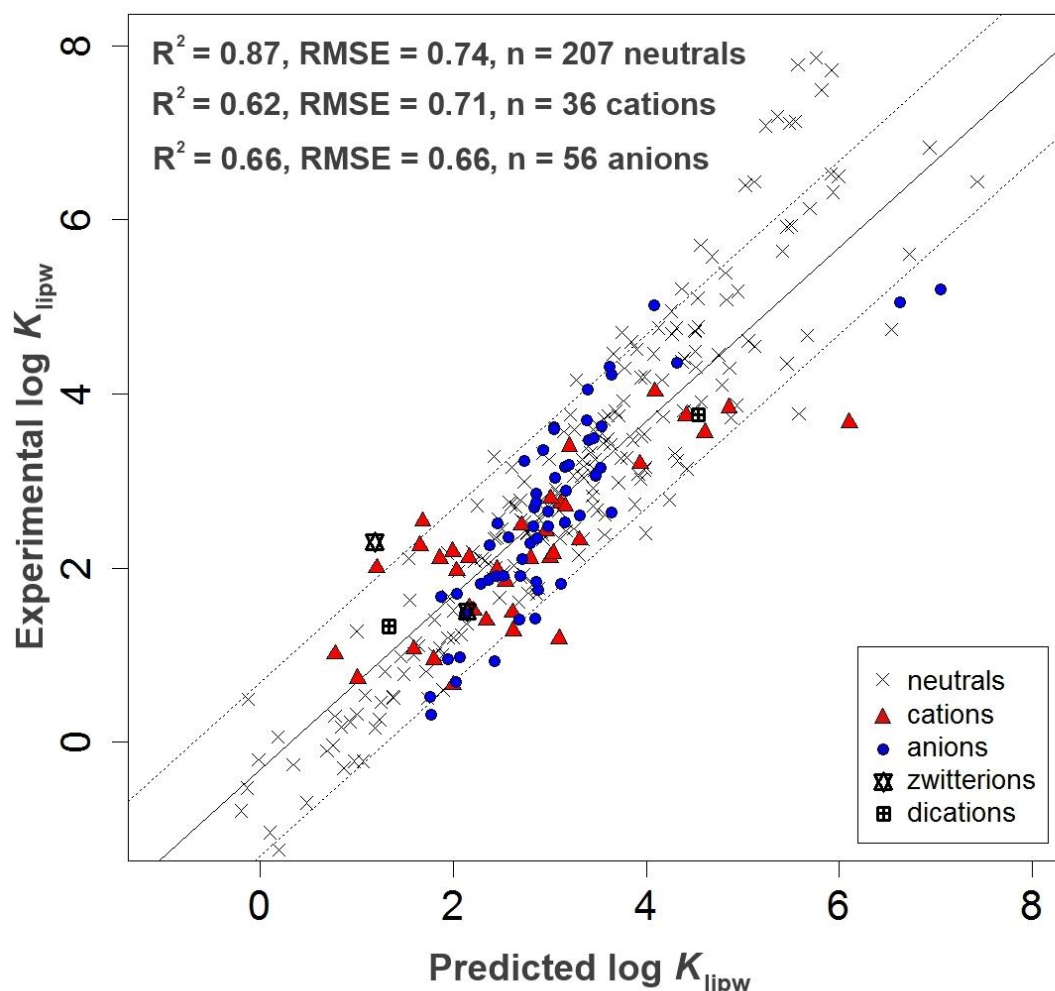
$$\log K_{lipw} = 0.44(\pm 0.12) + 0.99(\pm 0.06)E - 0.69(\pm 0.05)S - 0.19(\pm 0.10)A - 2.82(\pm 0.11)B + 2.83(\pm 0.12)V - 1.14(\pm 0.10)J^+ + 2.95(\pm 0.14)J^-; SD = 0.507, n(\text{neutrals} + \text{ions}) = 205, R^2 = 0.912 \quad (10)$$

It is important to note, that both the terms for neutral as well as for ionic interactions are altered considerably in this Eq. 10 with respect to Eq. 9:  $\bar{j}$  and  $b$  become 1.03 and 1.02 units smaller, respectively, while  $a$  even changes its sign from plus to minus. This is again an indication that the parameters in the general Eq. 7 are not describing the membrane-water partitioning system sufficiently (when ions are included), because the membrane anisotropy cannot be accounted for. Zwitterions and divalent ions are not yet explicitly addressed within the pp-LFER approach. However, it also has to be pointed out that for some solute descriptors the values for ions are much bigger than for the neutral chemicals. E.g., B and S for the 42 anions go up to values of 4.39 and 16.59, respectively, while B and S for the 207 neutral compounds do not exceed the values of 2.19 and 3.29, respectively. In order to cover the full physicochemical space occupied by the solute descriptors, it seems to be more meaningful to fit ions and neutral chemicals together for the derivation of a pp-LFER equation. This is not the procedure recommended by Abraham et al. (Abraham and Acree, Jr, 2010a, 2010b, 2010c; Saifullah et al., 2011; Zhao and Abraham, 2005), but it could be successfully used recently to fit the partitioning of 46 neutral, 34 anionic, and 6 cationic chemicals to muscle protein ( $R^2=0.89$ ,  $RMSE=0.29$ ) (Henneberger et al., 2016).

#### 1.4.2.3 COSMO-RS and COSMOmic

As expected from the results shown above with a slightly smaller data set, COSMOmic was predicting the  $K_{lipw}$  of neutral compounds ( $R^2=0.87$ ,  $RMSE=0.74$ ) with roughly the same accuracy as the  $K_{lipw}$  of ionic compounds (36 cations:  $R^2=0.62$ ,  $RMSE=0.71$ ; 56 anions:  $R^2=0.66$ ,  $RMSE=0.66$ ). Compared to the previous two models, COSMOmic performed slightly worse regarding the neutral compounds, but it was clearly the best with respect to ionic compounds. Given that the model has a sound mechanistic basis and can be used

independently of the charge as outlined above, it is expected to serve as the most reliable model when it comes to the prediction of  $K_{lipw}$  of ions. The strong outliers which are more than 1.2 log units off in the prediction are mainly large molecules with molecular weights above 300. Reasons for the increased occurrence of strong outliers for these chemicals are already discussed above. Regarding the additionally predicted ions it needs to be emphasized that two emerging pollutants, the perfluorinated PFOS (perfluorooctane-1-sulfonic acid,  $\log K_{lipw}(exp)= 3.15$  (Lehmler et al., 2006),  $\log K_{lipw}(calcd)= 3.53$ ) and PFOA (perfluorooctanoic acid,  $\log K_{lipw}(exp)= 2.34$  (Inoue et al., 1988),  $\log K_{lipw}(calcd)= 2.88$ ), were well-predicted. These chemicals are of special concern because they are highly persistent, bioaccumulative and detected globally (Houde et al., 2006) ; moreover they are difficult to assess with traditional approaches because they are essentially permanently charged in the environment due to their very low  $pK_a$  (Goss, 2008). In contrast to the two models presented above, also zwitterions and divalent ions could be calculated in COSMOmic: For the few data available on zwitterionic chemicals there was good agreement between experiment and predicted values: the model correctly predicted the cetirizine zwitterion ( $\log K_{lipw}(exp)=2.30$  (Plempers van Balen et al., 2001),  $\log K_{lipw}(calcd)= 1.19$ ) to have a lower  $K_{lipw}$  than the corresponding cation ( $\log K_{lipw}(exp)= 3.20$  (Plempers van Balen et al., 2001),  $\log K_{lipw}(calcd)= 3.94$ ). Similarly, the acrivastine zwitterion ( $\log K_{lipw}(exp)= 1.50$  (Plempers van Balen et al., 2001),  $\log K_{lipw}(calcd)= 2.15$ ) was correctly predicted to have a lower  $K_{lipw}$  than the corresponding anion ( $\log K_{lipw}(exp)= 2.60$  (Plempers van Balen et al., 2001),  $\log K_{lipw}(calcd)= 3.31$ ).



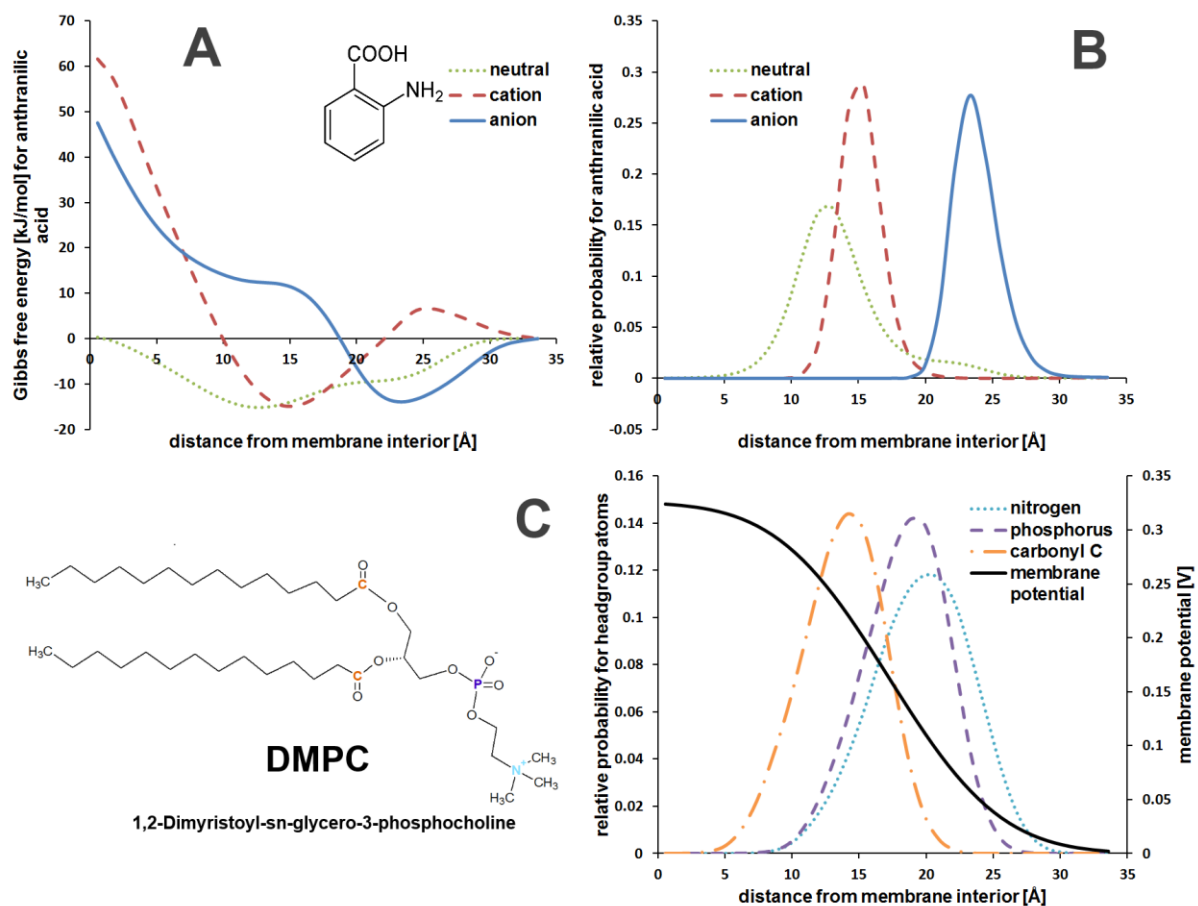
**Figure 9.** Comparison between the experimental and predicted  $\log K_{lipw}$  values using the model COSMOmic (version C30-1501). A constant offset of 0.32 log units has been subtracted from all predicted values (but not from the points depicted in the Fig. 9) for the calculation of the RMSE. The straight identity line as well as the deviations of one log unit (dotted lines) are shifted according to the offset.

The Gibbs free energy profile and the corresponding relative distribution of a molecule can also be calculated with COSMOmic as shown in Fig. 10. For neutral chemicals it has been shown that they are in good agreement with computationally more costly molecular dynamics simulations (Ingram et al., 2013; Jakobtorweihen et al., 2014). This kind of data is almost not accessible with experimental methods. It facilitates the understanding of the different sorption behavior depending on the different speciation: anthranilic acid for example, which can be present in a neutral, cationic or anionic form, reveals very differently shaped energy minima

depending on the speciation. Mirroring these minima, the respective relative distribution profiles show that neutral, cationic and anionic form even of the same molecule are located at very different depths in the membrane. Given that the phospholipid membrane is a highly anisotropic system and not a bulk phase, the sorption to different depths results in different solvent-solute interactions. Thus the membrane appears as a different sorption medium for each of the three species, which elucidates the difficulties inherent in the other two approaches presented above.

Fig. 10 and SI-2, Fig. 4 suggest that cations rather sorb on the ‘interior’ side of the headgroup (i.e., closer to the negatively charged phosphate), whereas anions rather sorb on the ‘exterior’ side of the headgroup (i.e., closer to the positively charged quaternary amine), which intuitively makes sense. However, the overall sorption of ionic chemicals to a phospholipid membrane is not only a function of these electrostatic forces, but also of the specific depth depended chemical environment and the resulting van der Waals and H-bond interactions (see SI-2, section 2.2), which additionally is superimposed by the membrane potential (Flewelling and Hubbell, 1986b; Wang, 2012), as also discussed above.

Beside the in-depth data discussed above, *COSMOmic* can also serve as a screening tool for partition coefficients (Jakobtorweihen et al., 2013). Critical for this purpose are the availability of a depth-dependent membrane composition (Jakobtorweihen et al., 2013) with an optimized membrane potential discussed in the previous section as well as a database of cosmo files for the molecules of interest (see SI-2, section 2.1 for one possible workflow for the generation of cosmo files). The time requirement for the calculation of cosmo files is mainly dependent on the number of atoms: on a standard CPU 12 atoms need minutes, 20 need hours, 40 need days and 100+ are in the range of weeks (*COSMOconf* version 3.0 manual).



**Figure 10.** A: Gibbs free energy of anthranilic acid in the neutral, cationic and anionic speciation calculated with the enhanced COSMOmic version, incorporating a membrane potential (version C30-1501); B: resulting solute distribution and C: relative probability profile of the membrane headgroup atoms (nitrogen and phosphorus) and carbonyl carbon atoms (left axis), as well as membrane potential (right axis). According to the color code in plot C a representative DMPC lipid molecule is shown in the bottom left corner. The distance from the membrane interior on the x-axis shows only one half of the mirror-imaged membrane: while 0 Å is in the middle of the membrane, the headgroup region is at around 20 Å and the bulk water phase begins at 30 Å.

## 1.5 Assessing the Toxicity of Ionic Liquids – Application of the Critical Membrane Concentration Approach

### 1.5.1 Materials and Methods

#### 1.5.1.1 Basic Assumptions and Considerations

A principal assumption within the baseline toxicity concept is that baseline toxicity can be described independently of the organism with only one partition coefficient, i.e., that the membranes in different organisms exhibit similar sorption characteristics. It is important to keep in mind, as already discussed above, that the  $K_{lipw}$  is based on a pure phospholipid membrane whereas the  $K_{mw}$  describes the sorption to a real (and complex) biological membrane, including other components such as cholesterol, different kinds of phospholipids and proteins. Here, the  $K_{lipw}$  is taken as surrogate for the  $K_{mw}$ , irrespective of the kind of organism or cell culture, which is generally a well-accepted assumption (Endo et al., 2011). For phenols it has been shown that liposomes composed of zwitterionic phosphatidylcholine mimic the sorption behavior of isolated membranes from *Rhodobacter sphaeroides* well (Escher and Schwarzenbach, 1996). The second crucial assumption of the baseline toxicity concept is the non-specificity, i.e., that the critical toxic concentration in the membrane is fairly independent of the nature of the chemical. Hence, the baseline toxicity model described here is based on an organism independent  $K_{mw}$  and a toxic threshold concentration in the phospholipid membrane, which is independent from the type of chemical. It has to be noted that the ‘target lipid model’ (Kipka and Di Toro, 2009) works in exact analogy - the ‘target lipid’ equals the membrane in the baseline toxicity concept. The fundamental relationship between the constant toxic membrane concentration and the toxic water concentration (LC50) is given by the membrane water partition coefficient

$$c_{mem}^{tox} = K_{mw} * LC50 \quad (11)$$

In principle, the LC50 in Eq. 11 is defined as the freely dissolved water concentration and not as the nominal water concentration, which is nonetheless often reported in toxicity

experiments (Escher and Hermens, 2002). The critical membrane concentration,  $c_{mem}^{tox}$ , causing a toxic effect (i.e., mortality of 50% of the organisms in the case of Eq.11, where  $c_{mem}^{tox}$  is related to the LC50) is analogous to the  $ILC50_{membrane\ lipid}$  discussed in (Escher and Hermens, 2002). A different abbreviation was chosen in this work because the prefixed “I” refers to the “internal” concentration. Although it may seem as a negligible detail, this work is focused only on the total membrane lipids of organisms or cell cultures regardless whether the membrane is inside or at the outer border of the respective organism or cell culture. Thus the term  $ILC50_{membrane\ lipid}$  is avoided in this work, although it is generally used in analogy to  $c_{mem}^{tox}$  defined in Eq. 11. Rearranging Eq. 11 and taking the logarithms leads to

$$\log LC50 = -1 * \log K_{mw} + \log c_{mem}^{tox} \quad (12)$$

Eq. 12 is the most frequently plotted correlation between  $\log K_{mw}$  and  $\log LC50$  when baseline toxicity is assessed. A crucial condition for this relationship is that the concentrations in water and membrane are in equilibrium with each other.

In the case of permanently charged chemicals - ionic liquids (ILs) in the presented work - both the anionic as well as the cationic compound have individual  $K_{mw}(ion)$  values that need to be considered. Hence, no single  $\log K_{mw}$  can be given for an IL salt and plotted against  $\log LC50$  as suggested by Eq. 12. Alternatively, the total membrane concentration at LC50,  $c_{mem}^{tox}(total)$ , was calculated from additive contributions of the respective anionic and cationic chemical via

$$c_{mem}^{tox}(total) = K_{mw}(anion) * LC50(IL) + K_{mw}(cation) * LC50(IL) \quad (13)$$

In Eq. 13 additive contributions of the respective anionic and cationic chemical is assumed both for the overall concentration in the membrane (via partitioning) as well as for the toxic mode of action (i.e., both anionic and cationic species act as baseline toxicants). Thus, the baseline toxicity concept in the case of ILs can be validated by checking whether  $c_{mem}^{tox}(total)$  falls into a similar range as previously determined for neutral chemicals, independent of the combination of anionic and cationic chemicals. Within this approach it is



implicitly assumed that potentially formed ion pairs in water or in the membrane are negligible (Escher et al., 2000) and that there are always enough background electrolytes so that differential partitioning of the anion and the cation does not infringe electroneutrality.

In order to assess whether the IL salts act according to a specific mode of toxic action (e.g., uncoupling) or according to baseline toxicity, the toxic ratio (TR) is assessed. The TR was originally introduced by (Verhaar et al., 1992) as the ratio between  $LC50_{baseline\ QSAR}$ , the predicted baseline effect concentration, and  $LC50_{experimental}$ , the experimental concentration for a given toxic endpoint:

$$TR = \frac{LC50_{baseline\ QSAR}}{LC50_{experimental}} = \frac{c_{mem}^{tox}(total) / [K_{mw}(anion) + K_{mw}(cation)]}{LC50_{experimental}} \quad (14)$$

Using the geometrical mean value of  $c_{mem}^{tox}(total)$  determined for neutral chemicals in the first part of this work the  $LC50_{baseline\ QSAR}$  can be calculated by simply rearranging Eq. 13 as shown in Eq. 14. Due to the scatter of baseline toxicity it is a concentration range rather than a fixed membrane concentration in which baseline toxicity is expected. Consequently, according to (Escher et al., 2017), all chemicals in the range  $0.1 < TR < 10$  are considered to be baseline toxicants; chemicals with  $TR > 10$  are considered to have modes of action causing excess toxicity and chemicals with  $TR < 0.1$  are less toxic than expected by the assumptions of the baseline toxicity concept.

In previous publications, e.g., (Endo, 2016; Escher and Schwarzenbach, 2002) and other publications cited above, the total concentration in an organism has been determined causing a toxic effect and thus the internal concentration at the target site (e.g., the membrane lipid) has been deduced by multi-compartment modelling. This approach yields a lot of insights into the internal distribution of a chemical, but requires a lot of experimental details (or at least sound assumptions). In this work nominal water concentrations were investigated in lieu of non-available freely dissolved concentrations causing a toxic effect and directly combined

this information with the respective  $K_{mw}$  values to yield the internal membrane concentration, assuming steady state distribution.

### 1.5.1.2 Compilation of Experimental Data

#### 1.5.1.2.1 Toxicity Data for Neutral Chemicals

The toxicity data set for neutral organic chemicals is based on a thorough revision of a published data set (Kipka and Di Toro, 2009), which originally comprises 1687 experimental LC50 values for 42 aquatic organisms (368 chemicals). By re-analyzing the original data set by (Kipka and Di Toro, 2009) it became evident that for seven chemicals the water solubility was below the reported LC50 values (experimental water solubility taken from PhysProp database accessed via Episuite<sup>3</sup> – for details see SI-3, Table 1). These seven chemicals were excluded from further analysis, although an oversaturation does not necessarily render the experiment futile (as long as the chemicals do not precipitate). Further 23 acidic chemicals (mostly phenols) were sorted out, whose  $pK_a$  values are smaller than 9 as well as five bases (anilines and pyridines) whose conjugated protonated acids have  $pK_a$  values larger than 5 (see SI-3, Tables 2 and 3, respectively).

After the data-quality check described above 1591 experimental LC50 values are left for 42 aquatic organisms (333 chemicals). While all 1687 LC50 values have been used in the original publication (Kipka and Di Toro, 2009) to parametrize a pp-LFER model based on the ‘target lipid’ concept, in this work the revised, concise data set is re-evaluated with regard to the concept of baseline toxicity. Thus, the revised data set is summarized in a two-step procedure:

First, the arithmetic mean was calculated for all reported experimental LC50 values for identical chemicals for each organism, resulting in 1072 organism- and chemical-specific

---

<sup>3</sup> U.S. EPA, EPISuite Exposure Assessment Tools and Models, US Environmental Protection Agency, 2009, <https://www.epa.gov/>.

LC50 values, clustered in 4 up to 216 chemicals for each of the 42 organism (see SI-3, Table 4). These organism- and chemical-specific LC50 values are the test set for evaluating the ‘baseline toxicity-QSAR’ according to Eq. 11 as discussed in SI-3, section 4.

In a second step the 1072 data points above were further summarized by taking the geometric mean of the different chemicals irrespective of the corresponding organisms. This boils down the revised data set to a maximum degree in order to comply with principle assumptions of the baseline toxicity concept (being that toxicity can be described independently of the organism). The resulting 333 LC50 values go from  $2.5 \cdot 10^{-4}$  to 1080 mmol/L(water).

It has to be noted that the LC50 values given here are nominal water concentrations which can be substantially different from freely dissolved water concentrations. This issue is discussed in the OECD guideline 203 for acute toxicity testing for fish. However, in the case of flow through tests (which is e.g. often done for fish (Cowan-Ellsberry et al., 2014)), actual concentrations are often measured and animals are often not fed in acute toxicity tests. Hence, large deviation between the reported nominal concentrations and the actually freely dissolved concentrations were not expect, but it has to be kept in mind that the validity of the present data set is restricted.

#### **1.5.1.2.2 Toxicity Data for Ionic Liquids (ILs)**

ILs have been chosen in this work as the object of investigation to shed light on whether the baseline toxicity concept can be applied also to organic ionic chemicals, because the ILs in this work are predominantly permanently charged. The two anions, bis(trifluoromethylsulfonyl)imide,  $pK_a(\text{JChem}^4) = -0.54$ , and dicyanamide,  $pK_a < 1$  (Gazitúa et al., 2014), can theoretically be protonated and thus neutral, but this can be neglected

---

<sup>4</sup> JChem for Excel, version 15.10.2600.341, Copyright 2008-2015 ChemAxon Ltd. <https://www.chemaxon.com/>.

because care has been taken in the experiments to exclude pH effects (see references in SI-3, Table 7). The anions tetrafluoroborate, hexafluorophosphate, chloride and bromide are the only ones that are not organic (see SI-3, Table 6), but they were also included in the calculation of the toxic membrane concentration as outlined by Eq. 13. The herein investigated toxicity data are based on a recently published review (Thuy Pham et al., 2010), including toxicity data for ILs composed of 39 organic cations and 6 anions (resulting in 96 different salt combinations and 169 different experimental toxicity values, see SI-3, Table 8 for all IL structures). Overall, the experimental EC50 values go from  $2.7 \cdot 10^{-7}$  to 178 mmol/L(water) (the IL toxicity data set is not only comprised of LC50 values, but also of EC50 values such as growth inhibition). As also discussed above for the toxicity data set of neutral chemicals it has to be kept in mind, that the LC50 and EC50 values for ILs are based on nominal water concentrations. This can lead to substantial artefacts in the case of the cell assay toxicity tests which are part of the IL toxicity data set. The discrepancy between nominal and freely dissolved concentration has been discussed in (Armitage et al., 2014) for neutral chemicals and in (Fischer et al., 2017) for ionizable chemicals. Unfortunately, it is very difficult to determine the freely dissolved water concentrations for cell assays (if at all possible).

### 1.5.1.3 Calculation Methods

#### 1.5.1.3.1 pp-LFER for $K_{mw}$ (neutral)

In contrast to the original publication (Kipka and Di Toro, 2009), where ADME Boxes Version 3.0 Absolv package was used to predict all pp-LFER descriptors, the UFZ-LSER database was used in this work (Ulrich S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U., 2017), in order to get a maximum of experimental descriptors ('UFZ preselected published values'). The use of experimental descriptors is superior to predicted values. If no experimental descriptors were available, they were predicted with the

UFZ-LSER-QSAR by Trevor Brown, accessed via the same database (Ulrich S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U., 2017).

Finally,  $K_{mw}$  of the neutral chemical was predicted with Eq. 8 introduced in section 1.4.2.2 (Endo et al., 2011).

#### 1.5.1.3.2 COSMOmic for $K_{mw}$ (ion)

The  $K_{lipw}(\text{ion})$  was again calculated with COSMOmic as outlined above, i.e., including the membrane dipole potential. Analogous to the assumptions made for neutral chemicals, the calculated  $K_{lipw}(\text{ion})$  was taken as a surrogate for the (real) biological membrane-water partition coefficient of the respective ion,  $K_{mw}(\text{ion})$ .

As also outlined above, a constant offset value of 0.3 log units from the COSMOmic calculated log  $K_{lipw}$  values was subtracted. This offset value is most likely due to an energy contribution needed to deform the membrane in order to make space for the sorbing molecule. The COSMOmic model does not account for this 'volume work', which makes partitioning into the membrane less favorable. Given that the implementation of the COSMO-RS theory in COSMOtherm (which is the 'engine' behind the COSMOmic calculation) is basically free of fitting factors (Klamt, 2015) and therefore not limited to certain chemical classes or structures, the initially deduced offset value is considered appropriate: the molecular weight (which is proportional to the volume) of the ILs investigated here goes from 35.5 to 349.6 (median 178.3). This is well within the 'volume range' of the chemicals used to calibrate COSMOmic in section 1.3, whose molecular weights go from 122.2 to 487.6 (median 230.1).

### 1.5.2 Results and Discussion

#### 1.5.2.1 Reviewing the Toxicity of Neutral Chemicals

Within the data set (which is based on the data set of (Kipka and Di Toro, 2009)), there are 320 cases where two up to eight experimental LC50 values are reported for a specific chemical for the same species. These duplications can be taken as a quantitative measure for

the experimental variance: the LC50 values for the same species and the same chemical differ only in twelve cases by more than one log unit (maximum 1.7 log units, median 0.047 log units). After summarizing the LC50 values for identical chemicals for each species the interspecies variability can be assessed: for 179 out of the 333 different chemicals LC50 values are reported for multiple species (for up to 18 different species). Out of these, the LC50 values of 37 chemicals differ by more than one log unit and the LC50 values of three chemicals differ even by more than two log units (median 0.42). It is crucial to keep this variance inherent in the experimental toxicity data in mind in order to reliably differentiate between chemicals acting via baseline toxicity or via a specific mode of toxic action (as discussed below).

The toxic membrane concentrations according to Eq. 11 have a median of 116 mmol/kg (membrane lipid) and a geometric mean of 105 mmol/kg(membrane lipid) with a standard deviation of the log-normal distribution of 26 to 425 mmol/kg (membrane lipid) (see Fig. 11B). This standard deviation of the log-normal distribution of membrane concentrations based on the data compilation differentiating only between the 333 chemicals is somewhat larger but still comparable to previously determined standard deviations of the log-normal distributions from 41 to 215 mmol/kg (membrane lipid) (19 chemicals measured for guppy (*Poecilia reticulata*), recalculated from (Vaes et al., 1998)), from 40 to 160 mmol/kg (membrane lipid) (65 industrial chemicals measured with the fathead minnow) (Wezel and Opperhuizen, 1995), from 91 to 120 mmol/kg (membrane lipid) (29 chemicals) (McCarty et al., 2013) and 80 to 250 mmol/kg (membrane lipid) (6 chemicals for 3 aquatic organisms) (Endo, 2016; van der Heijden et al., 2015). Similarly, the geometric mean determined in this work is almost identical to the geometric mean of 94.4 mmol/kg (membrane lipid), calculated from the data of (Vaes et al., 1998) (see SI-3, section 2 for details), and also close to the 140 mmol/kg (membrane lipid) reported in (Endo, 2016). The latter and most recent analysis of (Endo, 2016) is based on a single high-quality data set, measured in one lab (van der Heijden et al.,

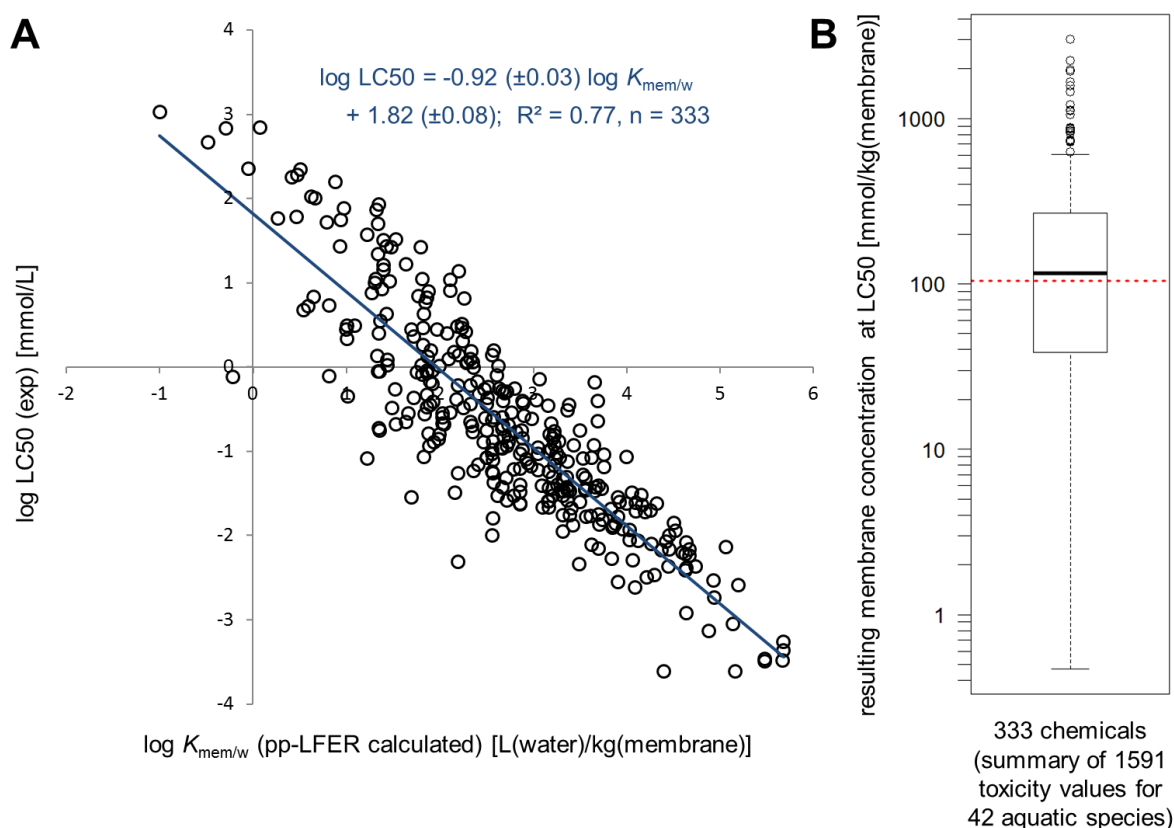
2015), and is based on the distribution of chemicals in the different body compartments rather than on water concentrations. The approach presented in (Endo, 2016) to calculate toxic threshold membrane concentration should be comparably accurate as a calculation based on freely dissolved water concentrations. Interestingly, the study of (Endo, 2016) reports the highest toxic membrane concentrations although the use of nominal water concentrations as done in (Vaes et al., 1998) would potentially overestimate toxic membrane concentrations.

Overall, the re-analysis of the revised data set of (Kipka and Di Toro, 2009) confirms the above cited earlier works and underlines the earlier finding of a toxic membrane concentration for neutral organic chemicals that ranges from 26 to 425 mmol/kg for the investigated 42 organisms and 333 chemicals. Consequently, the pp-LFER for  $\log K_{mw}$  (Eq. 8) is very similar to the pp-LFER calibrated by (Kipka and Di Toro, 2009) describing the partitioning to the ‘target lipid’ (see SI-3, section 3).

The regression equation resulting from Fig. 11A is

$$\log LC50 = -0.92(\pm 0.03) \log K_{mw} + 1.82(\pm 0.08); R^2 = 0.77 \quad (15)$$

The very small deviation of the slope to the ideal value of -1 in the regression Eq. 15 might well be due to the biological variability that is also expressed in the experimental toxicity data. It might also be a hint that  $K_{lipw}$  is not a perfect surrogate for  $K_{mw}$  for every organism, which comes as no surprise. While  $K_{lipw}$  is based on pure phospholipids,  $K_{mw}$  should also account for everything else that makes up real biological membranes beyond pure phospholipids as discussed above. Cholesterol, e.g., is ubiquitous in eukaryotic cell membranes to varying amounts and influences not only the transition phase temperature, but also the sorption characteristics (Endo et al., 2011). Moreover, almost one third of naturally occurring proteins are believed to be located in biological membranes (Tan et al., 2008).



**Figure 11.** A) log LC50 values for 333 different neutral organic chemicals (summarized as described above from 1591 experimental toxicity values measured for 42 aquatic organisms) against their log  $K_{mw}$ , predicted with pp-LFER (Endo et al., 2011). The regression analysis was made with Origin 2015. B) Tukey boxplot of the resulting toxic membrane concentration calculated based on Eq. 11 (the bottom and top of the box represent the first and third quartiles, the thick black line inside box represents the median with 116 mmol/kg (membrane lipid); whiskers set at lowest/highest data point still within 1.5 interquartile range of the lower/upper quartile) with the geometric mean toxic membrane concentration of 105 mmol/kg (membrane lipid) shown as dotted red line. The analysis was done with R version 2.14.2.

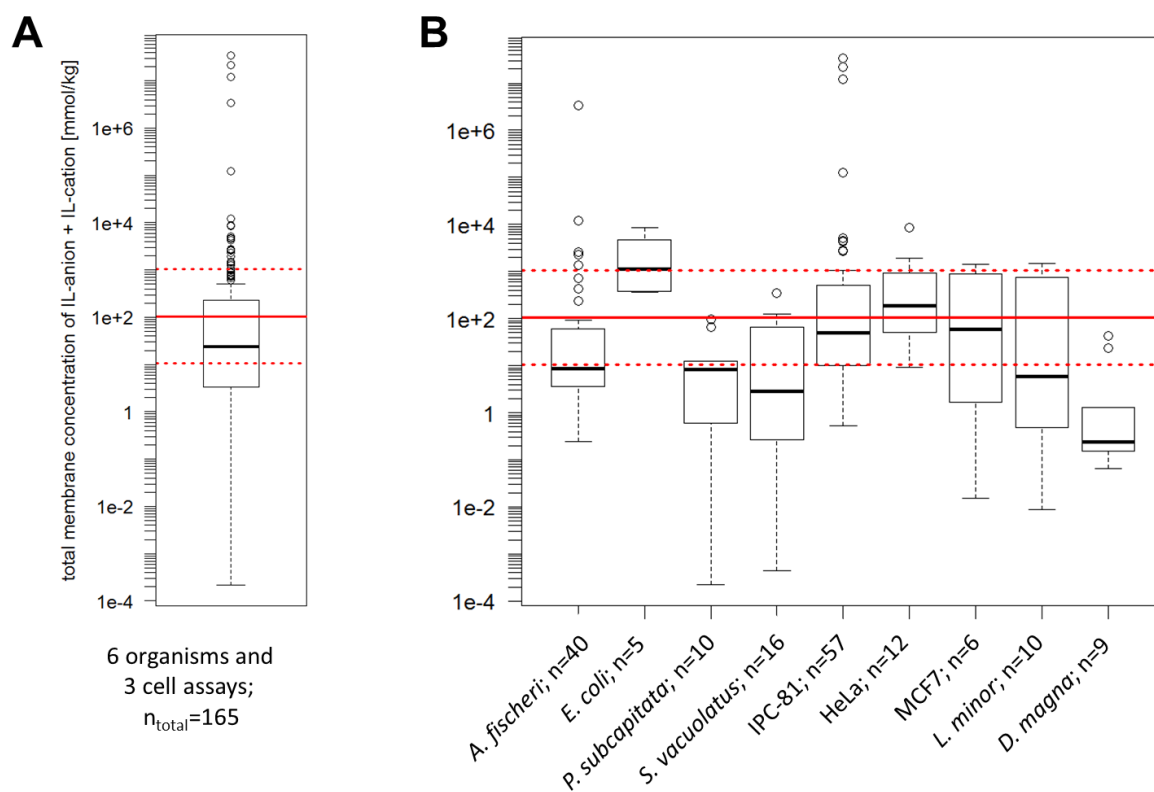
### 1.5.2.2 Toxicity f ILs

The log  $K_{mw}$  [L/kg] values of the cations were calculated to go from -0.80 to 12.06, while the predicted log  $K_{mw}$  [L/kg] values of the corresponding anions go from 0.16 to 3.02 (see SI-3, Table 6). The predicted log  $K_{mw}$  of 12.06 for the trihexyl(tetradecyl)phosphonium cation (P666-14) is far beyond the validation data set of COSMOmic (and it would also experimentally not be feasible to investigate such a high partition coefficient). In principle COSMOmic should be applicable for extrapolations well beyond the validation data set as discussed above, but to be on the safe side P666-14 was excluded from further analysis. This



leaves 38 cations (with a maximum  $\log K_{mw}$  of 8.69) and 6 anions (see SI-3, Table 6) in 77 different salt combinations yielding a total of 165 toxicity values for six different test organisms and 3 cell assays: the bacteria *Aliivibrio fischeri* (n= 40), and *E. coli* (n= 5), the algae *Pseudokirchneriella subcapitata* (n= 10) and *Scenedesmus vacuolatus* (n= 16), the cell lines IPC-81 (rat leukemic cells, n= 57), HeLa cells (n= 12) and MCF7 cells (breast cancer cell line, n= 6), the water plant *Lemna minor* (n= 10) and the water flea *Daphnia magna* (n= 9).

Fig. 12 B does not seem to be a confirmation that the baseline toxicity approach is valid for ILs at the first sight: the additive toxic membrane concentrations for the different IL salts are not close enough to the expected range of baseline toxicity determined for neutral chemicals. However, it cannot be taken for granted that all of the investigated ILs do only exhibit baseline toxicity. Hence, the considerable scatter shown in Fig. 12 B can neither be per se taken as a falsification of the baseline toxicity concept for ILs, but the (partially substantial) deviations from the expected toxic membrane threshold concentration need to be addressed. Taking the concept as a plausible assumption and assuming the same critical membrane concentration range for ILs as determined above for the neutral chemicals, those toxicity values can be tracked down that seem to be based on a specific mode of toxic action and those that are most likely experimental artefacts. Overall, the medians of the total toxic membrane concentrations for ILs range from 0.2 (*D. magna*) to 1100 mmol/kg (membrane lipid) (*E. coli*) and the corresponding geometrical means range from 0.6 (*D. magna*) to 1432 (*E. coli*) mmol/kg(lipid). The toxic membrane concentrations for ILs calculated with Eq. 13 for the different organisms/cell assays have a standard deviation of the log-normal distribution of 0.5 to 1407 mmol/kg (membrane lipid) (for details see SI-3, Table 5).



**Figure 12.** A) Tukey boxplot of the total membrane concentrations resulting from all 165 IL toxicity values for the six different organisms and three cell assays. B) Tukey boxplots of the total membrane concentrations resulting for the six different organisms and three cell assays individually. The bottom and top of the boxes represent the first and third quartiles, the thick black lines inside the boxes represent the medians; whiskers are set at lowest/highest data points still within 1.5 interquartile range of the lower/upper quartile. The red line indicates the geometric mean toxic membrane concentration of 105 mmol/kg (membrane lipid) determined above (Fig. 11 B) for 333 neutral chemicals (42 aquatic organisms), while the dotted red lines correspond to the range of membrane concentrations regarded as baseline toxicity ( $0.1 < \text{TR} < 10$ ), given by Eq. 14. The analysis was done with R version 2.14.2.

As summarized in Table 3, only 22 to 58% of the toxic ratios (TRs) in present data set can be classified as baseline toxicants for the different organisms/cell assays, according to the Eq. 14 (i.e.,  $0.1 < \text{TR} < 10$ ) (see SI-3, Table 9 for all TR values). Only few TRs are classified as less toxic than expected according to the baseline toxicity concept (*P. subcapitata*, *S. vacuolatus* and *D. magna* have no values reported and the remaining organisms/cell assays are below the 25% reported for HeLa, except *E. coli* with 60%). On the other side there is a considerable amount of TRs pointing towards a specific toxic mode of action ( $\text{TR} > 10$ ), up to

78% for *D. magna*; the exception is again *E. coli* with no IL toxicity values being above baseline toxicity.

Fig. 12 B seems to suggest that the values for the organisms *D. magna* and *E. coli* are systematic outliers from the range defined by the baseline toxicity concept. However, this is a false conclusion due to the random selection of tested ILs: in fact, *D. magna* should not be regarded as a specifically sensitive test system, because most of the ILs tested for *D. magna* show similarly elevated TRs as they do for other tested organisms/cell assays. On the other hand *E. coli* should not be regarded as a specifically insensitive test system based on Fig. 12 B, because most of the ILs tested for *E. coli* also show TR values below 10 or even below 0.1 (in this regard the 1-butyl-3-methyl-1H-imidazol-3-ium tetrafluoroborate (IM14 BF<sub>4</sub>) gives the most heterogeneous picture: while it has a TR of 0.095 for *E. coli*, it is classified as baseline toxicant for *A. fischeri*, IPC-81 and HeLa, while exerting excess toxicity to *S. vacuolatus*, *L. minor* and *D. magna*; see SI-3, Table 9 and further discussion below). The ions tested for *E. coli* are not particularly prone to exhibit large differences between nominal and freely-dissolved concentrations (the respective log  $K_{ipw}$  values do not exceed 3.42).

**Table 3.** Summary of the TR analysis via binning the IL toxicity data into chemicals being less toxic than expected according to the baseline toxicity concept ( $TR < 0.1$ ), baseline toxicants ( $0.1 < TR < 10$ ) and specifically acting toxic chemicals ( $TR > 10$ ).

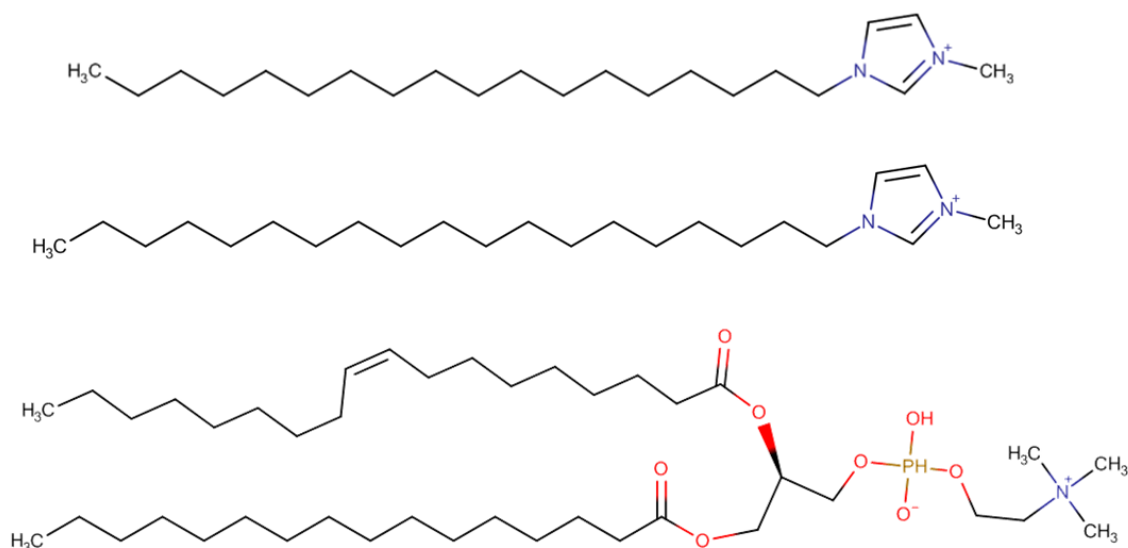
organism/ cell assay	<i>A. fi- scheri</i>	<i>E. coli</i>	<i>P. sub- capitata</i>	<i>S. vacu- olatus</i>	IPC- 81	HeLa	MCF7	<i>L. mi- nor</i>	<i>D. mag- na</i>
TR < 0.1	5	3	0	0	11	3	1	1	0
0.1 < TR < 10	14	2	3	5	31	7	3	3	2
TR > 10	21	0	7	11	15	2	2	6	7

Experimental artefacts are the obvious suspicion for those ILs exerting much less than the expected baseline toxicity ( $TR < 0.1$ ). None of the calculated TRs go below 0.01, except those of the very long chained imidazolium cations 3-methyl-1-octadecyl-1H-imidazol-3-ium (IM1-18) and 3-methyl-1-nonadecyl-1H-imidazol-3-ium (IM1-19), whose TRs go from  $8 \cdot 10^{-4}$  to

$9 \times 10^{-6}$  (data for *A. fischeri* and ICP-81). This is most likely an experimental artefact, because IM1-18 and IM1-19 also exhibit the strongest hydrophobicity of all ions investigated ( $\log K_{lipw} = 8.08$  and  $8.69$ , respectively), probably causing big artefacts due to the use of nominal concentrations instead of freely dissolved concentrations as discussed in (Fischer et al., 2017). The corresponding toxic membrane concentration for the IM1-18 and IM1-19 salts would go from  $1.2 \times 10^5$  to  $3.4 \times 10^7$  mmol/kg (membrane lipid) (determined for *A. fischeri* and IPC-81 cell test), which seems to be far too high. A different, rather statistical argument classifying the IM1-18 and IM1-19 salt toxicity data as artefacts is that out of the 85 toxicity values for imidazolium salts with side chain lengths going from 2 to 14, only 8 have  $TR < 0.1$  and 36 are even supposed to act via specific toxicity – while the majority (41 of the toxicity values) are in the range of baseline toxicity (data for all organisms/cell assays, except MCF-7).

Another, albeit rather hypothetical explanation for the high apparent toxic membrane concentrations of the IM1-18 and IM1-19 salts might be due to the structure of these cations exhibiting the same characteristic features as phospholipids: the IM1-18 and IM1-19 cations have a charged head group and a long apolar tail. They might be less toxic than expected by the general assumptions of the baseline toxicity concept, because they might not alter the physico-chemistry of the membrane as much as other chemicals do. This explanation is purely speculative but seems reasonable based on the structure of the IM1-18 and IM1-19 cations and POPC as a representative phospholipid constituting biological membranes (see Fig. 13). Accordingly, it has been shown that the anesthetic effect to tadpoles for a homologous series of saturated aliphatic alcohols exhibits a cutoff in potency (Pringle et al., 1981). This is seen as a strong hint that the crucial sorption sites determining baseline toxicity are not the pure membranes but hydrophobic pockets of proteins with well-defined volumes. Contrary to this hypothesis we could show that the surfactant-like hexadecyltrimethyl-ammonium-, hexadecylpyridinium- and didecyldimethylammonium cations are all baseline toxicants

(Baumer et al., 2017; Escher et al., 2017). These three structures are similar to IM1-18 and IM1-19. Finally, it could also well be that COSMOmic overpredicted the respective  $K_{mw}$  values, leading to erroneously high membrane concentrations.



**Figure 13.** From top to bottom: structures of 3-methyl-1-octadecyl-1H-imidazol-3-ium (IM1-18;  $\log K_{mw} = 8.08$ ), 3-methyl-1-nonadecyl-1H-imidazol-3-ium cation (IM1-19 cation;  $\log K_{mw} = 8.69$ ) and the phospholipid 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC).

The duration of different toxicity tests ranges from 4 to 24 h, except for *A. fischeri* (30 min) and *L. minor* (7 days) (see SI-3, Table 7 for details). It is well known that the membrane permeability of the ionic species of an ionizable chemical is orders of magnitudes lower than the permeability of the corresponding neutral species (Saparov et al., 2006) which therefore governs the uptake into the organism. In recent work, MDCK cells did not show any uptake of the charged chemical 9,10-dimethoxyanthracene-2-sulfonate within 24 hours (Abele, 2016). This is a strong hint that in the presented cases for ILs the toxicity experiments are (mostly) not conducted long enough for the charged chemicals to reach equilibrium partitioning between the organisms/cell assays membranes' and water. In the case of permanently charged ILs no corresponding neutral species is present so that the uptake into the organism can be very slow. This might also partly explain the very high apparent toxic

membrane concentrations for *A. fischeri* (30 min exposure) and IPC-81 (4 h exposure), which exhibit the lowest TRs, as also discussed above.

While experimental artefacts are likely to explain (at least most of) the toxicity values with  $TR < 0.1$ , this is not the case for toxicity values that are classified to act via a specific mode of action ( $TR > 10$ ). E.g., only 5 out of 23 toxicity values for pyridinium cations fall into the range of baseline toxicity, the remaining 18 toxicity values are all classified as specifically acting toxicants (data for all organisms/cell assays, except *E. coli*). This finding is in line with a recent study that has an albeit different focus: (Peric et al., 2013) showed that the 1-butyl-3-methyl-1H-imidazol-3-ium (IM14) chloride and the 1-butylpyridin-1-ium (Py4) chloride are considerably more toxic to *A. fischeri*, *P. subcapitata*, *L. minor* and IPC-81 than protic ILs (the latter are not part of the present work because they are also prone to ion trapping, which needs a different modelling that brings along additional uncertainties as shortly discussed above and in more detail e.g. in (Baumer et al., 2017)). Interestingly, the toxicity of IM14 in the present data set seems to depend on the nature of the anionic counterion. While IM14 chloride and bromide show specific toxicity for all investigated organisms/cell assays in the data set ( $n = 12$ ), IM14 in combination with the  $BF_4$  anion ( $n = 7$ ) exhibits three values classified as specific toxicity, three values classified as baseline toxicity and one value even classified as less toxic than expected by the baseline toxicity approach. For the remaining three anions (hexafluorophosphate ( $PF_6$ ), bis((trifluoromethyl)sulfonyl)amide (( $CF_3SO_2$ ) $_2N$ ), and dicyanamide (( $CN$ ) $_2N$ )), all of the IM14 toxicity values are classified as baseline toxicity ( $n = 11$ ) or less toxic than expected by the baseline toxicity approach ( $n = 4$ ). While this analysis for the different IM14 salts may also be affected by interspecies differences in sensitivity, it is worth to have a detailed look into the TR analysis for *A. fischeri* and IPC-81, both of which have a complete set of toxicity values for all investigated anions (see Table 4).

**Table 4.** Toxic ratios (TRs) as defined by Eq. 14 for the IM14 cation with the 6 different anions investigated.

	Cl	Br	(CN)2N	BF4	PF6	(CF3SO2)2N
<i>A. fischeri</i>	16	12	2.7	2.1	0.14	0.041
IPC-81	11	12	4.0	2.9	0.14	0.21

Based on the TRs shown in Table 4 it can be speculated that, according to the Pearson acid base concept, the soft IM14 cation forms stronger ion pairs with the soft (CN)2N, BF4, PF6 and (CF3SO2)2N anions than with the hard chloride and bromide anions. These potential ion pairs might hamper a specific mode of toxic action exerted by the unbound IM14 cation. The highest TR of 476005 (which corresponds to the lowest toxic membrane concentration of  $2.2 \cdot 10^{-4}$  mmol/kg (membrane lipid)) is determined for 1-hexyl-3-methyl-1H-imidazol-3-ium (IM16) chloride salt for *P. subcapitata*. IM16 chloride has also high TRs for *A. fischeri* (TR = 23) and *S. vacuolatus* (TR = 4760), but the TR for IPC-81 corresponds to the expectation for baseline toxicity (TR = 8.1). Similar to the pattern discussed above for IM14, all three TRs reported for IM16 bromide indicate excess toxicity, while all remaining four TRs with the anionic counterions BF4, PF6 and (CF3SO2)2N fall into the range of baseline toxicity.

The IL salts containing quinoline cations ( $n = 5$ ) do all show excess toxicity. However, it remains unclear in the present data set whether this finding can be generalized, because all of experimental values are for the ICP-81 cell line test. Interestingly, already Vaes et al. identified the neutral form of quinolone as a chemical exerting excess toxicity towards guppy in their early work advocating the  $K_{lipw}$  over the  $K_{ow}$  for the description of baseline toxicity (Vaes et al., 1998).

In the literature, the anions present in the present data set are discussed to “show none or just trivial” cytotoxicity, except the (CF3SO2)2N anion, which is supposed to “demonstrate a noteworthy effect on the cytotoxicity” (Zhao et al., 2007). This finding can be nicely explained by the baseline toxicity approach: the (CF3SO2)2N anion simply has the highest  $K_{lipw}$  value

---

of all anions investigated in the present work ( $\log K_{\text{lipw}} = 3.02$ ). In fact, none of the 25 toxicity values of IL salts containing the  $(\text{CF}_3\text{SO}_2)_2\text{N}$  anion do show excess toxicity (ten even have  $\text{TRs} < 0.1$ ).



## 1.6 Conclusions and Outlook

### 1.6.1 Prediction of $K_{lipw}(\text{ion})$ with COSMOmic

In order to predict how strongly ions sorb to phospholipid membranes, the membrane potential has to be adequately accounted for, although it cannot be deduced directly from the membrane structure. In the presented enhancement of COSMOmic, the membrane potential has been implemented as a Gaussian-type error function that is optimized with the experimental sorption data, yielding satisfying  $K_{lipw}$  predictions for neutral and ionic compounds. The overall prediction accuracy of the revised COSMOmic model presented in this work is well within the expected accuracy of COSMOtherm, which is reported to be 0.65 to 0.93 log units for the prediction of the partitioning between different liquid/liquid systems for highly diverse data sets (Stenzel et al., 2014). Although there seems to be still some scatter in the prediction especially for cations, the presented enhancement of COSMOmic is, to the best of our knowledge, the first mechanistic model that is able to predict the sorption of both ions and neutral species in such a complex anisotropic phase as membranes are.

In future research, the energy profiles derived with COSMOmic might be used to predict the permeability of ions through membranes. This, however, will need further experimental confirmation, as the permeability of membranes depends on the main resistances (i.e., Gibbs free maxima), while the partition coefficients are more related to the energy minima of the calculated profiles. This is specifically important when it comes to the toxicity of uncouplers, which involves the transfer of ions through energy transducing membranes (Spycher et al., 2008). The presented improvement of COSMOmic for the use with ions may also have implications in drug design, where the ‘lipophilic efficiency’ of ionogenic drugs is still often quantified by using an empirically estimated octanol-water partition coefficient (i.e.,  $\log D_{ow}$ ). The  $K_{lipw}$  predictions of per- and polyfluorinated alkyl chemicals (which are experimentally very difficult to handle) has already successfully been used to enhance the description of their bioaccumulation potential (Ng and Hungerbuehler, 2015).

### 1.6.2 Assessment of Different Models Predicting $K_{lipw}$ (ion)

In terms of  $K_{lipw}$  prediction of neutral chemicals (which were not the focus of this work but were considered for consistency), both the  $K_{ow}$  based model as well as the pp-LFER model are computationally less demanding than COSMOmic. If any mechanistic understanding of the partitioning process of neutral chemicals is desired, the pp-LFER approach will be the more suited one of these two.

However, in terms of  $K_{lipw}$  prediction of charged chemicals, the usage of the pp-LFER approach is constrained due to the lack of a sound mechanistic basis (although the model seems to be applicable for the sorption to anisotropic muscle protein (Henneberger et al., 2016)). The correlation approach with  $\log K_{ow}$ , though being strictly empirical and lacking a strict mechanistic reasoning, performs better than could be expected, at least with regard to anionic chemicals. However, permanent ions, zwitterions and polyvalent ions cannot be handled with the  $K_{ow}$  approach. The mechanistic approach underlying COSMOmic allows the calculation of  $K_{lipw}$  independent of charge and chemical classes and seems to have the potential to handle also zwitter- and divalent ions; it only needs the molecular structure as input. Concerning new pollutants outside the chemical space of the present fitting data set, COSMOmic seems to be the only model that can be used with some confidence to make  $K_{lipw}$  predictions.

### 1.6.3 Applying the Baseline Toxicity Concept on Ions

In this work previous studies on the baseline toxicity concept for neutral organic chemicals could be substantiated. The nonspecific toxicity of neutral chemicals can satisfyingly be explained for a large variety of different aquatic organisms using a toxic membrane concentration range around 100 mmol/kg (membrane lipid) and the equilibrium partition coefficients between phospholipid membranes and water. Applying this concept to permanently charged ILs and assuming independent additive contributions of cationic and anionic chemicals suggests that as many of the investigated toxicity data comply with the

baseline toxicity concept as do seem to exert specific toxicity (about 42% each). The baseline toxicity concept not only enables the differentiation between baseline and excess toxicity but it furthermore points out data that are most likely based on experimental artefacts. Moreover, the analysis of toxic ratios (TRs) is useful to investigate the toxicity of a single IL component (e.g., a cation) when combined with different counterions (i.e., different anions). However, there are still plenty of uncertainties and problems that need to be solved before it can conclusively be argued that baseline toxicity as investigated in this work is the ubiquitous driver of the minimal toxic effect exerted by every chemical (which is the basis for every further analysis): Further toxicity experiments are necessary that take the slower uptake kinetics of ionized organic chemicals into account. The experiments need to be designed in such a way that artefacts caused by the difference between nominal and freely dissolved concentrations can be ruled out (which is specifically important for cell-assay toxicity tests). Studies on membrane permeability of charged chemicals should help to get a clearer picture of the baseline toxicity of permanently charged as well as ionizable chemicals present in charged and non-charged form. The ionizable chemicals are prone to an ion-trapping effect and therefore the difference in the respective membrane permeability of the charged and non-charged form of a chemical becomes crucial, given that the external pH differs from the internal (cytoplasmic) pH. In that regard pH-dependent toxicity tests will help to better understand the ion-trapping effect and thus its implications on baseline toxicity. Finally, also the  $K_{lipw}$  predictions of ionized chemicals need to be further validated, especially in the case of multiply charged chemicals (which are important, e.g., when the toxicity of pharmaceuticals or pesticides is assessed). All of these raised issues will necessarily have to be tackled in order to distinguish reliably between specific and non-specific toxicity and to clarify the exact mechanism of non-specific, baseline toxicity. Nevertheless, the baseline toxicity concept can already be regarded as a useful tool also for charged chemicals, e.g., for regulation purposes, if an estimate of the minimal expected toxicity is needed.

## 1.7 Bibliography

- Abele, C., 2016. Bestimmung von Aufnahmekinetiken organischer Chemikalien in Zellen mittels Fluoreszenzmikroskopie. Bachelor thesis (german). Eberhard Karls University of Tübingen.
- Abraham, M.H., 1993. Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes. *Chem. Soc. Rev.* 22, 73. doi:10.1039/cs9932200073
- Abraham, M.H., Acree, Jr, W.E., 2010a. The transfer of neutral molecules, ions and ionic species from water to wet octanol. *Phys. Chem. Chem. Phys.* 12, 13182. doi:10.1039/c0cp00695e
- Abraham, M.H., Acree, Jr, W.E., 2010b. The transfer of neutral molecules, ions and ionic species from water to ethylene glycol and to propylene carbonate; descriptors for pyridinium cations. *New J. Chem.* 34, 2298. doi:10.1039/c0nj00222d
- Abraham, M.H., Acree, Jr, W.E., 2010c. Equations for the Transfer of Neutral Molecules and Ionic Species from Water to Organic phases. *J. Org. Chem.* 75, 1006–1015. doi:10.1021/jo902388n
- Abraham, M.H., Acree, Jr, W.E., 2010d. Solute Descriptors for Phenoxide Anions and Their Use To Establish Correlations of Rates of Reaction of Anions with Iodomethane. *J. Org. Chem.* 75, 3021–3026. doi:10.1021/jo100292j
- Abraham, M.H., Ibrahim, A., Zissimos, A.M., 2004. Determination of sets of solute descriptors from chromatographic measurements. *J. Chromatogr. A* 1037, 29–47. doi:10.1016/j.chroma.2003.12.004
- Abraham, M.H., Smith, R.E., Luchtefeld, R., Boorem, A.J., Luo, R., Acree, Jr, W.E., 2010. Prediction of solubility of drugs and other compounds in organic solvents. *J. Pharm. Sci.* 99, 1500–1515. doi:10.1002/jps.21922
- Ahlich, R., Furche, F., Hättig, C., Klopper, W., Sierka, M., Weigend, F., 2015. TURBOMOLE, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH: since 2007.
- Armitage, J.M., Arnot, J.A., Wania, F., Mackay, D., 2013. Development and evaluation of a mechanistic bioconcentration model for ionogenic organic chemicals in fish. *Environ. Toxicol. Chem.* 32, 115–128. doi:10.1002/etc.2020
- Armitage, J.M., Wania, F., Arnot, J. a, 2014. Application of Mass Balance Models and the Chemical Activity Concept To Facilitate the Use of in Vitro Toxicity Data for Risk Assessment. *Environ. Sci. Technol.* 48, 9770–9779. doi:10.1021/es501955g
- Avdeef, A., Box, K.J., Comer, J.E.A., Hibbert, C., Tam, K., 1998. pH-Metric logP 10. Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharm. Res.* 15, 209–215. doi:10.1023/A:1011954332221
- Bangham, A.D., Standish, M.M., Watkins, J.C., 1965. Diffusion of univalent ions across the lamellae of swollen phospholipids. *J. Mol. Biol.* 13, 238–IN27. doi:10.1016/S0022-2836(65)80093-6
- Baumer, A., Bittermann, K., Klüver, N., Escher, B.I.I., 2017. Baseline toxicity and ion-trapping models to describe the pH-dependence of bacterial toxicity of pharmaceuticals. *Environ. Sci. Process. Impacts* Accepted M. doi:10.1039/C7EM00099E
- Becke, A.D., 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38, 3098–3100. doi:10.1103/PhysRevA.38.3098
- Chowdhary, J., Harder, E., Lopes, P.E.M., Huang, L., MacKerell, A.D., Roux, B., 2013. A polarizable force field of dipalmitoylphosphatidylcholine based on the classical Drude model for molecular dynamics simulations of lipids. *J. Phys. Chem. B* 117, 9142–60. doi:10.1021/jp402860e
- Clarke, R.J., 2001. The dipole potential of phospholipid membranes and methods for its detection. *Adv. Colloid Interface Sci.* 89–90, 263–281. doi:10.1016/S0001-

- 8686(00)00061-0
- Cowan-Ellsberry, C., Belanger, S., Dorn, P., Dyer, S., McAvoy, D., Sanderson, H., Versteeg, D., Ferrer, D., Stanton, K., 2014. Environmental Safety of the Use of Major Surfactant Classes in North America. *Crit. Rev. Environ. Sci. Technol.* 44, 1893–1993. doi:10.1080/10739149.2013.803777
- Demura, M., Kamo, N., Kobatake, Y., 1987. Binding of lipophilic cations to the liposomal membrane: thermodynamic analysis. *Biochim. Biophys. Acta - Biomembr.* 903, 303–308. doi:10.1016/0005-2736(87)90220-3
- Deneer, J.W., Sinnige, T.L., Seinen, W., Hermens, J.L.M., 1988. The joint acute toxicity to *Daphnia magna* of industrial organic chemicals at low concentrations. *Aquat. Toxicol.* 12, 33–38. doi:10.1016/0166-445X(88)90017-3
- Eichkorn, K., Treutler, O., Öhm, H., Häser, M., Ahlrichs, R., 1995. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* 240, 283–290. doi:10.1016/0009-2614(95)00621-A
- Endo, S., 2016. Re-analysis of narcotic critical body residue data using the equilibrium distribution concept and refined partition coefficients. *Environ. Sci. Process. Impacts* 18, 1024–1029. doi:10.1039/C6EM00180G
- Endo, S., Escher, B.I., Goss, K.-U., 2011. Capacities of Membrane Lipids to Accumulate Neutral Organic Chemicals. *Environ. Sci. Technol.* 45, 5912–5921. doi:10.1021/es200855w
- Endo, S., Goss, K., 2014. Applications of Polyparameter Linear Free Energy Relationships in Environmental Chemistry. *Environ. Sci. Technol.* 48, 12477–12491. doi:10.1021/es503369t
- Escher, B.I., Baumer, A., Bittermann, K., Henneberger, L., König, M., Kühnert, C., Klüver, N., 2017. General baseline toxicity QSAR for nonpolar, polar and ionisable chemicals and their mixtures in the bioluminescence inhibition assay with *Aliivibrio fischeri*. *Environ. Sci. Process. Impacts* 19, 414–428. doi:10.1039/C6EM00692B
- Escher, B.I., Baumgartner, R., Koller, M., Treyer, K., Lienert, J., McArdell, C.S., 2011. Environmental toxicology and risk assessment of pharmaceuticals from hospital wastewater. *Water Res.* 45, 75–92. doi:10.1016/j.watres.2010.08.019
- Escher, B.I., Eggen, R.I.L., Schreiber, U., Schreiber, Z., Vye, E., Wisner, B., Schwarzenbach, R.P., 2002. Baseline Toxicity (Narcosis) of Organic Chemicals Determined by In Vitro Membrane Potential Measurements in Energy-Transducing Membranes. *Environ. Sci. Technol.* 36, 1971–1979. doi:10.1021/es015844c
- Escher, B.I., Hermens, J.L.M., 2002. Modes of Action in Ecotoxicology: Their Role in Body Burdens, Species Sensitivity, QSARs, and Mixture Effects. *Environ. Sci. Technol.* 36, 4201–4217. doi:10.1021/es015848h
- Escher, B.I., Schwarzenbach, R.P., 2002. Mechanistic studies on baseline toxicity and uncoupling of organic compounds as a basis for modeling effective membrane concentrations in aquatic organisms. *Aquat. Sci.* 64, 20–35. doi:10.1007/s00027-002-8052-2
- Escher, B.I., Schwarzenbach, R.P., 1996. Partitioning of Substituted Phenols in Liposome–Water, Biomembrane–Water, and Octanol–Water Systems. *Environ. Sci. Technol.* 30, 260–270. doi:10.1021/es9503084
- Escher, B.I., Schwarzenbach, R.P., Westall, J.C., 2000. Evaluation of Liposome–Water Partitioning of Organic Acids and Bases. 1. Development of a Sorption Model. *Environ. Sci. Technol.* 34, 3954–3961. doi:10.1021/es0010709
- Escher, B.I., Sigg, L., 2004. Chemical Speciation of Organics and of Metals at Biological Interphases, in: v. Leeuwen, H.P., Köster, W. (Eds.), *Physicochemical Kinetics and Transport at Biointerfaces*. John Wiley & Sons, Ltd, Chichester, UK, pp. 205–269. doi:10.1002/0470094044.ch5

- Escher, B.I., Snozzi, M., Schwarzenbach, R.P., 1996. Uptake, Speciation, and Uncoupling Activity of Substituted Phenols in Energy Transducing Membranes. *Environ. Sci. Technol.* 30, 3071–3079. doi:10.1021/es960153f
- Fischer, F.C., Henneberger, L., König, M., Bittermann, K., Linden, L., Goss, K.-U., Escher, B.I., 2017. Modeling Exposure in the Tox21 in Vitro Bioassays. *Chem. Res. Toxicol.* 30, 1197–1208. doi:10.1021/acs.chemrestox.7b00023
- Flewelling, R.F., Hubbell, W.L., 1986a. The membrane dipole potential in a total membrane potential model. Applications to hydrophobic ion interactions with membranes. *Biophys. J.* 49, 541–552. doi:10.1016/S0006-3495(86)83664-5
- Flewelling, R.F., Hubbell, W.L., 1986b. Hydrophobic ion interactions with membranes. Thermodynamic analysis of tetraphenylphosphonium binding to vesicles. *Biophys. J.* 49, 531–540. doi:10.1016/S0006-3495(86)83663-3
- Franco, A., Ferranti, A., Davidsen, C., Trapp, S., 2010. An unexpected challenge: ionizable compounds in the REACH chemical space. *Int. J. Life Cycle Assess.* 15, 321–325. doi:10.1007/s11367-010-0165-6
- Franklin, J.C., Cafiso, D.S., 1993. Internal electrostatic potentials in bilayers: measuring and controlling dipole potentials in lipid vesicles. *Biophys. J.* 65, 289–99. doi:10.1016/S0006-3495(93)81051-8
- Franks, N.P., Abraham, M.H., Lieb, W.R., 1993. Molecular organization of liquid n-octanol: An X-ray diffraction analysis. *J. Pharm. Sci.* 82, 466–470. doi:10.1002/jps.2600820507
- Franks, N.P., Lieb, W.R., 1990. Mechanisms of General Anesthesia. *Environ. Health Perspect.* 87, 199. doi:10.2307/3431025
- Fruttero, R., Caron, G., Fornatto, E., Boschi, D., Ermondi, G., Gasco, A., Carrupt, P.A., Testa, B., 1998. Mechanisms of liposomes/water partitioning of (p- methylbenzyl)alkylamines. *Pharm. Res.* 15, 1407–1413. doi:10.1023/A:1011953622052
- Gazitúa, M., Tapia, R.A., Contreras, R., Campodónico, P.R., 2014. Mechanistic pathways of aromatic nucleophilic substitution in conventional solvents and ionic liquids. *New J. Chem.* 38, 2611. doi:10.1039/c4nj00130c
- Goss, K., 2008. The pKa Values of PFOA and Other Highly Fluorinated Carboxylic Acids. *Environ. Sci. Technol.* 42, 456–458. doi:10.1021/es702192c
- Goss, K.-U., Endo, S., 2016. Comment on “Application of the Activity Framework for Assessing Aquatic Ecotoxicology Data for Organic Chemicals.” *Environ. Sci. Technol.* 50, 4139–4140. doi:10.1021/acs.est.5b05534
- Gurtovenko, A. a, Patra, M., Karttunen, M., Vattulainen, I., 2004. Cationic DMPC/DMTAP lipid bilayers: molecular dynamics study. *Biophys. J.* 86, 3461–72. doi:10.1529/biophysj.103.038760
- Harder, E., Mackerell, A.D., Roux, B., 2009. Many-body polarization effects and the membrane dipole potential. *J. Am. Chem. Soc.* 131, 2760–1. doi:10.1021/ja806825g
- Hefter, G., Marcus, Y., Waghorne, W.E., 2002. Enthalpies and Entropies of Transfer of Electrolytes and Ions from Water to Mixed Aqueous Organic Solvents. *Chem. Rev.* 102, 2773–2836. doi:10.1021/cr010031s
- Henneberger, L., Goss, K.-U., Endo, S., 2016. Partitioning of Organic Ions to Muscle Protein: Experimental Data, Modeling, and Implications for in Vivo Distribution of Organic Ions. *Environ. Sci. Technol.* 50, 7029–7036. doi:10.1021/acs.est.6b01417
- Honig, B.H., Hubbell, W.L., Flewelling, R.F., 1986. Electrostatic interactions in membranes and proteins. *Annu. Rev. Biophys. Chem.* 15, 163–93. doi:10.1146/annurev.bb.15.060186.001115
- Houde, M., Martin, J.W., Letcher, R.J., Solomon, K.R., Muir, D.C.G., 2006. Biological Monitoring of Polyfluoroalkyl Substances: A Review. *Environ. Sci. Technol.* 40, 3463–3473. doi:10.1021/es052580b
- Ingram, T., Storm, S., Kloss, L., Mehling, T., Jakobtorweihen, S., Smirnova, I., 2013.

- Prediction of micelle/water and liposome/water partition coefficients based on molecular dynamics simulations, COSMO-RS, and COSMOmic. *Langmuir* 29, 3527–37. doi:10.1021/la305035b
- Inoue, T., Iwanaga, T., Fukushima, K., Shimozawa, R., 1988. Effect of sodium octanoate and sodium perfluorooctanoate on gel-to-liquid-crystalline phase transition of dipalmitoylphosphatidylcholine vesicle membrane. *Chem. Phys. Lipids* 46, 25–30. doi:10.1016/0009-3084(88)90109-0
- Inoue, T., Miyakawa, K., Shimozawa, R., 1986. Interaction of surfactants with vesicle membrane of dipalmitoylphosphatidylcholine. Effect on gel-to-liquid-crystalline phase transition of lipid bilayer. *Chem. Phys. Lipids* 42, 261–270. doi:10.1016/0009-3084(86)90085-X
- Jafvert, C.T., Westall, J.C., Grieder, E., Schwarzenbach, R.P., 1990. Distribution of hydrophobic ionogenic organic compounds between octanol and water: organic acids. *Environ. Sci. Technol.* 24, 1795–1803. doi:10.1021/es00082a002
- Jakobtorweihen, S., Ingram, T., Smirnova, I., 2013. Combination of COSMOmic and molecular dynamics simulations for the calculation of membrane-water partition coefficients. *J. Comput. Chem.* 34, 1332–1340. doi:10.1002/jcc.23262
- Jakobtorweihen, S., Zuniga, a. C., Ingram, T., Gerlach, T., Keil, F.J., Smirnova, I., 2014. Predicting Solute Partitioning in Lipid Bilayers: Free Energies and Partition Coefficients from Molecular Dynamics Simulations and COSMOmic. *J. Chem. Phys.* 141, 45102. doi:10.1063/1.4890877
- Johnson, C.A., Westall, J.C., 1990. Effect of pH and potassium chloride concentration on the octanol-water distribution of methylanilines. *Environ. Sci. Technol.* 24, 1869–1875. doi:10.1021/es00082a014
- Kaiser, S., Escher, B.I., 2006. The evaluation of liposome-water partitioning of 8-hydroxyquinolines and their copper complexes. *Environ. Sci. Technol.* 40, 1784–1791.
- Kipka, U., Di Toro, D.M., 2009. Technical basis for polar and nonpolar narcotic chemicals and polycyclic aromatic hydrocarbon criteria. III. A polyparameter model for target lipid partitioning. 28, 1429–1438. doi:10.1897/08-364.1
- Klamt, A., 2015. COSMO-RS for aqueous solvation and interfaces. *Fluid Phase Equilib.* 407, 152–158. doi:10.1016/j.fluid.2015.05.027
- Klamt, A., Huniar, U., Spycher, S., Keldenich, J., 2008. COSMOmic: A Mechanistic Approach to the Calculation of Membrane–Water Partition Coefficients and Internal Distributions within Membranes and Micelles. *J. Phys. Chem. B* 112, 12148–12157. doi:10.1021/jp801736k
- Krämer, S.D., Wunderli-Allenspach, H., 2001. Physicochemical properties in pharmacokinetic lead optimization. *Farm.* 56, 145–148. doi:10.1016/S0014-827X(01)01028-X
- Lehmler, H., Xie, W., Bothun, G., Bummer, P.M., Knutson, B.L., 2006. Mixing of perfluorooctanesulfonic acid (PFOS) potassium salt with dipalmitoyl phosphatidylcholine (DPPC). *Colloids Surfaces B Biointerfaces* 51, 25–29. doi:10.1016/j.colsurfb.2006.05.013
- Li, X., Brownawell, B.J., 2010. Quaternary Ammonium Compounds in Urban Estuarine Sediment Environments - A Class of Contaminants in Need of Increased Attention? *Environ. Sci. Technol.* 44, 7561–7568. doi:10.1021/es1011669
- Loidl-Stahlhofen, A., Hartmann, T., Schöttner, M., Röhring, C., Brodowsky, H., Schmitt, J., Keldenich, J., 2001. Multilamellar liposomes and solid-supported lipid membranes (TRANSIL): screening of lipid-water partitioning toward a high-throughput scale. *Pharm. Res.* 18, 1782–8. doi:10.1023/A:1013343117979
- Lombardi, D., Cuenoud, B., Krämer, S.D., 2009. Lipid membrane interactions of indacaterol and salmeterol: do they influence their pharmacological properties? *Eur. J. Pharm. Sci.*

- 38, 533–47. doi:10.1016/j.ejps.2009.10.001
- McCarty, L.S., Arnot, J.A., Mackay, D., 2013. Evaluation of critical body residue data for acute narcosis in aquatic organisms. *Environ. Toxicol. Chem.* 32, 2301–2314. doi:10.1002/etc.2289
- McCarty, L.S., Mackay, D., 1993. Enhancing ecotoxicological modeling and assessment. *Body Residues and Modes Of Toxic Action. Environ. Sci. Technol.* 27, 1718–1728. doi:10.1021/es00046a001
- Mouritsen, O.G., Andersen, H.K., Andersen, J.S., Davidsen, J., Nielsen, L.K., Jørgensen, K., 2001. Pharmacokinetic Optimization in Drug Research, in: Testa, B., van de Waterbeemd, H., Folkers, G., Guy, R. (Eds.), *PHARMACOKINETIC OPTIMIZATION IN DRUG RESEARCH*. Verlag Helvetica Chimica Acta, Zürich, pp. 32–49. doi:10.1002/9783906390437
- Müller, M.T., Zehnder, A.J.B., Escher, B.I., 1999. Liposome-water and octanol-water partitioning of alcohol ethoxylates. *Environ. Toxicol. Chem.* 18, 2191–2198. doi:10.1002/etc.5620181011
- Neuwoehner, J., Fenner, K., Escher, B.I., 2009. Physiological Modes of Action of Fluoxetine and its Human Metabolites in Algae. *Environ. Sci. Technol.* 43, 6830–6837. doi:10.1021/es9005493
- Ng, C.A., Hungerbuehler, K., 2015. Exploring the Use of Molecular Docking to Identify Bioaccumulative Perfluorinated Alkyl Acids (PFAAs). *Environ. Sci. Technol.* 49, 12306–12314. doi:10.1021/acs.est.5b03000
- Ng, C.A., Hungerbuehler, K., 2013. Bioconcentration of Perfluorinated Alkyl Acids: How Important Is Specific Binding? *Environ. Sci. Technol.* 47, 130618120735000. doi:10.1021/es400981a
- Olson, F., Hunt, C.A., Szoka, F.C., Vail, W.J., Papahadjopoulos, D., 1979. Preparation of liposomes of defined size distribution by extrusion through polycarbonate membranes. *Biochim. Biophys. Acta - Biomembr.* 557, 9–23. doi:10.1016/0005-2736(79)90085-3
- Pagliara, A., Carrupt, P., Caron, G., Gaillard, P., Testa, B., 1997. Lipophilicity Profiles of Ampholytes. *Chem. Rev.* 97, 3385–3400. doi:10.1021/cr9601019
- Paloncýová, M., Devane, R., Murch, B., Berka, K., Otyepka, M., 2014a. Amphiphilic Drug-Like Molecules Accumulate in a Membrane below the Head Group Region. *J. Phys. Chem. B* 118, 1030–39. doi:10.1021/jp4112052
- Paloncýová, M., Fabre, G., DeVane, R.H., Trouillas, P., Berka, K., Otyepka, M., 2014b. Benchmarking of Force Fields for Molecule–Membrane Interactions. *J. Chem. Theory Comput.* 10, 4143–4151. doi:10.1021/ct500419b
- Perdew, J., 1986. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* 33, 8822–8824. doi:10.1103/PhysRevB.33.8822
- Peric, B., Sierra, J., Martí, E., Cruañas, R., Garau, M.A., Arning, J., Bottin-Weber, U., Stolte, S., 2013. (Eco)toxicity and biodegradability of selected protic and aprotic ionic liquids. *J. Hazard. Mater.* 261, 99–105. doi:10.1016/j.jhazmat.2013.06.070
- Platts, J.A., Butina, D., Abraham, M.H., Hersey, A., 1999. Estimation of Molecular Linear Free Energy Relation Descriptors Using a Group Contribution Approach. *J. Chem. Inf. Model.* 39, 835–845. doi:10.1021/ci980339t
- Plempers van Balen, G., Caron, G., Ermondi, G., Pagliara, A., Grandi, T., Bouchard, G., Fruttero, R., Carrupt, P.-A., Testa, B., 2001. Lipophilicity Behaviour of the Zwitterionic Antihistamine Cetirizine in Phosphatidylcholine Liposomes/Water Systems. *Pharm. Res.* 18, 694–701. doi:10.1023/A:1011049830615
- Pringle, M.J., Brown, K.B., Miller, K.W., 1981. Can the lipid theories of anesthesia account for the cutoff in anesthetic potency in homologous series of alcohols? *Mol. Pharmacol.* 19, 49–55.



- Saifullah, M., Ye, S., Grubbs, L.M., De La Rosa, N.E., Acree, Jr, W.E., Abraham, M.H., 2011. Abraham Model Correlations for Transfer of Neutral Molecules to Tetrahydrofuran and to 1,4-Dioxane, and for Transfer of Ions to Tetrahydrofuran. *J. Solution Chem.* 40, 2082–2094. doi:10.1007/s10953-011-9776-1
- Saparov, S.M., Antonenko, Y.N., Pohl, P., 2006. A New Model of Weak Acid Permeation through Membranes Revisited: Does Overton Still Rule? *Biophys. J.* 90, L86–L88. doi:10.1529/biophysj.106.084343
- Schäfer, A., Huber, C., Ahlrichs, R., 1994. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 100, 5829. doi:10.1063/1.467146
- Schmitt, W., 2008. General approach for the calculation of tissue to plasma partition coefficients. *Toxicol. In Vitro* 22, 457–467. doi:10.1016/j.tiv.2007.09.010
- Schwarzenbach, R.P., Gschwend, P., Imboden, D., 2003. Environmental organic chemistry, 2nd editio. ed, *Journal of Chemical Education*. John Wiley & Sons, Hoboken, New Jersey.
- Spycher, S., Smejtek, P., Netzeva, T.I., Escher, B.I., 2008. Toward a Class-Independent Quantitative Structure–Activity Relationship Model for Uncouplers of Oxidative Phosphorylation. *Chem. Res. Toxicol.* 21, 911–927. doi:10.1021/tx700391f
- Stenzel, A., Goss, K.-U., Endo, S., 2014. Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOtherm, ABSOLV, and SPARC. *Environ. Toxicol. Chem.* 33, 1537–1543. doi:10.1002/etc.2587
- Tan, S., Tan, H.T., Chung, M.C.M., 2008. Membrane proteins and membrane proteomics. *Proteomics* 8, 3924–3932. doi:10.1002/pmic.200800597
- Tang, J.Y.M., McCarty, S., Glenn, E., Neale, P. a, Warne, M.S.J., Escher, B.I., 2013. Mixture effects of organic micropollutants present in water: Towards the development of effect-based water quality trigger values for baseline toxicity. *Water Res.* 47, 3300–3314. doi:10.1016/j.watres.2013.03.011
- Thomae, A. V, Koch, T., Panse, C., Wunderli-Allenspach, H., Krämer, S.D., 2007. Comparing the lipid membrane affinity and permeation of drug-like acids: the intriguing effects of cholesterol and charged lipids. *Pharm. Res.* 24, 1457–1472. doi:10.1007/s11095-007-9263-y
- Thomas, P., Dawick, J., Lampi, M., Lemaire, P., Presow, S., van Egmond, R., Arnot, J.A., Mackay, D., Mayer, P., Galay Burgos, M., 2015. Application of the Activity Framework for Assessing Aquatic Ecotoxicology Data for Organic Chemicals. *Environ. Sci. Technol.* 49, 12289–12296. doi:10.1021/acs.est.5b02873
- Thuy Pham, T.P., Cho, C.-W., Yun, Y.-S., 2010. Environmental fate and toxicity of ionic liquids: A review. *Water Res.* 44, 352–372. doi:10.1016/j.watres.2009.09.030
- Ulrich S.; Brown, T.N.; Watanabe, N.; Bronner, G.; Abraham, M.H.; Goss, K.-U., N. E., 2017. UFZ-LSER database v 3.2 [Internet].
- Vaes, W.H.J., Ramos, E.U., Verhaar, H.J.M., Hermens, J.L.M., 1998. Acute toxicity of nonpolar versus polar narcosis: Is there a difference? *Environ. Toxicol. Chem.* 17, 1380–1384. doi:10.1002/etc.5620170723
- Vainio, M.J., Johnson, M.S., 2007. Generating Conformer Ensembles Using a Multiobjective Genetic Algorithm. *J. Chem. Inf. Model.* 47, 2462–2474. doi:10.1021/ci6005646
- van der Heijden, S.A., Hermens, J.L.M., Sinnige, T.L., Mayer, P., Gilbert, D., Jonker, M.T.O., 2015. Determining High-Quality Critical Body Residues for Multiple Species and Chemicals by Applying Improved Experimental Design and Data Interpretation Concepts. *Environ. Sci. Technol.* 49, 1879–1887. doi:10.1021/es505078r
- van der Heijden, S.A., Jonker, M.T.O., 2009. Evaluation of Liposome–Water Partitioning for Predicting Bioaccumulation Potential of Hydrophobic Organic Chemicals. *Environ. Sci. Technol.* 43, 8854–8859. doi:10.1021/es902278x

- van Wezel, A.P., Cornelissen, G., van Miltenburg, J.K., Opperhuizen, A., 1996. Membrane burdens of chlorinated benzenes lower the main phase transition temperature in dipalmitoyl-phosphatidylcholine vesicles: Implications for toxicity by narcotic chemicals. *Environ. Toxicol. Chem.* 15, 203–212. doi:10.1002/etc.5620150219
- Verhaar, H.J.M., van Leeuwen, C.J., Hermens, J.L.M., 1992. Classifying environmental pollutants. 1: Structure-activity relationships for prediction of aquatic toxicity. *Chemosphere* 25, 471–491. doi:10.1016/0045-6535(92)90280-5
- Wachter, W., Buchner, R., Hefter, G., 2006. Hydration of Tetraphenylphosphonium and Tetraphenylborate Ions by Dielectric Relaxation Spectroscopy. *J. Phys. Chem. B* 110, 5147–5154. doi:10.1021/jp057189r
- Wang, L., 2012. Measurements and Implications of the Membrane Dipole Potential. *Annu. Rev. Biochem.* 81, 615–635. doi:10.1146/annurev-biochem-070110-123033
- Wezel, A.P. van, Opperhuizen, A., 1995. Narcosis Due to Environmental Pollutants in Aquatic Organisms: Residue-Based Toxicity, Mechanisms, and Membrane Burdens. *Crit. Rev. Toxicol.* 25, 255–279. doi:10.3109/10408449509089890
- Zhao, D., Liao, Y., Zhang, Z., 2007. Toxicity of Ionic Liquids. *CLEAN – Soil, Air, Water* 35, 42–48. doi:10.1002/clen.200600015
- Zhao, Y.H., Abraham, M.H., 2005. Octanol/Water Partition of Ionic Species, Including 544 Cations. *J. Org. Chem.* 70, 2633–2640. doi:10.1021/jo048078b
- Zuccato, E., Castiglioni, S., Bagnati, R., Chiabrando, C., Grassi, P., Fanelli, R., 2008. Illicit drugs, a novel group of environmental contaminants. *Water Res.* 42, 961–968. doi:10.1016/j.watres.2007.09.010

## 1.8 Abbreviations

COSMO	conductor-like screening model
COSMOmic	COSMO-RS for micelles
COSMO-RS	conductor-like screening model for real solvents
DMPC	1,2-dimyristoyl-sn-glycero-3-phosphocholine
$D_{ow}$	sum of the neutral fraction times the respective $K_{ow}$ plus the ionized fraction times the respective $K_{ow}$
IL	ionic liquid
$K_{lipw}$	liposome-water partition coefficient
$K_{mw}$	(biological) membrane-water partition coefficient
$K_{ow}$	octanol-water partition coefficient
MD	molecular dynamics simulation
MW	molecular weight
PFOA	perfluorooctanoic acid
PFOS	heptadecafluoro-1-octanesulfonic acid
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
pp-LFER	poly-parameter linear free energy relationship
QSAR	quantitative structure activity relationship
RMSE	root-mean-square error
SI	supporting information, enumerated according to the order of appearance (see Preface)
TR	toxic ratio
$\Psi_d$	internal membrane dipole potential

## 2. Abstracts of original publications

### 2.1 Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals using the Mechanistic Model

#### COSMOmic

Kai Bittermann, Simon Spycher, Satoshi Endo, Larissa Pohler, Uwe Huniar, Kai-Uwe Goss and Andreas Klamt

*Journal of Physical Chemistry B.* **2014**, 118 (51) 14833–42. doi:10.1021/jp509348a

#### ABSTRACT

The partition coefficient of chemicals from water to phospholipid membrane,  $K_{lipw}$ , is of central importance for various fields. For neutral organic molecules,  $\log K_{lipw}$  correlates with the log of bulk solvent-water partition coefficients such as the octanol-water partition coefficient. However, this is not the case for charged compounds, for which a mechanistic modelling approach is highly necessary. In this work, we extend the model COSMOmic, which adapts the COSMO-RS theory for anisotropic phases and has been shown to reliably predict  $K_{lipw}$  for neutral compounds, to the use of ionic compounds. To make the COSMOmic model applicable for ionic solutes, we implemented the internal membrane dipole potential in COSMOmic. We empirically optimized the potential with experimental  $K_{lipw}$  data of 161 neutral and 75 ionic compounds, yielding potential shapes that agree well with experimentally determined potentials from the literature. This model refinement has no negative effect on the prediction accuracy of neutral compounds (root mean square error, RMSE = 0.62 log units), while it highly improves the prediction of ions (RMSE = 0.70 log units). The refined COSMOmic is, to our knowledge, the first mechanistic model that predicts  $K_{lipw}$  of both ionic and neutral species with accuracies better than 1 log unit.

## 2.2 Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds

Kai Bittermann, Simon Spycher and Kai-Uwe Goss

*Chemosphere*. 2016, (144) 382–391. doi:10.1016/j.chemosphere.2015.08.065

### ABSTRACT

A large fraction of commercially used chemicals is ionizable. This results in the need for mechanistic models to describe the physicochemical properties of ions, like the membrane-water partition coefficient ( $K_{mw}$ ), which is related to toxicity and bioaccumulation. In this work we compare 3 different and already existing modelling approaches to describe the liposome-water partition coefficient ( $K_{lipw}$ ) of organic ions, including 36 cations, 56 anions, 2 divalent cations and 2 zwitterions (plus 207 neutral compounds for ensuring model consistency). 1) The empirical correlation with the octanol-water partition coefficient of the corresponding neutral species yielded better results for the prediction of anions (RMSE=0.79) than for cations (RMSE=1.14). Though describing most anions reasonably well, the lack of mechanistic basis and the poor performance for cations constrain the usage of this model. 2) The polyparameter linear free energy relationship (pp-LFER) model performs worse (RMSE=1.26/1.12 for anions/cations). The different physicochemical environments, due to different sorption depths into the membrane of the different species, cannot be described with a single pp-LFER model. 3) COSMOmic is based on quantum chemistry and fluid phase thermodynamics and has the widest applicability domain. It was the only model applicable for multiply charged ions and gave the best results for anions (RMSE=0.66) and cations (RMSE=0.71). We expect COSMOmic to contribute to a better estimation of the environmental risk of ionizable emerging pollutants.

## 2.3 Erratum - Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds

Kai Bittermann, Simon Spycher and Kai-Uwe Goss

*Chemosphere*. 2016, (179) 405–406. doi:10.1016/j.chemosphere.2017.03.132

### SUMMARY

Due to a transcription error 26 pp-LFER solute descriptors of the charged compounds listed in the supporting information (SI) available online before 10.13.2015, Table 8 are not correct: the B and J<sup>-</sup> values for the cationic forms of acebutolol, alprenolol, bupranolol, labetalol, nadolol, oxprenolol, pindolol, toliprolol, ceterizine, chlorpromazine, hydroxyzine and morantel should be 0. Likewise, the J<sup>+</sup> values for the anionic forms of 2,4-dichlorophenoxyacetic acid and acrivastine should be 0. These errors have been corrected in the SI now available (<http://www.sciencedirect.com/science/article/pii/S0045653515300655>).

The simultaneous errors in the solute descriptors B and J<sup>-</sup> of the cations almost cancel, while the two errors in the J<sup>+</sup> values for the anions do not have a big influence on the multi linear regression. Eq. 6 in the article changes on the second digit of the regression coefficient for J<sup>-</sup> when the corrected solute descriptors are used and in Figure 3 the values of 2,4-dichlorophenoxyacetic acid anion and acrivastine anion are shifted by 1.80 and 3.54 log units, respectively (all other predictions change less than 0.9 log unit).

The observation in the published article that Eqs. 6 and 7 differ substantially is still valid for the two corrected Eqs. 6 and 7 above. Thus, the claim that one pp-LFER is not enough to describe the heterogeneous membrane-water partitioning system adequately for ions and neutral compounds as discussed in the article can be maintained. However, it also has to be pointed out that for some solute descriptors the values for ions are much bigger than for the neutral compounds. E.g. B and S for the 42 anions go up to values of 4.39 and 16.59, respectively, while B and S for the 207 neutral compounds do not exceed the values of 2.19 and 3.29, respectively. In order to cover the full physicochemical space occupied by the solute descriptors, it seems to be more meaningful to fit ions and neutral compounds together for the derivation of a pp-LFER equation (although this is not the procedure recommended by Abraham et al. (Abraham and Acree, Jr, 2010a, 2010b, 2010c; Saifullah et al., 2011; Zhao and Abraham, 2005)).

## 2.4 Assessing the toxicity of ionic liquids – Application of the Critical Membrane Concentration approach

Kai Bittermann, Kai-Uwe Goss

*Chemosphere*. 2017. doi: 10.1016/j.chemosphere.2017.05.097.

### ABSTRACT

Charged organic chemicals are a prevailing challenge for toxicity modelling. In this contribution we strive to recapitulate the lessons learned from the well-known modelling of narcosis (or baseline toxicity) of neutral chemicals and apply the concept to charged chemicals. First we reevaluate the organism- and chemical independent critical membrane concentration causing 50% mortality,  $c_{\text{mem}}^{\text{tox}}$ , based on a critical revision of a previously published toxicity dataset for neutral chemicals. In accordance to values reported in the literature we find a mean value for  $c_{\text{mem}}^{\text{tox}}$  of roughly 100 mmol/kg (membrane lipid) for a broad variety of 42 aquatic organisms (333 different chemicals), albeit with a considerable scatter. Then we apply this concept to permanently charged ionic liquids (ILs). Using COSMOmic, a quantum mechanically based mechanistic model that makes use of the COSMO-RS theory, we predict membrane-water partition coefficients ( $K_{\text{mem/w}}$ ) of the anionic and cationic IL components. Doing so,  $c_{\text{mem}}^{\text{tox}}(\text{total})$  for permanently charged ILs can be estimated assuming independent, concentration additive contributions of the cationic and its respective anionic species. The resulting values for some of the toxicity data for ionic liquids are consistent with the expected range for baseline toxicity for neutral chemicals while other values are consistently greater or smaller. Based on the calculation of toxic ratios we identify ILs that exert a specific mode of toxic action. Limitations of the modelling approach especially but not exclusively due to the use of nominal concentrations instead of freely-dissolved concentrations in the published literature are critically discussed.

## 2.5 Modeling Exposure in the Tox21 in Vitro Bioassays

Fabian Fischer, Luise Henneberger, Maria König, Kai Bittermann, Lukas Linden, Kai-Uwe Goss and Beate Escher

*Chemical Research in Toxicology*. 2017, 30 (5) 1197-1208.

doi: 10.1021/acs.chemrestox.7b00023

### ABSTRACT

High-throughput *in vitro* bioassays are becoming increasingly important in the risk characterization of anthropogenic chemicals. Large databases gather nominal effect concentrations ( $C_{\text{nom}}$ ) for diverse modes of action. However, the biologically effective concentration can substantially deviate due to differences in chemical partitioning. In this study, we modeled freely dissolved ( $C_{\text{free}}$ ), cellular ( $C_{\text{cell}}$ ), and membrane concentrations ( $C_{\text{mem}}$ ) in the Tox21 GeneBLAzer bioassays for a set of neutral and ionogenic organic chemicals covering a large physicochemical space. Cells and medium constituents were experimentally characterized for their lipid and protein content, and partition constants were either collected from the literature or predicted by mechanistic models. The chemicals exhibited multifaceted partitioning to proteins and lipids with distribution ratios spanning over 8 orders of magnitude. Modeled  $C_{\text{free}}$  deviated over 5 orders of magnitude from  $C_{\text{nom}}$  and can be compared to *in vivo* effect data, environmental concentrations, and the unbound fraction in plasma, which is needed for the *in vitro* to *in vivo* extrapolation.  $C_{\text{cell}}$  was relatively constant for chemicals with membrane lipid–water distribution ratios of 1000 or higher and proportional to  $C_{\text{nom}}$ . Representing a sum parameter for exposure that integrates the entire dose from intracellular partitioning,  $C_{\text{cell}}$  is particularly suitable for the effect characterization of chemicals with multiple target sites and the calculation of their relative effect potencies. Effective membrane concentrations indicated that the specific effects of very hydrophobic chemicals in multiple bioassays are occurring at concentrations close to baseline toxicity. The equilibrium partitioning model including all relevant system parameters and a generic bioassay setup is attached as an excel workbook to this paper and can readily be applied to diverse *in vitro* bioassays.



## 2.6 General baseline toxicity QSAR for nonpolar, polar and ionisable chemicals and their mixtures in the bioluminescence inhibition assay with *Aliivibrio fischeri*

Beate I. Escher, Andreas Baumer, Kai Bittermann, Luise Henneberger, Maria König, Christian Kühnert and Nils Klüver

*Environmental Science: Processes Impacts*. **2017**, 19 (3) 414-428.  
doi: 10.1039/C7EM00099E

### ABSTRACT

The Microtox assay, a bioluminescence inhibition assay with the marine bacterium *Aliivibrio fischeri*, is one of the most popular bioassays for assessing the cytotoxicity of organic chemicals, mixtures and environmental samples. Most environmental chemicals act as baseline toxicants in this short-term screening assay, which is typically run with only 30 min of exposure duration. Numerous Quantitative Structure–Activity Relationships (QSARs) exist for the Microtox assay for nonpolar and polar narcosis. However, typical water pollutants, which have highly diverse structures covering a wide range of hydrophobicity and speciation from neutral to anionic and cationic, are often outside the applicability domain of these QSARs. To include all types of environmentally relevant organic pollutants we developed a general baseline toxicity QSAR using liposome–water distribution ratios as descriptors. Previous limitations in availability of experimental liposome–water partition constants were overcome by reliable prediction models based on polyparameter linear free energy relationships for neutral chemicals and the COSMOmic model for charged chemicals. With this QSAR and targeted mixture experiments we could demonstrate that ionisable chemicals fall in the applicability domain. Most investigated water pollutants acted as baseline toxicants in this bioassay, with the few outliers identified as uncouplers or reactive toxicants. The main limitation of the Microtox assay is that chemicals with a high melting point and/or high hydrophobicity were outside of the applicability domain because of their low water solubility. We quantitatively derived a solubility cut-off but also demonstrated with mixture experiments that chemicals inactive on their own can contribute to mixture toxicity, which is highly relevant for complex environmental mixtures, where these chemicals may be present at concentrations below the solubility cut-off.

## 2.7 Baseline toxicity and ion-trapping models to describe the pH-dependence of bacterial toxicity of pharmaceuticals

Andreas Baumer, Kai Bittermann, Nils Klüver and Beate Escher

*Environmental Science: Processes Impacts*. 2017, doi: 10.1039/C7EM00099E

### ABSTRACT

In numerous studies on the toxicity of ionisable organic chemicals, it has been shown that the toxicity was typically higher, when larger fractions of the neutral species were present. This observation was explained in some cases by slower uptake of charged species. In other cases it was suggested that the neutral species has intrinsically higher toxicity than the charged species or is alone responsible for the toxicity. However, even permanently charged and organic chemicals with multiple acid and base functional groups and zwitterions are toxic. We set out to reconcile the divergent views and to compare the various existing models for describing the pH-dependence of toxicity with the goal to derive one model that is valid independent of the type and number of charges on the molecule. To achieve this goal we measured the cytotoxicity of 18 acidic, 15 basic and 9 multiprotic/zwitterionic pharmaceuticals at pH 5.5 to pH 9 with the bioluminescence inhibition test using *Aliivibrio fischeri* (Microtox assay). This assay is useful for an evaluation of various models to describe pH-dependent toxicity because the majority of chemicals act as baseline toxicants in this 30 min cytotoxicity assay. Therefore baseline toxicity with constant membrane concentrations of the sum of all chemical species of approximately  $200 \text{ mmol kg}_{\text{lip}}^{-1}$  served for the validation of the suitability of the various tested models. We confirmed that most tested pharmaceuticals acted as baseline toxicants in this assay at all examined pH values, when toxicity was modeled with a mixture model of concentration addition between the neutral species and all charged species. An ion trapping model, that assumes that the membrane permeability of charged species is kinetically limited, improved model predictions for some pharmaceuticals and pH values. However, neither unhindered uptake nor no uptake of the charged species were ideal models; the reality lies presumably between the two limiting cases with a slower uptake of the charged species than the neutral species. For practical applications a previously developed QSAR model with the ionisation-corrected liposome–water distribution ratio as the sole physicochemical descriptor proved to be generally applicable for all ionisable organic chemicals including those with multiple charges and zwitterions..



## **Eidesstattliche Erklärung**

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Weiterhin erkläre ich, dass ich noch keine vergeblichen Promotionsversuche unternommen habe und die Dissertation weder in der gegenwärtigen noch in einer anderen Fassung bereits einer anderen Fakultät vorgelegen hat.

Leipzig, 28. Mai 2017

Kai Bittermann

**Angaben zur Person und zum Bildungsgang**

Name	Kai Bittermann
Geburtsdatum	29.07.1983
Geburtsort	Neustadt an der Aisch
1988 - 2002	<b>Deutsche Botschaftsschule (Istanbul), Grundschule (Lauf), Peter Vischer Gymnasium (Nürnberg)</b>
2003 - 2005	<b>Grundstudium der Chemie auf Diplom, Friedrich-Alexander Universität Erlangen/Nürnberg</b>
2005 - 2006	Mitarbeit in der Arbeitsgruppe von Prof. María Luisa Moya über die <i>"Struktur und Reaktivität von Mizellen"</i> and der Universidad de Sevilla (Veröffentlichung im Journal of Colloid and Interface Science 313 (2007) 542-550)
2006 - 2010	<b>Hauptstudium der Chemie an der F.-A. Universität Erl/Nbg</b>
2009/2010	<b>Diplomarbeit an der University of Florida:</b> <i>„Building and refining a model of the oxalate/formate antiporter – a membrane bound protein that is essential to the metabolism of an oxalate-degrading bacterium“</i>
2011 – 2017	<b>Helmholtz Zentrum für Umweltforschung</b> Department: Analytische Umweltchemie Promotion: <i>“Equilibrium Partitioning of Ionic Organic Chemicals in Environmental Systems: Experiments and Model Predictions”</i> Mitarbeit im AiF ZIM Projekt: <i>“Entwicklung einer kommerziellen Software zur Vorhersage der Permeabilität von neutralen und ionischen organischen Chemikalien in biologischen Membranen”</i>

## Publikationsliste

### Veröffentlichungen

- Rodríguez, A., del Mar Graciani, M., Bittermann, K., Carmona, A.T., Moyá, M.L., 2007. Micellar kinetic effects in gemini micellar solutions: Influence of sphere-to-rod transitions on kinetics. *J. Colloid Interface Sci.* 313, 542–550.
- Bittermann, K., Spycher, S., Endo, S., Pohler, L., Huniar, U., Goss, K.-U., Klamt, A., 2014. Prediction of Phospholipid–Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic. *J. Phys. Chem. B* 118, 14833–42.
- Bittermann, K., Spycher, S., Goss, K.-U., 2016. Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds. *Chemosphere* 144, 382–391.
- Bittermann, K., Spycher, S., Goss, K.-U., 2017. Erratum to “Comparison of different models predicting the phospholipid-membrane water partition coefficients of charged compounds” [*Chemosphere* 144C (2016) 382–391]. *Chemosphere* 179, 405–406.
- Fischer, F.C., Henneberger, L., König, M., Bittermann, K., Linden, L., Goss, K.-U., Escher, B.I., 2017. Modeling Exposure in the Tox21 in Vitro Bioassays. *Chem. Res. Toxicol.* 30, 1197–1208.
- Escher, B.I., Baumer, A., Bittermann, K., Henneberger, L., König, M., Kühnert, C., Klüver, N., 2017. General baseline toxicity QSAR for nonpolar, polar and ionisable chemicals and their mixtures in the bioluminescence inhibition assay with *Aliivibrio fischeri*. *Environ. Sci. Process. Impacts* 19, 414–428.
- Baumer, A., Bittermann, K., Klüver, N., Escher, B.I.I., 2017. Baseline toxicity and ion-trapping models to describe the pH-dependence of bacterial toxicity of pharmaceuticals. *Environ. Sci. Process. Impacts* Accepted M.
- Bittermann, K., Goss, K.-U., Predicting apparent passive permeability of Caco-2 and MDCK cell-monolayers. In preparation
- Schwöbel, J., Bittermann, K., Huniar, U., Goss, K.-U., Klamt, A., Mechanistic Prediction of Membrane Permeability with COSMOmic: Neutral Compounds, Ionizable Compounds and Ions. In preparation

**Konferenzbeiträge**

- 24.09.2013 SETAC, Essen. Vortrag: *Vorhersage von Phospholipidmembran-Wasser-Verteilungskoeffizienten anionischer organischer Substanzen*
- 18.02.2014 Umweltbundesamt, Dessau. Vortrag: *siehe SETAC Essen*
- 14.05.2014 SETAC, Basel. Vortrag: *A Mechanistic Model to predict Phospholipid-Water Partition Coefficients of Ions*  
Poster: *Modelling the Passive Membrane Permeability for neutral molecules using experimental blood-brain barrier (BBB) data*
- 27.06.2014 Bioakkumulationsworkshop im Umweltbundesamt, Dessau. Vortrag: *siehe SETAC Basel*
- 11.11.2014 SETAC, Vancouver. Poster: *Predicting Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals with a Mechanistically Based Model*
- 16.03.2015 COSMO-RS Symposium, Bonn, Vortrag: *Prediction of Phospholipid-Water Partition Coefficients of Ionic Organic Chemicals using the Mechanistic Model COSMOmic*
- 20.09.2015 ICCE, Leipzig. Poster: *Comparing different models for the prediction of the phospholipid-membrane water partition coefficients of charged chemicals*
- 15.11.2016 LSER workshop, Leipzig. Vortrag: *The membrane-water partition coefficient and what it might be useful for*

## Danksagung

Ich bin meinem Chef, Kai-Uwe Goss, sehr dankbar, dass er mich so lange angestellt hat, dass ich dabei einiges lernen konnte und dass er so integer ist, was mir über die zwangsläufigen Widrigkeiten der Arbeit mit hinweggeholfen hat. Beate Escher danke ich fürs Protegieren, die vielen kritischen Verbesserungen und eine Einladung nach Vancouver. Ich danke Simon Spycher dafür, dass er mir v.a. auf dem Weg zum und durchs erste Paper so nett und motivierend beigestanden hat, mir geduldig die ersten schlechten Manuskripte verbesserte und mich begeistert in R eingeführt hat. Danke an Sven Jakobtorweihen, der mir nicht nur seine MD Simulationen zur Verfügung gestellt hat, sondern auch so bereitwillig sein Wissen um die COSMOmic Geheimnisse teilte. Ich danke meiner Arbeitsgruppe für ihre Hilfsbereitschaft; Satoshi Endo, der mir in Theorie und Laborpraxis immer mal wieder den rechten Weg gewiesen hat, Luise Henneberger für Labor- und Formatierungshilfen und v.a. Lukas Linden für die vielen gewinnbringenden Diskussionen, für (Text)verbesserungen und für das gemeinsame Erreichen neuer Horizonte am Kickertisch. Der Volleyballbetriebsportgruppe und affilierten UFZlern danke ich für die stets schöne Atmosphäre beim Volleyball, Skifahren und im Alltag; Steffen Schock und Felix Witing danke ich für den tollen Arbeitsausgleich, das Klettern; Simon Klüpfel und Vincent Schultheiß danke ich für die verlässliche Freundschaft und die vielen kleineren und größeren Hilfen und last but not least danke ich natürlich auch meiner Familie, für die ich diese Diss nicht hätte machen müssen, die mir aber trotzdem immer die Daumen gedrückt hat.





## Appendix

### **Supporting Information 1: Prediction of Phospholipid–Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic**

#### **1 Experimental Section**

##### **1.1 Chemicals details**

Fen and 246TriBP were purchased from Sigma Aldrich, OBS from TCI Europe, Flu from Fluka and 5-NB from Fluorochem. The buffers used were MOPS [3-(N-morpholino)propanesulfonic acid],  $pK_a=7.2$  from Roth and CHES [2-(N-cyclohexylamino)ethanesulfonic acid],  $pK_a=9.3$  from Sigma Aldrich (>99%); KCl was from Fluka (>99.5%) and used for adjusting the ionic strength. All water used in the experiments was purified by a MilliQ Gradient A10 system (Millipore). The synthetic POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine) came from Avanti Polar Lipids (>99%). Phosphoric acid (85%), methanol (SupraSolv) and acetonitril (LiChrosolv HPLC grade) was from Merck.

##### **1.2 Buffer Solution Preparation**

For buffer preparation all ingredients (buffer, KOH and KCl) were first weighed into a volume metric flask in the desired concentration. Then the glass was filled up with water and finally pH was measured with a pH meter. We did not adjust the pH by adding additional acid or base to ensure accurate concentrations of  $K^+$  in the buffer solutions. Instead, the amounts of buffer and KOH were optimized so that the pH falls into the desired range.

For 1 L MOPS buffer with an ionic strength of 100 mM KCl the following amounts were used: 0.561 g (10 mM) KOH, 0.209 g (10 mM) MOPS and 6.710 g KCl (90 mM). The experimentally determined pH was between 8.18 and 8.34 while the  $pK_a$  of MOPS is 7.2.

For 1 L CHES buffer with an ionic strength of 100 mM KCl the same amounts of KCl and KOH were used as above for the MOPS buffer solution. Additionally, 2.073 g (10 mM) CHES has been added. The experimentally determined pH was 10.0.

Note that, strictly speaking, it is not correct to assume a concentration of 100 mM KCl in the buffer solutions used, but a concentration of 100 mM K<sup>+</sup>.

### **1.3 Liposome Preparation**

Pure POPC was dissolved in CHCl<sub>3</sub>, transferred into a round-bottom flask and dried down to a film in a rotary evaporator. Residual traces of solvent were removed under vacuum in a desiccator (without silica gel to avoid possible contamination) overnight, while the lipid was protected against light using aluminum foil. The remaining film was hydrated and suspended by gentle shaking with 10 mL buffer solution and transferred into a cryo-vial. In order to increase the size of the multi-lamellar vesicles (MLV), the liposome samples were shock-frozen in liquid nitrogen and thawed in a 40°C water bath 10 times. Finally the suspension was filtered 10 times through a membrane extruder (Lipex Biomembranes, Vancouver, BC, Canada with Whatman polycarbonate filter membrane pore size 0.1 μm) to strip off the outer lipid layers and form unilamellar vesicles (ULV). The filtrate was stored in an amber glass bottle in the fridge and used within 10 days. The sorption properties of liposomes were assumed unchanged within the storage time, as shown before (Kaiser and Escher, 2006).

The final concentration of POPC liposomes in the suspension was analyzed photometrically via the amount of phosphate after digesting with peroxodisulfate and autoclaving at 120°C for 0.5h, according to DIN ISO 15923-1.

#### 1.4 Dialysis cell experiments

The home built equilibrium dialysis cells consisted of two glass chambers with 2 mL volume each. One chamber was filled with 2 mL liposome suspension, and the other with 2 mL buffer solution. The chemical was added to the liposome side. This has the advantage of faster equilibration compared to the addition of the chemical to the buffer side, because at equilibrium, the liposome side of the cell contains a larger amount of the chemical than the buffer side and thus adding the chemical to the liposome side makes the initial state closer to the equilibrium state. Samples (200  $\mu$ L) were taken from the liposome-free side on the 4th and 6th days. All chemicals except OBS were measured with an HPLC system from JASCO, equipped with a UV detector (UV-970 M). Separation was done on an Eclipse Plus C18 column from Agilent (4.6 mm  $\times$  100 mm, 5  $\mu$ m particle size) with gradient or isocratic elution of acetonitrile and water (both containing 0.1 % orthophosphoric acid) at a rate of 1 mL/min. For OBS a Shimadzu HPLC system was used equipped with a diode array detector (SPD-M10AVP) and a Phenomenex Luna HILIC column (4.6 mm  $\times$  100 mm, 5  $\mu$ m particle size). A mixture of acetonitrile and water (10:90) with 5 mM ammonium acetate served as eluent, flowing with 1 mL/min.

The partition coefficient of fluazinam (CAS 79622-59-6) could not be quantified due to alkaline hydrolysis in the pH range of interest (more than 50 % mass loss after 4 days).

## 1.5 Data collection anions

SI-1, Table 1: Data collection for anions. Values annotated with 'P' are taken from the PhysProp-Database (<http://esc.syrres.com/fatepointer/search.asp>), while 'o' denote the own measurements as described above. Egg stands for egg-phosphatidylcholine, DOPC for dioleoylphosphatidylcholine, DPPC for 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine and POPC for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine. If more than one value for  $\log K_{lipw}$  (ion) was found in literature this is marked with a shading. For comparison with calculated  $\log K_{lipw}$  values, the arithmetic means of the experimental  $\log K_{lipw}$  values were taken.

CAS	compoundname	abbreviation	pK <sub>a</sub>	$\log K_{lipw}$ (ion) [L/kg]	lipid	method	T [°C]
4901-51-3	2,3,4,5-tetrachlorophenol	2345TeCP	6.35 (Schellenberg et al., 1984)	3.90 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
				3.48 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25
58-90-2	2,3,4,6-tetrachlorophenol	2346TeCP	5.40 (Schellenberg et al., 1984)	3.46 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
935-95-5	2,3,5,6-tetrachlorophenol	2356TeCP	5.14 <sup>P</sup>	3.49 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25
95-95-4	2,4,5-trichlorophenol	245TriCP	6.94 (Schellenberg et al., 1984)	2.98 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
				2.79 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25
118-79-6	2,4,6-tribromophenol	246TriBP	6.80 <sup>P</sup>	3.07 <sup>o</sup>	popc	equilibrium dialysis	25
88-06-2	2,4,6-trichlorophenol	246TriCP	6.15 (Schellenberg et al., 1984)	2.50 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
				2.54 (Smejtek et al., 1996)	egg	electrophoretic mobility	25

						measurements	
120-83-2	2,4-dichloro-phenol	24DCP	7.85 (Schellenberg et al., 1984)	2.69 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
51-28-5	2,4-dinitrophenol	24DNP	3.94 (Schwarzenbach et al., 1988)	1.90 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
87-65-0	2,6-dichloro-phenol	26DCP	6.97 (Escher and Schwarzenbach, 1996)	1.43 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
				1.40 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25
573-56-8	2,6-dinitrophenol	26DNP	3.70 (Escher and Schwarzenbach, 1996)	1.86 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
95-57-8	2-chlorophenol	2CP	8.56 (Escher and Schwarzenbach, 1996)	0.92 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
534-52-1	2-methyl-4,6-dinitrophenol	DNOC	4.31 (Schwarzenbach et al., 1988)	2.35 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
88-75-5	2-nitrophenol	2NP	7.23 (Schwarzenbach et al., 1988)	0.69 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
88-85-7	2-s-butyl-4,6-dinitrophenol	Dinoseb	4.62 (Schwarzenbach et al., 1988)	3.35 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
1420-07-1	2-tert-butyl-4,6-dinitrophenol	Dino2terb	4.80 (Miyoshi et al., 1987)	3.59 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
609-19-8	3,4,5-trichlorophenol	345TriCP	7.73 (Schellenberg et al., 1984)	3.16 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
95-77-2	3,4-dichlorophenol	34DCP	8.59 (Escher and Schwarzenbach, 1996)	2.85 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
577-71-9	3,4-dinitrophenol	34DNP	5.48 (Schwarzenbach et al., 2003)	1.90 (Escher and Schwarzenbach, 1996)	DPPC/DOPC	equilibrium dialysis	20
1689-84-5	3,5-dibromo-4-hydroxy-benzonitrile	Bromox	4.09 (Escher et al., 2001)	2.10 (Escher et al., 2001)	POPC	TRANSIL	NA
13979-	3,5-dibromo-4-	35DBC	8.28 (Escher et al., 2001)	3.18 (Escher et al., 2001)	POPC	TRANSIL	NA

81-2	methylphenol							
591-35-5	3,5-dichlorophenol	35DCP	8.26 (Schwarzenbach et al., 2003)	2.09 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25	
2338-29-6	4,5,6,7-tetrachloro-2-(trifluoromethyl)-1H-benzimidazole	TTFB	5.30 (Dilger and McLaughlin, 1979)	4.35 (Dilger and McLaughlin, 1979)	egg	equilibrium dialysis	22.5	
106-48-9	4-chlorophenol	4CP	9.38 (Escher and Schwarzenbach, 1996)	2.51 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20	
609-93-8	4-methyl-2,6-dinitrophenol	DNPC	4.06 (Schwarzenbach et al., 1988)	2.26 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20	
100-02-7	4-nitrophenol	4NP	7.08 (Schwarzenbach et al., 1988)	0.95 (Escher and Schwarzenbach, 1996)	DPPC/DOPC	equilibrium dialysis	20	
6149-03-7	4-octylbenzene-1-sulfonate	OBS	NA	3.63 °	POPC	equilibrium dialysis	25	
4097-49-8	4-tert-butyl-2,6-dinitrophenol	Dino4terb	4.11 (Schwarzenbach et al., 1988)	3.23 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20	
2338-25-2	5,6-dichloro-2-(trifluoromethyl)-benzimidazole	DTFB	7.30 (Cohen et al., 1977)	3.05 (Cohen et al., 1977)	egg	equilibrium dialysis	NA	
521-74-4	5,7-dibromo-8-hydroxyquinoline	Dibromox	2.90 (Kaiser and Escher, 2006)	3.03 (Kaiser and Escher, 2006)	POPC	equilibrium dialysis	25	
773-76-2	5,7-dichloro-8-hydroxyquinoline	Dichlorox	2.60 (Kaiser and Escher, 2006)	2.47 (Kaiser and Escher, 2006)	POPC	equilibrium dialysis	25	
16128-96-4	5-chloro-3-tert-butyl-2'-chloro-4'-nitrosalicylanilide	S-13	5.80 (Kasianowicz et al., 1987)	5.05 (Kasianowicz et al., 1987)	egg	equilibrium dialysis	21	
130-16-5	5-chloro-8-hydroxyquinoline	Chlorox	3.71 (Kaiser and Escher, 2006)	1.91 (Kaiser and Escher, 2006)	POPC	equilibrium dialysis	25	

327-19-5	5-nitro-2-trifluoromethyl-benzimidazole	5-NB	NA	1.81 °	POPC	equilibrium dialysis	25
2270-20-4	5-phenylvaleric acid	5-PA	4.88 <sup>P</sup>	1.66 (Avdeef et al., 1998)	DOPC	pH metric technique	NA
148-24-3	8-hydroxy-quinoline	Oxine	4.89 (Kaiser and Escher, 2006)	1.47 (Kaiser and Escher, 2006)	POPC	equilibrium dialysis	25
118-92-3	anthranilic acid	AA	4.76 (Thomae et al., 2007)	0.13 (Thomae et al., 2007)	egg	equilibrium dialysis	26
555-60-2	carbonyl cyanide m-chlorophenyl-hydrazone	CCCP	5.95 (Kasianowicz et al., 1987)	4.05 (Kasianowicz et al., 1987)	egg	equilibrium dialysis	21
370-86-5	carbonyl cyanide p-methoxyphenylhydrazone	FCCP	6.20 (Kasianowicz et al., 1987)	4.22 (Kasianowicz et al., 1987)	egg	equilibrium dialysis	21
15307-86-5	diclofenac	Dic	3.99 (Avdeef et al., 1998)	2.64 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
22494-42-4	diflunisal	Dif	3.00 (Pallicer and Krämer, 2012)	2.73 (Pallicer and Krämer, 2012)	egg	equilibrium dialysis	25
91-40-7	fenamic acid	Fen	3.99 <sup>P</sup>	2.28 °	POPC	equilibrium dialysis	25
530-78-9	flufenamic acid	Flu	NA	3.61 °	POPC	equilibrium dialysis	25
15687-27-1	ibuprofen	Ibu	4.45 (Avdeef et al., 1998)	1.81 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
36894-69-6	labetalol	Lab	7.35 (Pallicer and Krämer, 2012)	1.84 (Pallicer and Krämer, 2012)	egg	equilibrium dialysis	25
608-71-9	pentabromo-phenol	PBrP	4.62 <sup>P</sup>	5.02 (Smejtek et al., 1996)	egg	electrophoretic mobility measurements	25
87-86-5	pentachloro-phenol	PCP	4.75 (Schellenberg et al., 1984)	4.35 (Escher et al., 2000) 4.28 (Smejtek et al., 1996)	DPPC/DOPC egg	equilibrium dialysis electrophoretic mobility	20 25



771-61-9	pentafluoro-phenol	PFP	5.53 <sup>P</sup>		1.74 (Smejtek et al., 1996)	egg	measurements electrophoretic mobility	25
69-72-7	salicylic acid	SA	2.75 (Thomae et al., 2005)		1.03 (Thomae et al., 2005)	DPPC	equilibrium dialysis	37
			3.00 (Ottiger and Wunderli-Allenspach, 1997)		0.85 (Thomae et al., 2007)	egg	equilibrium dialysis	25
					1.04 (Ottiger and Wunderli-Allenspach, 1997)	egg	equilibrium dialysis	37
1198-55-6	tetrachloro-catechol	TeCC	5.97 (Schweigert et al., 2001)		2.63 (Schweigert et al., 2001)	DOPC	potentiometric titration	25
4358-26-3	tetraphenylborate	TPB	NA		5.05 (Flewelling and Hubbell, 1986)	egg	electron paramagnetic resonance	25
					5.35 (Flewelling and Hubbell, 1986)	egg	electron paramagnetic resonance	25
81-81-2	warfarin	Warf	4.90 (Ottiger and Wunderli-Allenspach, 1997)		1.40 (Ottiger and Wunderli-Allenspach, 1997)	egg	equilibrium dialysis	37

The partition coefficient of TPB is given as the ratio of the bound probe molecule surface density to the probe free concentration  $\beta$  in the units of length. It is reported to be 0.02 cm to 0.04 cm (Flewelling and Hubbell, 1986), which results in a log  $K$ 's [L/kg] of 5.05 and 5.35, respectively. The conversion of units can be done as follows:

$$K \left[ \frac{kg}{L} \right] = \frac{0.1 * \beta * S * N_A}{M}$$

S is the surface area of a single membrane lipid molecule and estimated to be  $7 * 10^{-17}$  dm<sup>2</sup>/molecule,  $N_A$  is the Avogadro constant ( $6.022 * 10^{23}$  mol<sup>-1</sup>) and M is the molar mass of the membrane lipid molecules (chosen to be 760.09 g/mol – which is the mass of POPC, a major part in egg phosphatidylcholine).

In the same way the  $\beta$  values have been converted for CCCP (0.002 cm (Kasianowicz et al., 1987)), 26DCP (0.0000045 cm (Smejtek et al., 1996)), 35DCP (0.000022 cm (Smejtek et al., 1996)), 246TriCP (0.000063 cm (Smejtek et al., 1996)), 245 TriCP (0.00011 cm (Smejtek et al., 1996)), 2356TeCP (0.00056 cm (Smejtek et al., 1996)), 2345TeCP (0.00055 cm (Smejtek et al., 1996)), PCP (0.0034 cm (Smejtek et al., 1996)), PFP (0.00001 cm (Smejtek et al., 1996)) and PBrP (0.019 cm (Smejtek et al., 1996)).

## 1.6 Data collection cations

SI-1, Table 2: Data collection for cations. Values annotated with ‘P’ are taken from the PhysProp-Database (<http://esc.syrres.com/fatepointer/search.asp>). Egg stands for egg-phosphatidylcholine, DOPC for dioleoylphosphatidylcholine, DPPC for 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine and POPC for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine. If more than one value for  $\log K_{lipw}$  (ion) was found in literature this is marked with a shading. For comparison with calculated  $\log K_{lipw}$  values, the arithmetic means of the experimental  $\log K_{lipw}$  values were taken.

CAS	compoundname	abbreviat ion	pK <sub>a</sub>	log K <sub>lipw</sub> (ion) [L/kg]	lipid	method	T [°C]
88-05-1	2,4,6- trimethylaniline	246TMA	4.38 (Escher et al., 2000)	2.12 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
95-68-1	3,4-dimethylaniline	34DMA	5.23 (Escher et al., 2000)	1.99 (Escher et al., 2000)	DPPC/DOPC	equilibrium dialysis	20
13214- 66-9	4-phenylbutylamine	4- PhenButA	10.54 (Austin et al., 1995)	2.12 (Austin et al., 1995)	DMPC	ultrafiltration	37
88150- 42-9	amlodipine	Amlodip	9.02 (Austin et al., 1995)	3.75 (Austin et al., 1995)	DMPC	ultrafiltration	37
118-92-3	anthranilic acid	AA	2.15 (Thomae et al., 2007)	1.97 (Thomae et al., 2007)	egg	equilibrium dialysis	NA
29122- 68-7	atenolol	Aten	9.55 (Betageri and Rogers, 1987)	0.51 (Escher et al., 2006)	POPC	equilibrium dialysis	NA

54910-89-3	fluoxetine	Fluox	10.06 (Brooks et al., 2003)	1.5 (Yamamoto et al., 2005)	POPC	equilibrium dialysis	NA
				4.08 (Neuwoehner et al., 2009)	POPC	equilibrium dialysis	NA
312753-06-3	indacaterol	Indac	6.7 (Lombardi et al., 2009)	3.79 (Yamamoto et al., 2005)	POPC	equilibrium dialysis	NA
				4.23 (Nakamura et al., 2008)	POPC	equilibrium dialysis	NA
137-58-6	lidocaine	Lido	7.86 (Ottiger and Wunderli-Allenspach, 1997)	0.91 (Ottiger and Wunderli-Allenspach, 1997)	egg	equilibrium dialysis	37
			7.96 (Avdeef et al., 1998)	1.22 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
51384-51-1	metoprolol	Metro	9.7 (Betageri and Rogers, 1987)	1.43 (Escher et al., 2006)	POPC	equilibrium dialysis	NA
83891-03-6	norfluoxetine	Norfluox	9.05 (Brooks et al., 2003)	3.84 (Neuwoehner et al., 2009)	POPC	equilibrium dialysis	NA
16183-21-4	p-methylbenzylbutylamine	MBButA	9.98 (Fruttero et al., 1998)	1.54 (Fruttero et al., 1998)	egg	potentiometric titration	NA
39099-13-3	p-methylbenzylethylamine	MBEthA	10.04 (Fruttero et al., 1998)	2.26 (Fruttero et al., 1998)	egg	potentiometric titration	NA
215177-24-5	p-methylbenzylhepotentiometric titrationylamine	MBHepA	10.02 (Fruttero et al., 1998)	2.71 (Fruttero et al., 1998)	egg	potentiometric titration	NA
215177-23-4	p-methylbenzylhexylamine	MBHexA	10.17 (Fruttero et al., 1998)	2.43 (Fruttero et al., 1998)	egg	potentiometric titration	NA
699-04-7	p-methylbenzylmethylamine	MBMetA	9.93 (Fruttero et al., 1998)	2.54 (Fruttero et al., 1998)	egg	potentiometric titration	NA
170303-38-5	p-methylbenzylpentylamine	MBPentA	10.08 (Fruttero et al., 1998)	1.84 (Fruttero et al., 1998)	egg	potentiometric titration	NA

39190-96-0	p-methylbenzyl-propylamine	MBPropA	9.98 (Fruttero et al., 1998)	2.11 (Fruttero et al., 1998)	egg	potentiometric titration	NA
59-46-1	procaine	Proc	9.04 (Avdeef et al., 1998)	0.76 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
525-66-6	propranolol	Prop	9.24 (Ottiger and Wunderli-Allenspach, 1997)	2.76 (Ottiger and Wunderli-Allenspach, 1997)	egg	equilibrium dialysis	37
				3.06 (Escher et al., 2006)	POPC	equilibrium dialysis	NA
				9.45 (Pallicer and Krämer, 2012)	egg	equilibrium dialysis	25
				9.53 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
130-95-0	quinine	Quinine	8.63 (Pallicer and Krämer, 2012)	2.47 (Pallicer and Krämer, 2012)	egg	equilibrium dialysis	25
89365-50-4	salmeterol	Salmet	8.8 (Lombardi et al., 2009)	3.67 (Lombardi et al., 2009)	DMPC	equilibrium dialysis	37
94-24-6	tetracaine	Tetrac	8.49 (Avdeef et al., 1998)	2.11 (Avdeef et al., 1998)	DOPC	potentiometric titration	25
18198-39-5	Tetraphenyl-phosphonium	TPP	NA	1.37 (Flewelling and Hubbell, 1986)	egg	equilibrium dialysis	NA
				1.01 (Demura et al., 1987)	DPPC	electron paramagnetic resonance	45

The log  $K$  for TPP is calculated as shown above – the value for  $\beta$  is reported to be  $4.2 \cdot 10^{-6}$  cm (Flewelling and Hubbell, 1986).

## 2 Theory

### 2.1 Estimation of variance

The RMSE has been calculated with the well-known formula:

$$RMSE = \sqrt{\frac{\sum_1^n (experiment - prediction)^2}{n}}$$

In contrast to this ‘normal’ RMSE that directly compares experimental with predicted values, we introduced the RMSE with respect to the offset. For this RMSE (offset), a constant offset (being the average overprediction calculated in the model) is subtracted from predicted values, resulting in the following formula:

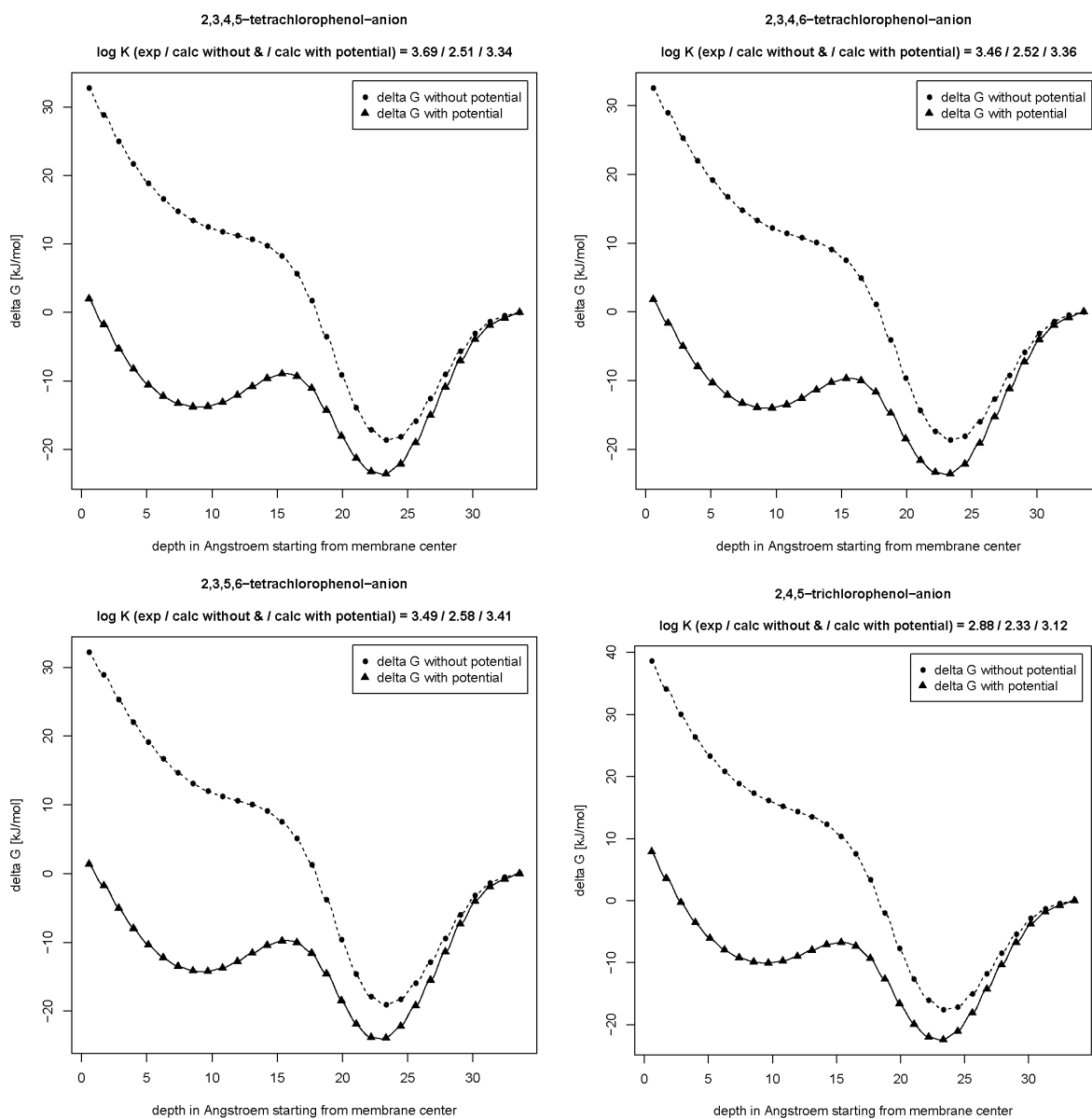
$$RMSE (offset) = \sqrt{\frac{\sum_1^n [experiment - (prediction - offset)]^2}{n}}$$

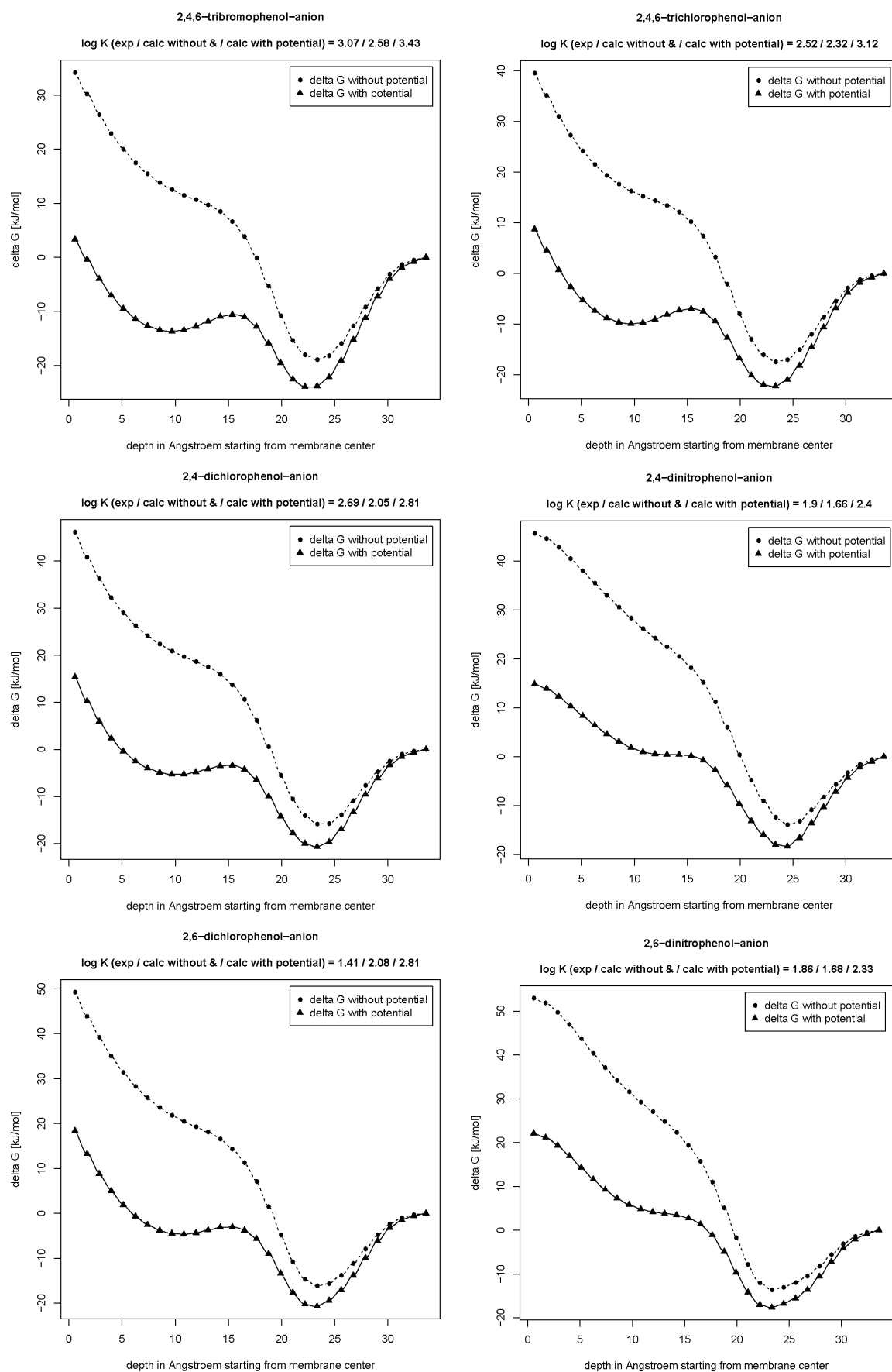
Because the predicted values for  $K_{lipw}$  of ions are considerably off when the membrane potential is not accounted for (cations are 0.9 to 2.3 log units overestimated while anions are up to 1.9 log units underestimated), but at the same time the predictions show a reasonable good fit (cations have an  $R^2$  of 0.45, while anions have an  $R^2$  of 0.76 for the DMPC membrane), we decided to relate the RMSE in that special case to the regression line. This would be according to the use of COSMOmic as a semi-empirical model, as it has been done previously (Spycher et al., 2008), but is not recommended in this work, because the subsequent introduction of the membrane potential makes such a fit unnecessary. The RMSE in respect to the regression line is calculated as follows:

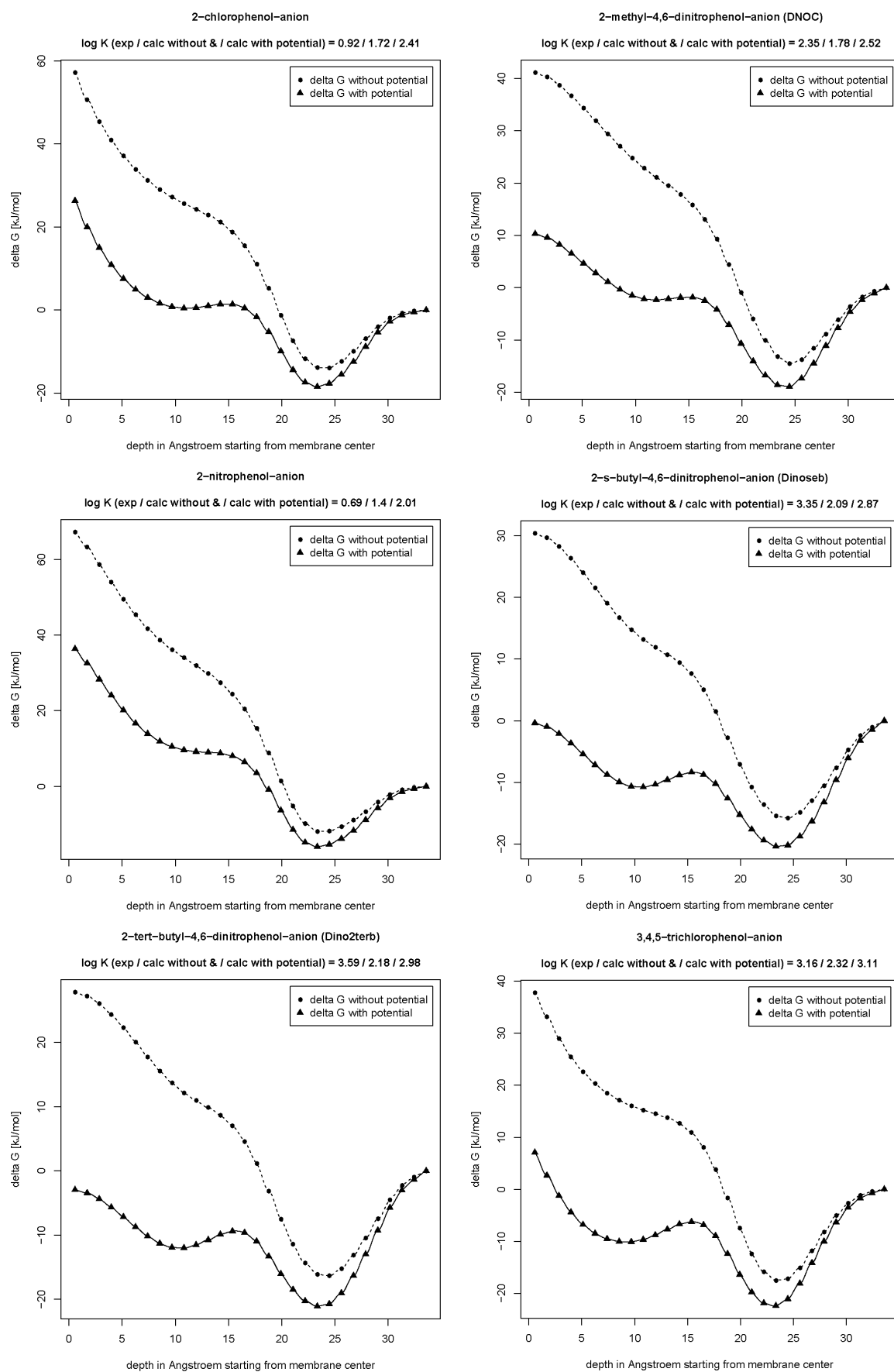
*RMSE (regression)*

$$= \sqrt{\frac{\sum_1^n [experiment - (\frac{1}{slope} * \log K_{lipw} (calculated) - \frac{intercept}{slope})]^2}{n}}$$

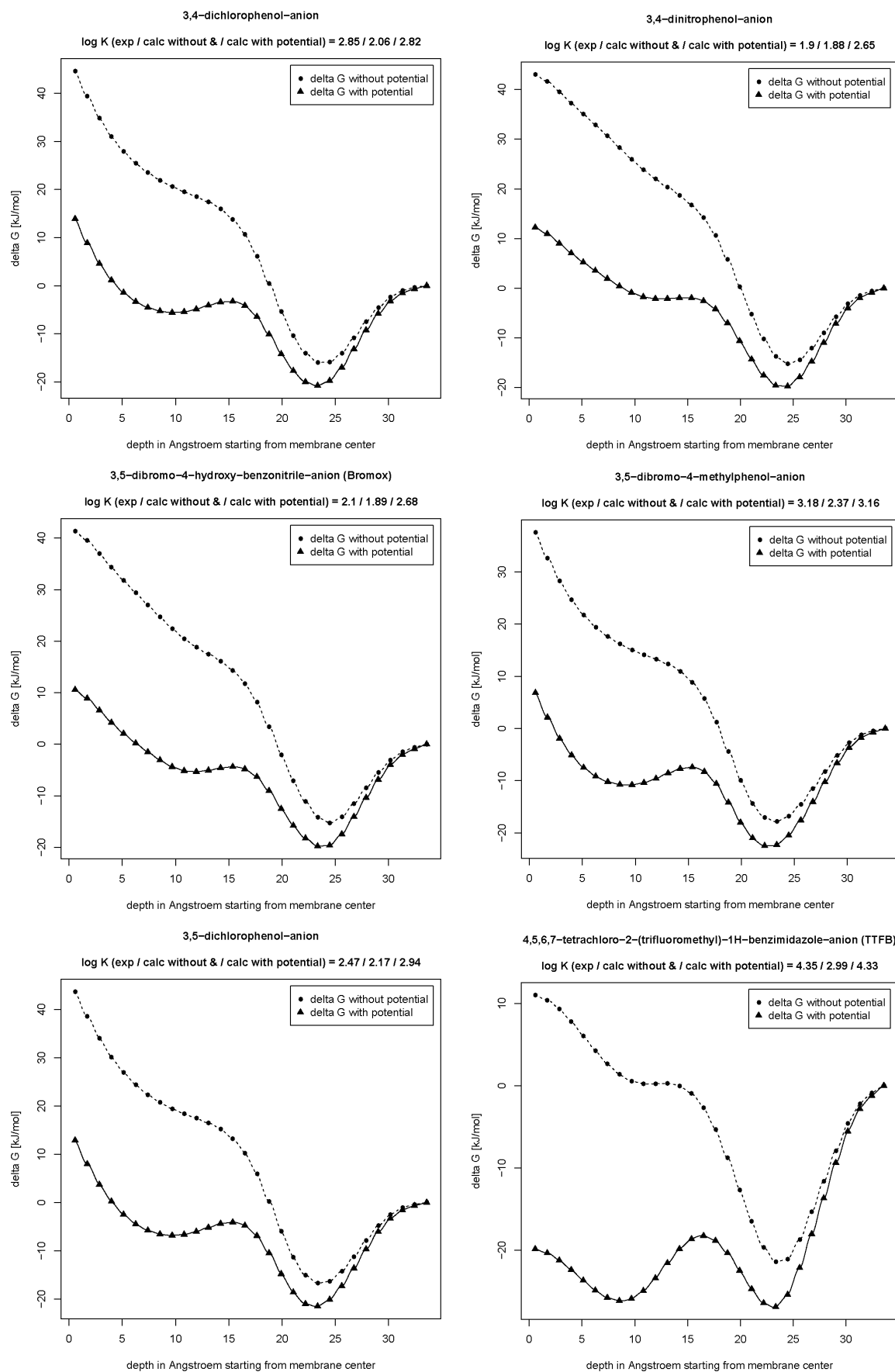
## 2.2 Influence of the membrane potential on the $\Delta G$ profiles of the anions

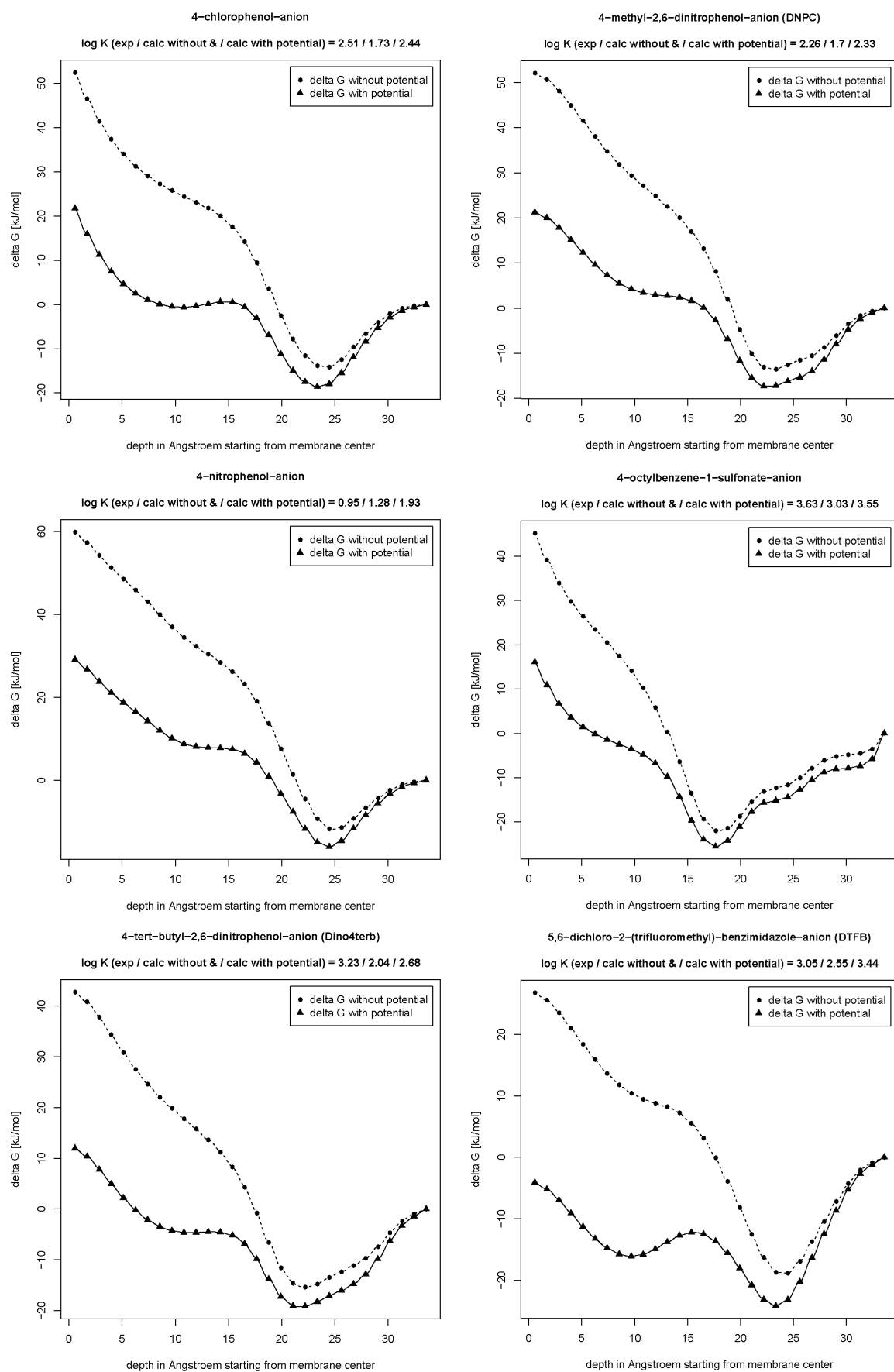


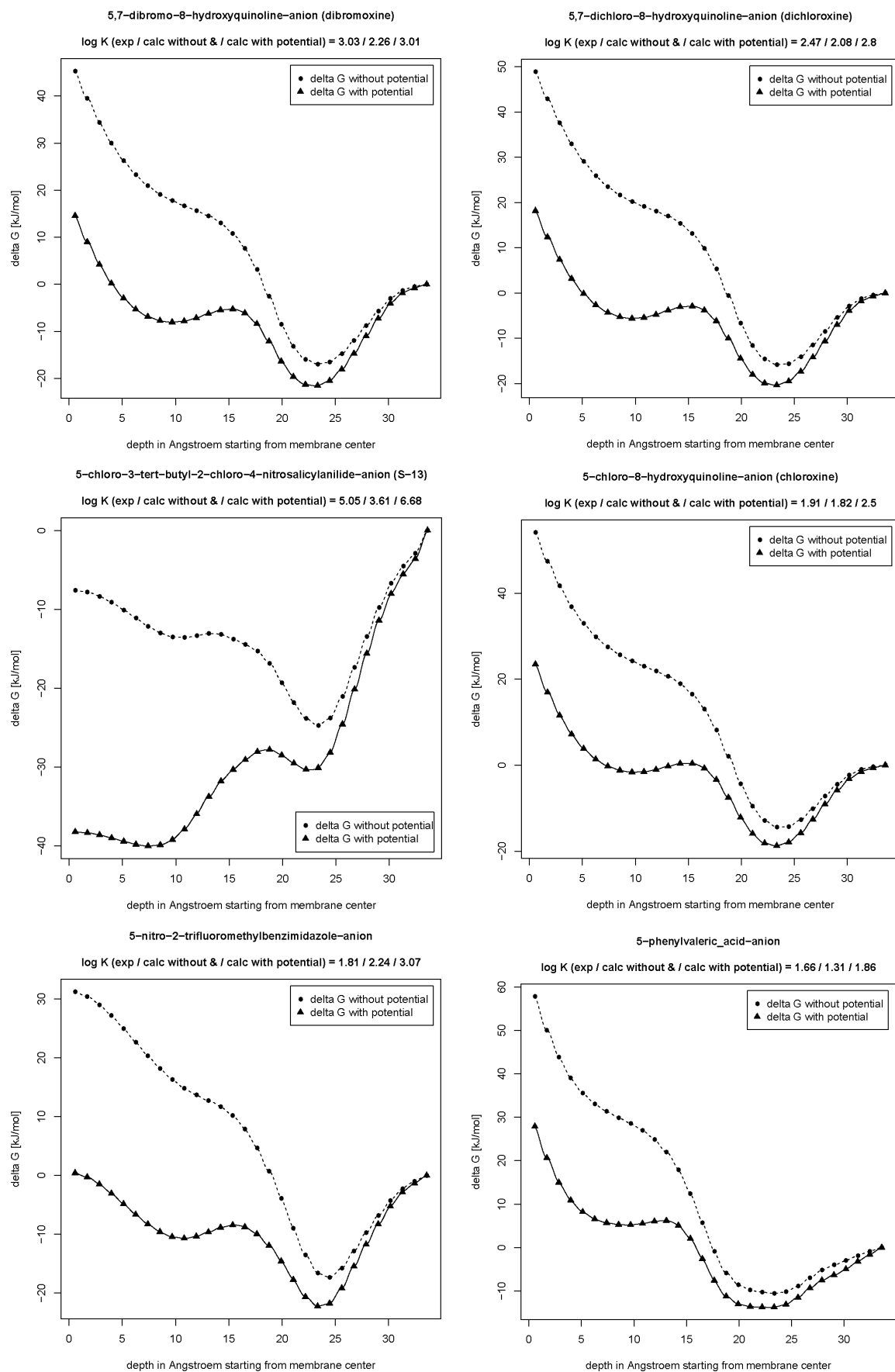


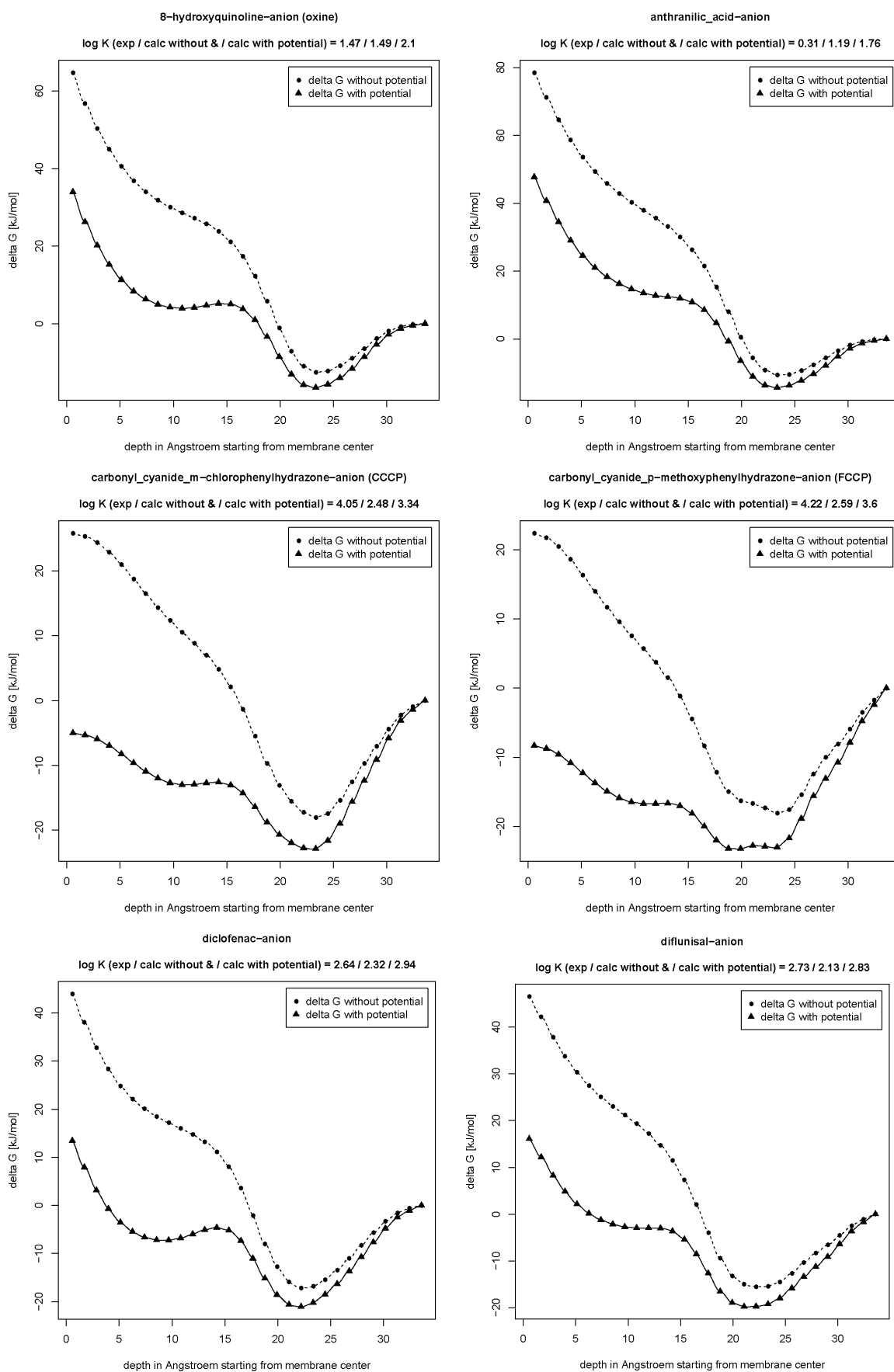


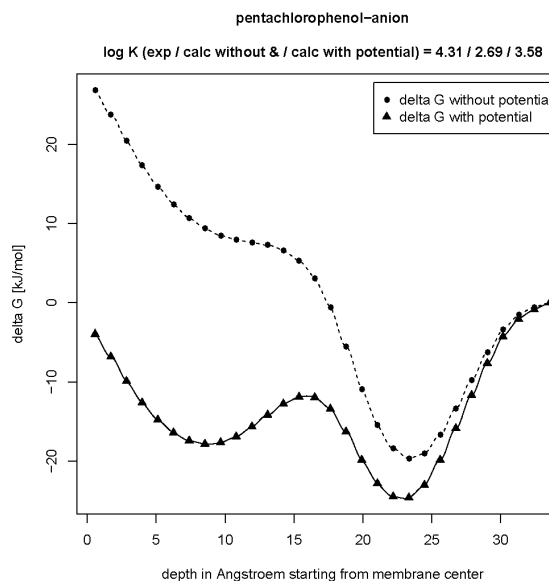
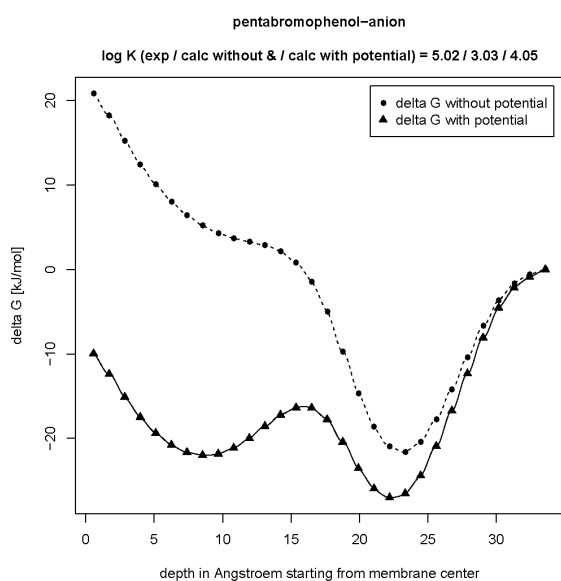
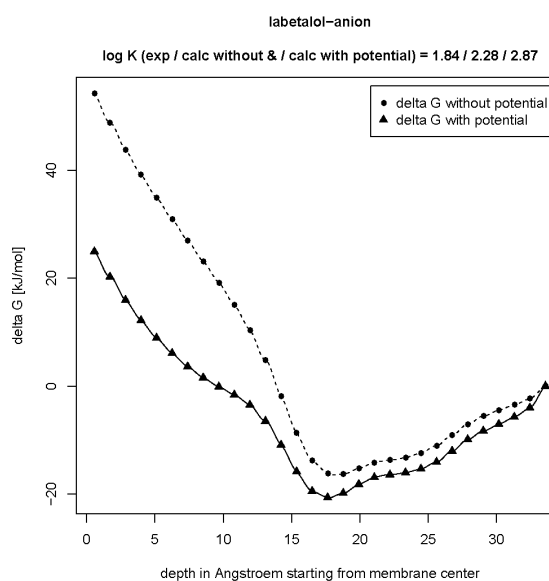
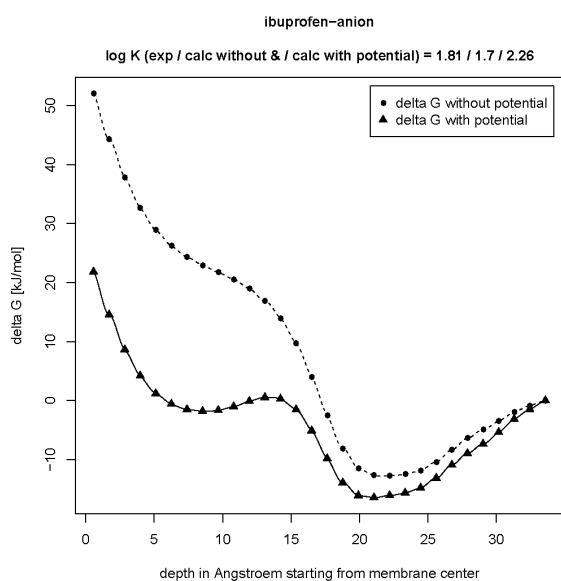
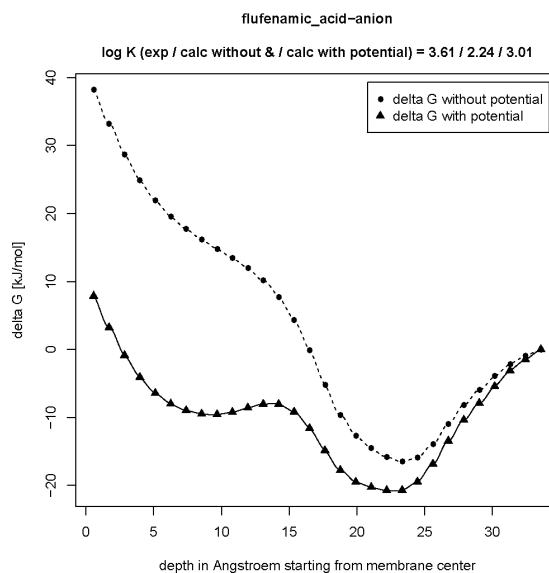
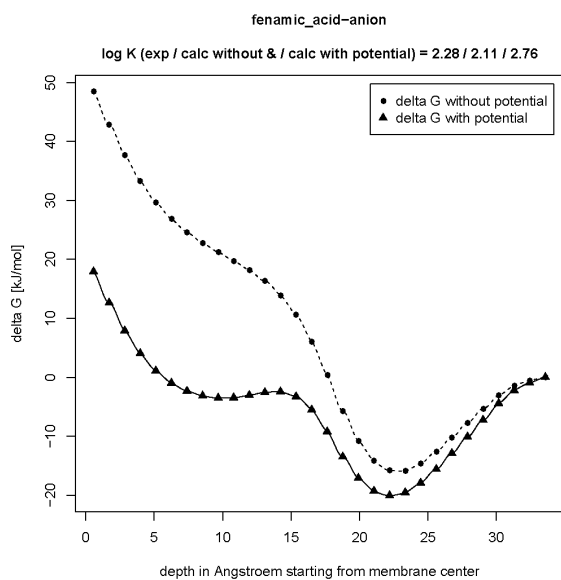


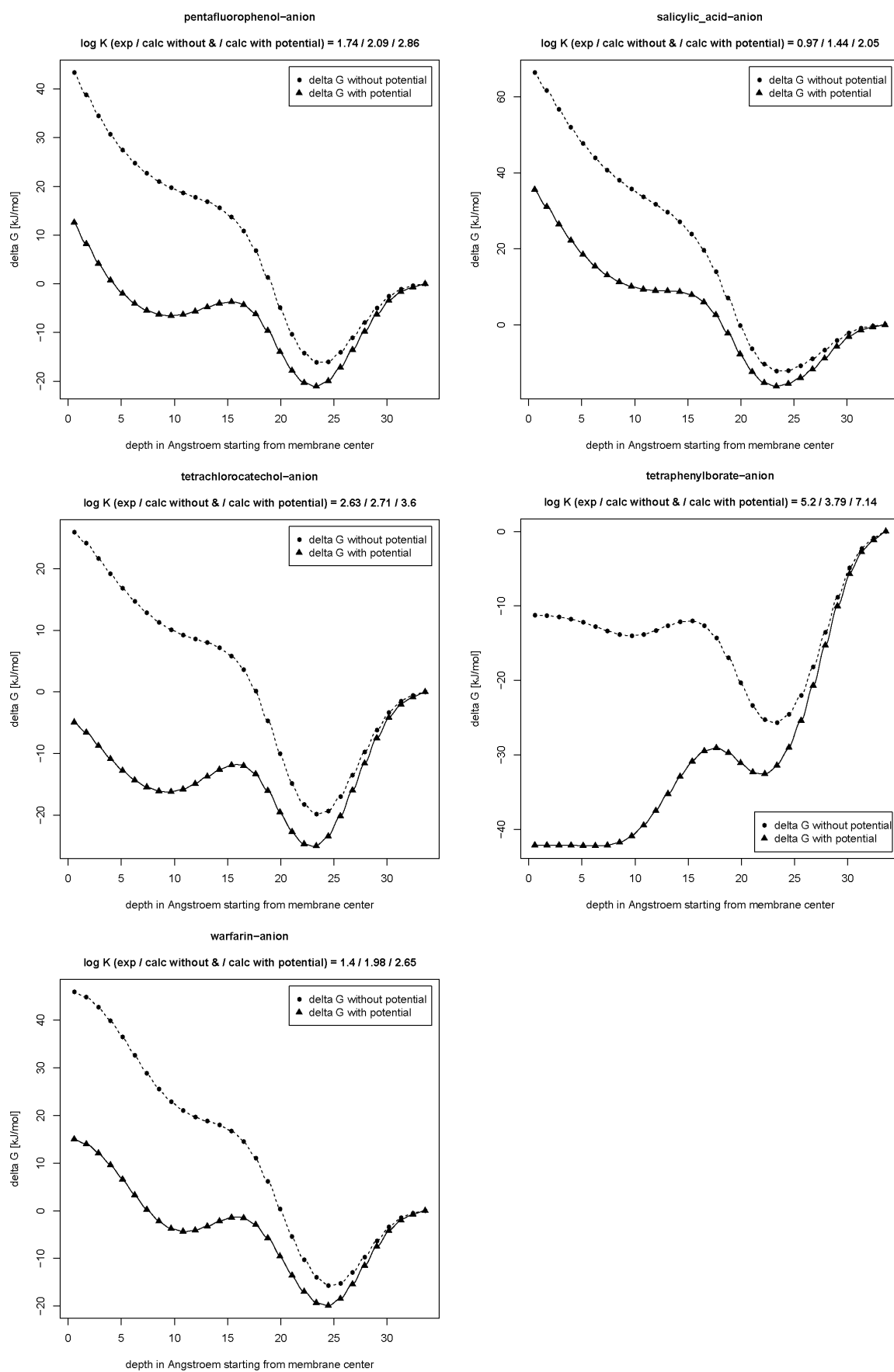




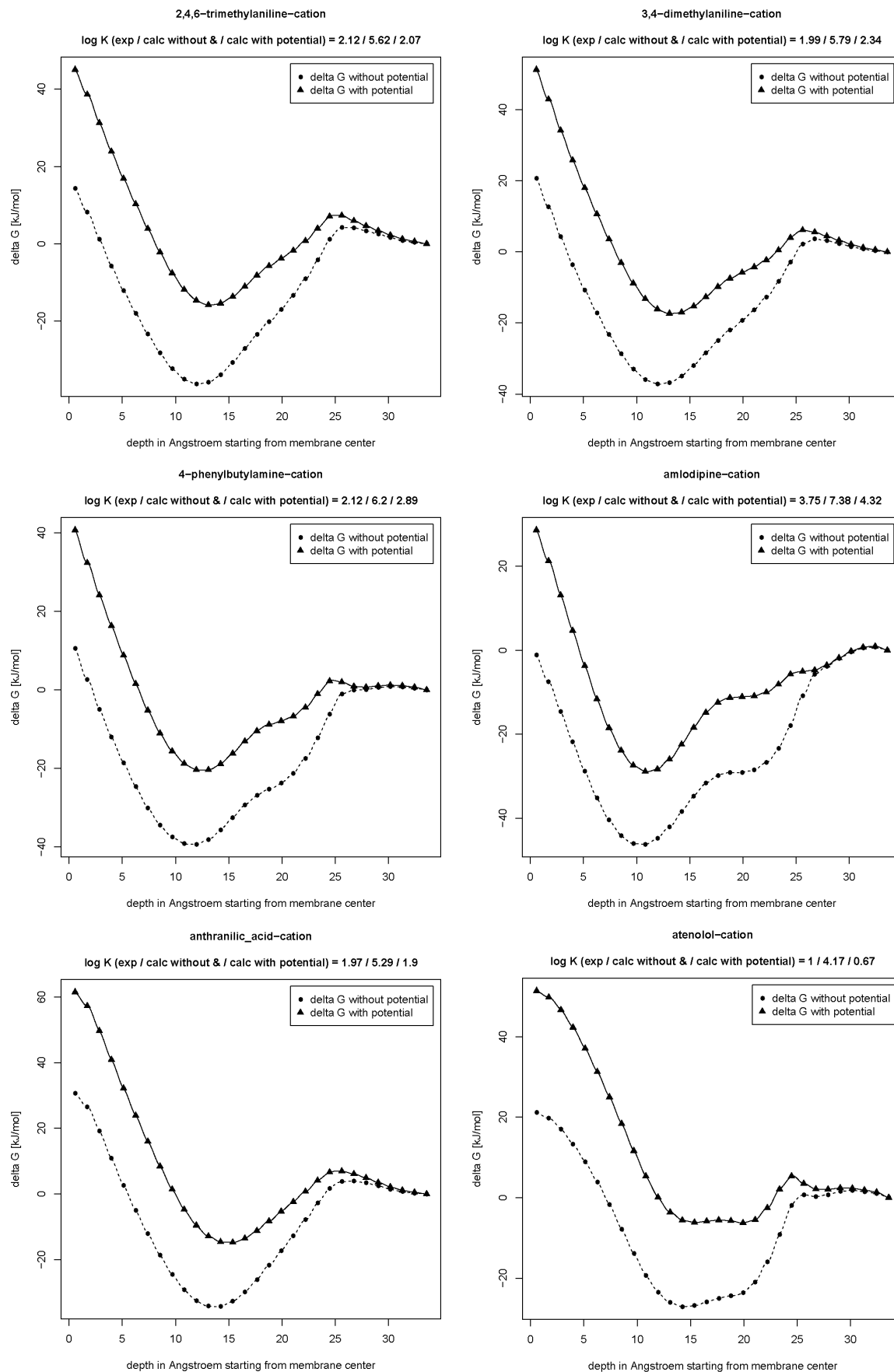


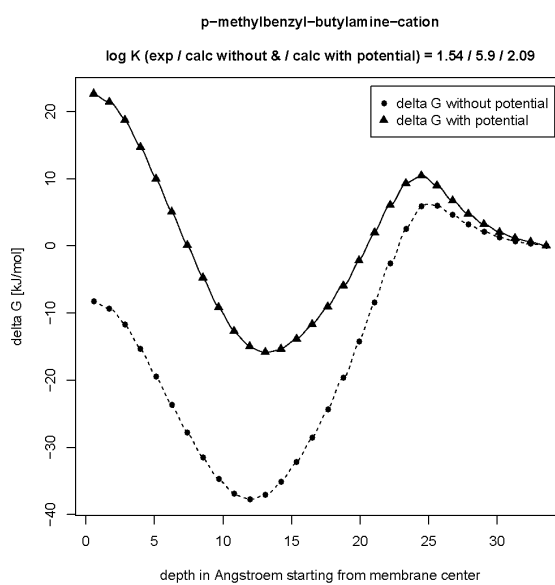
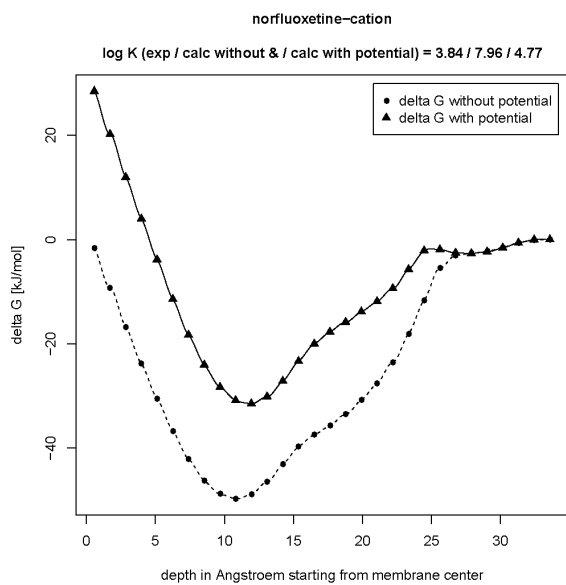
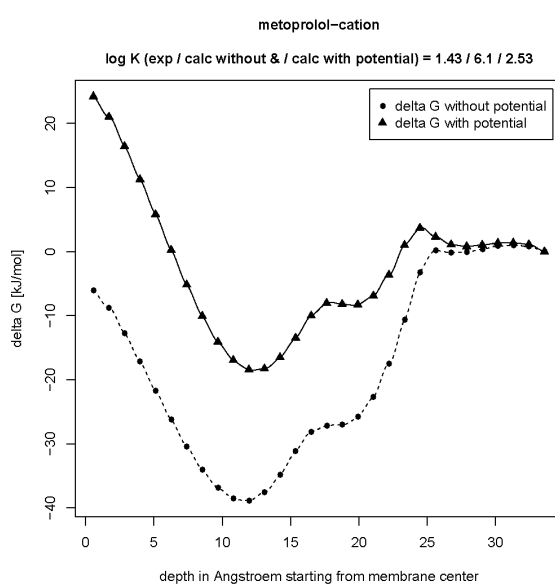
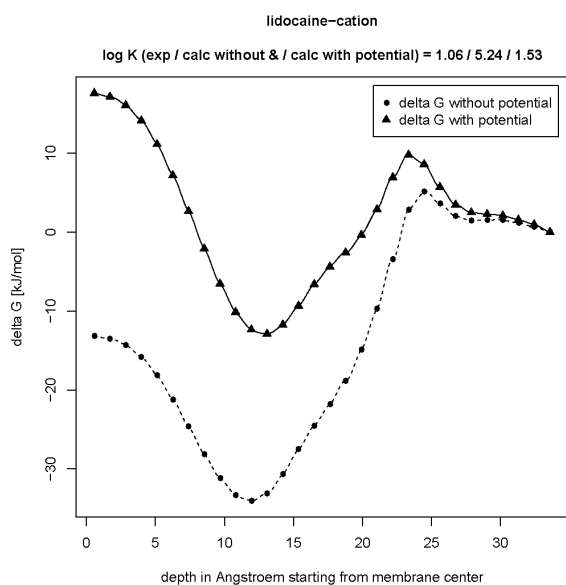
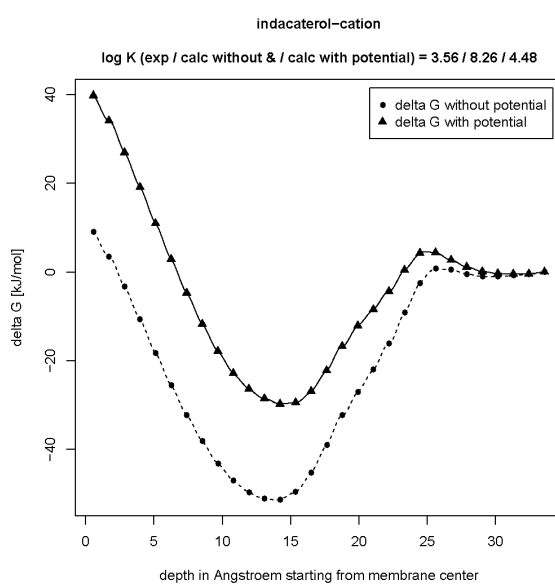
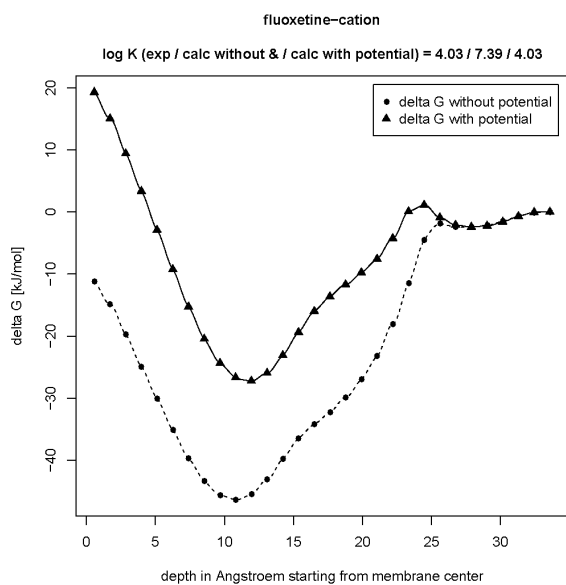




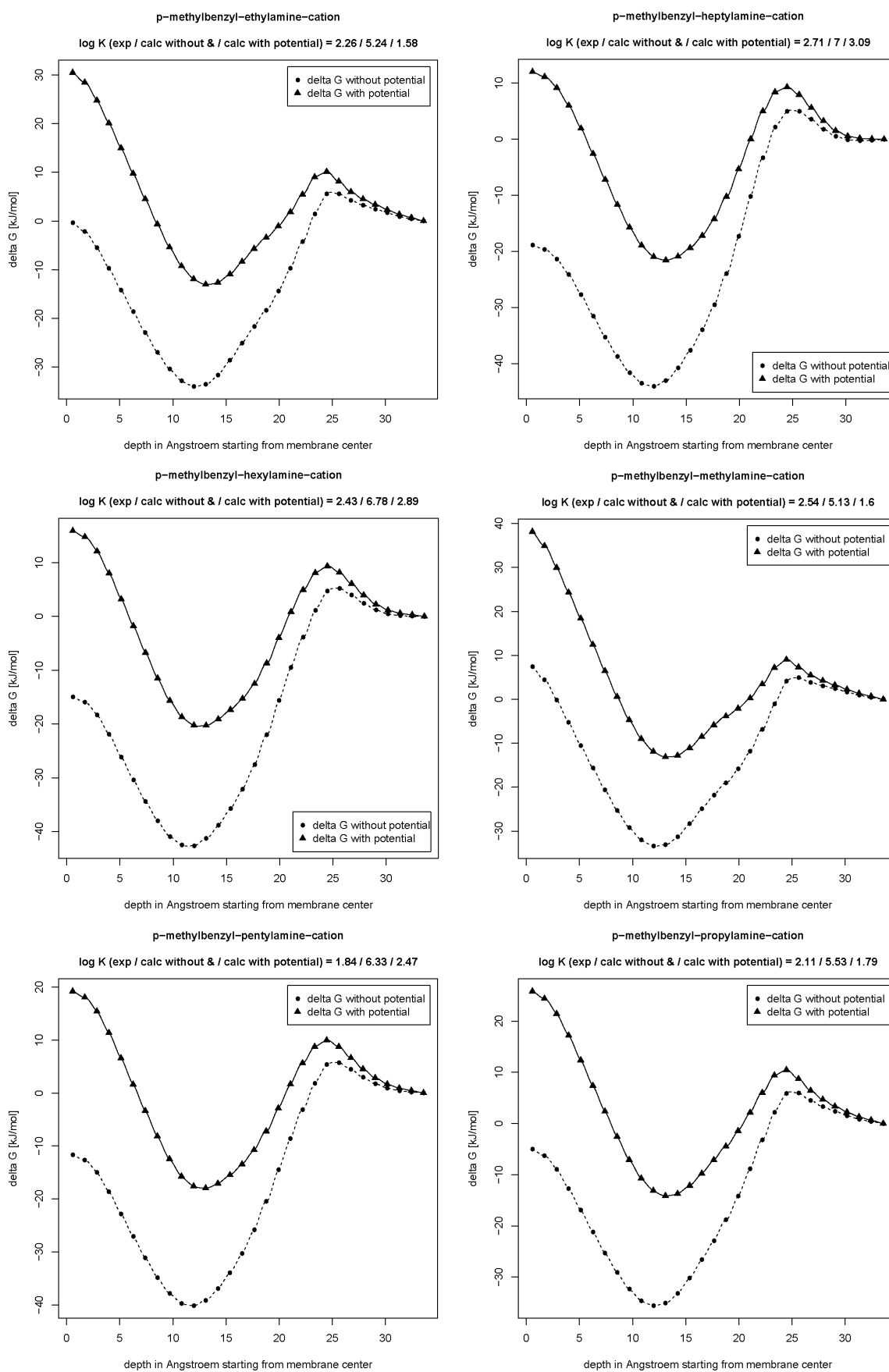


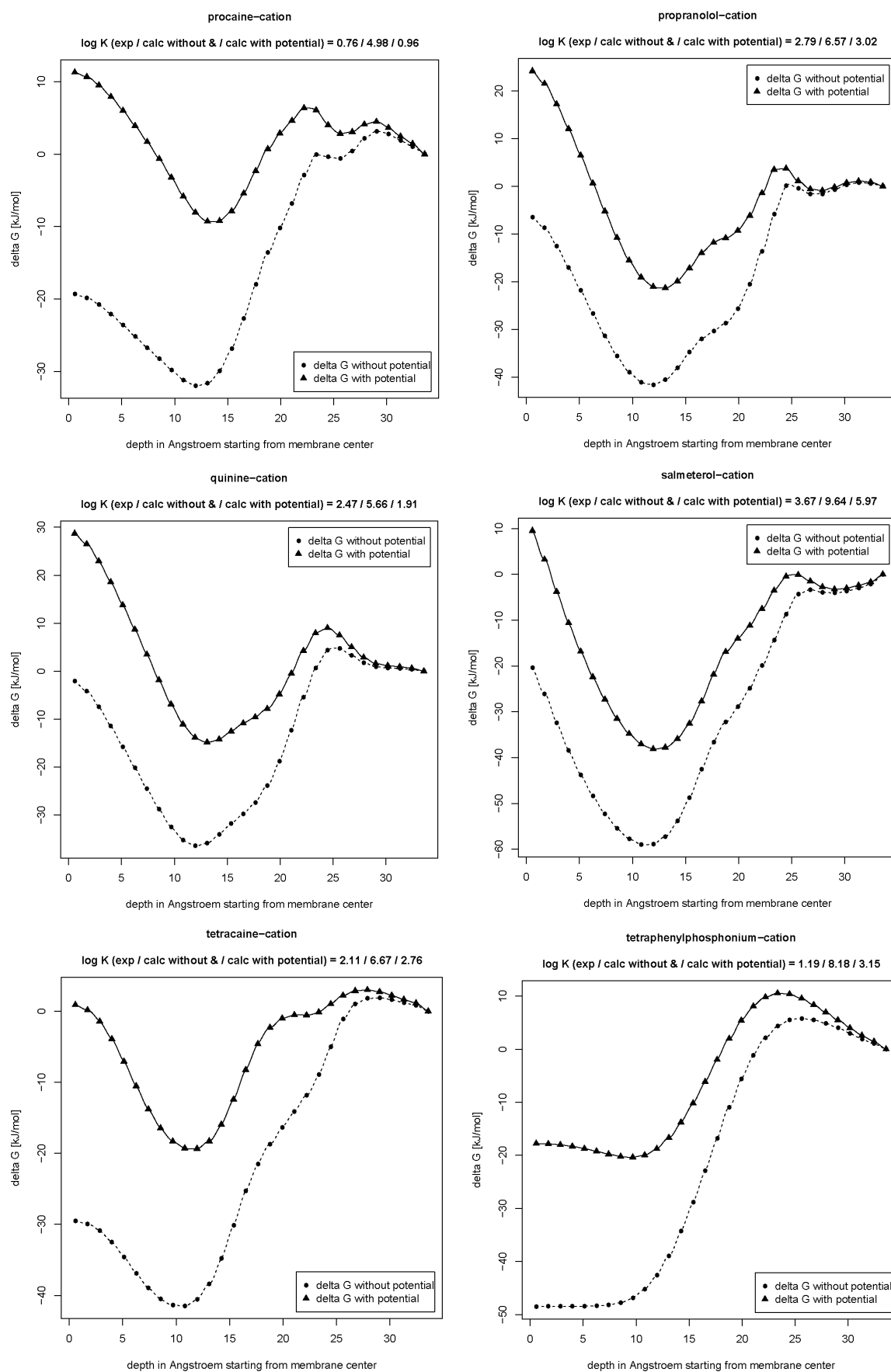
## 2.3 Influence of the membrane potential on the $\Delta G$ profiles of the cations



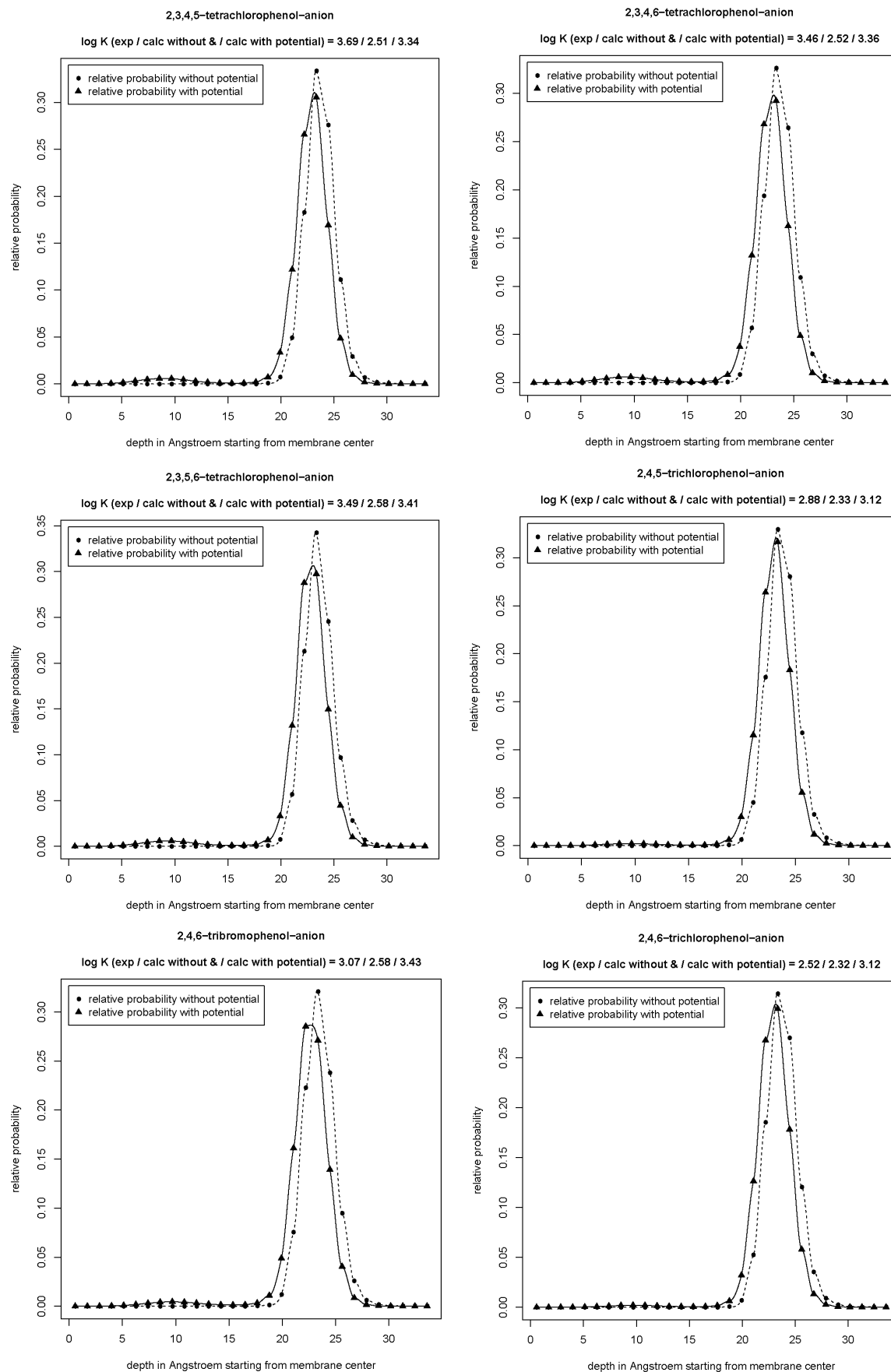


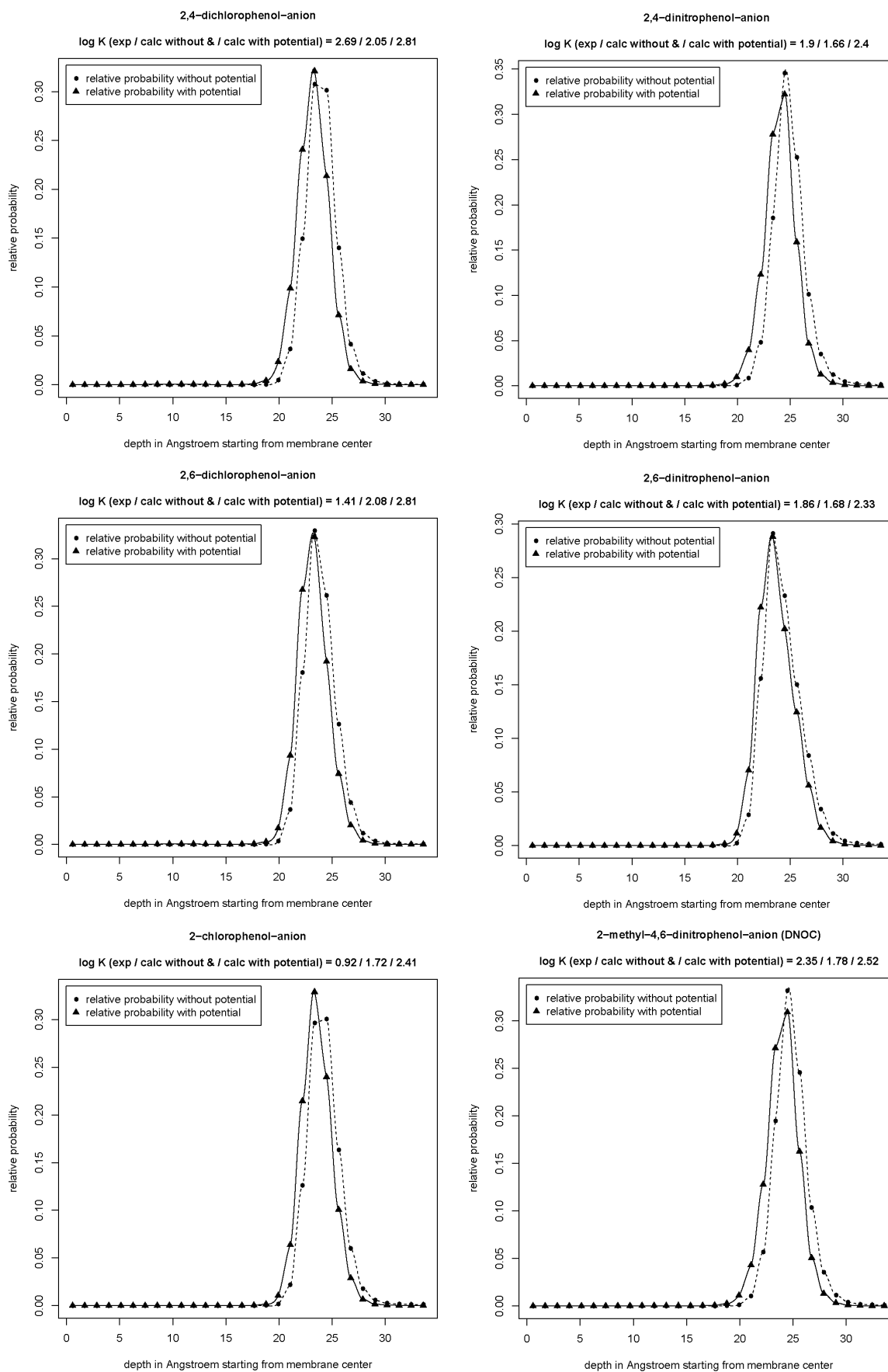


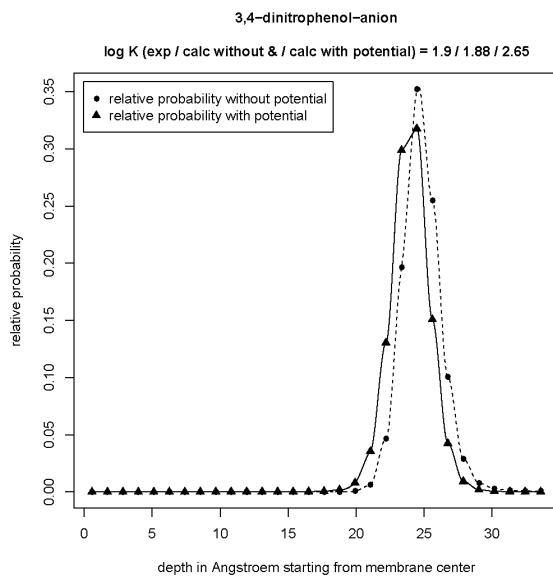
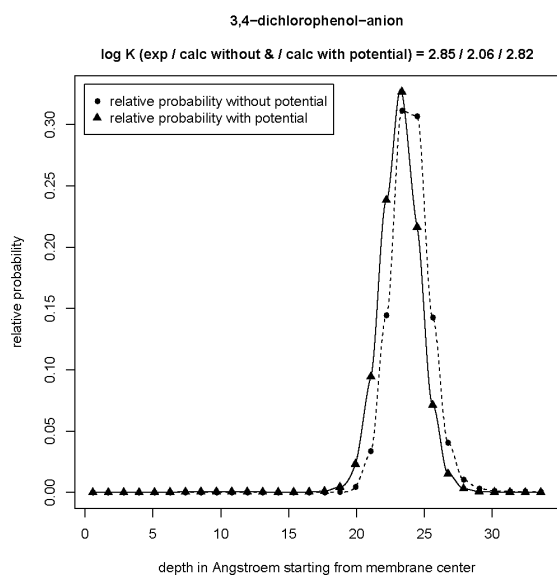
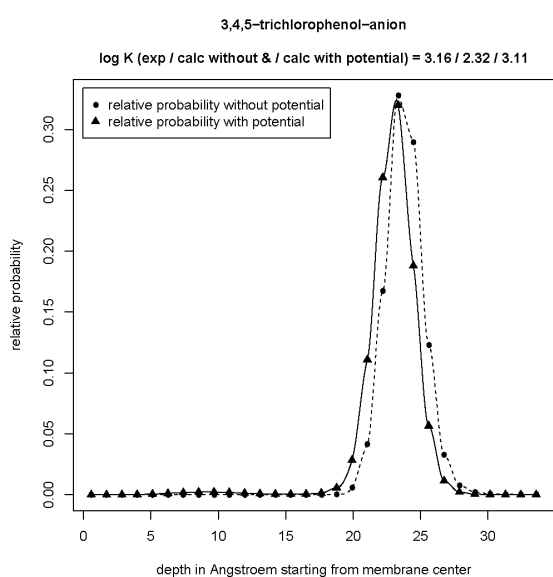
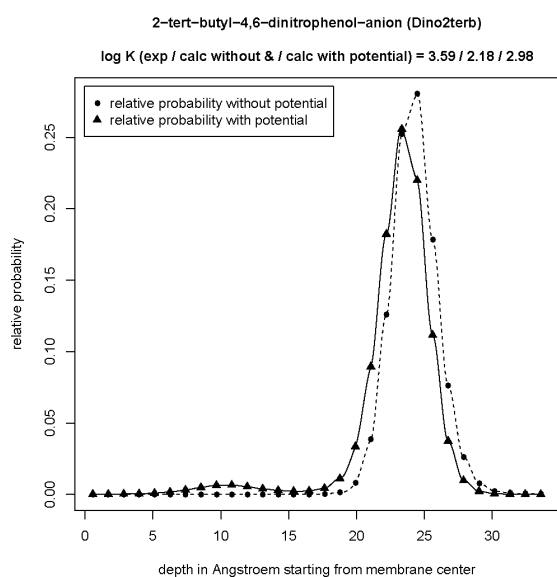
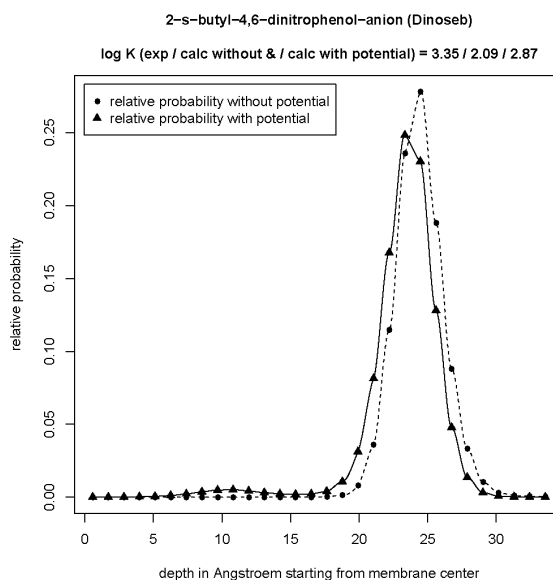
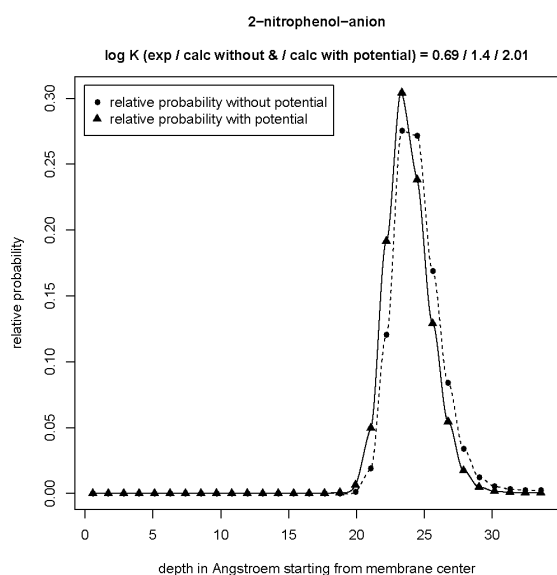


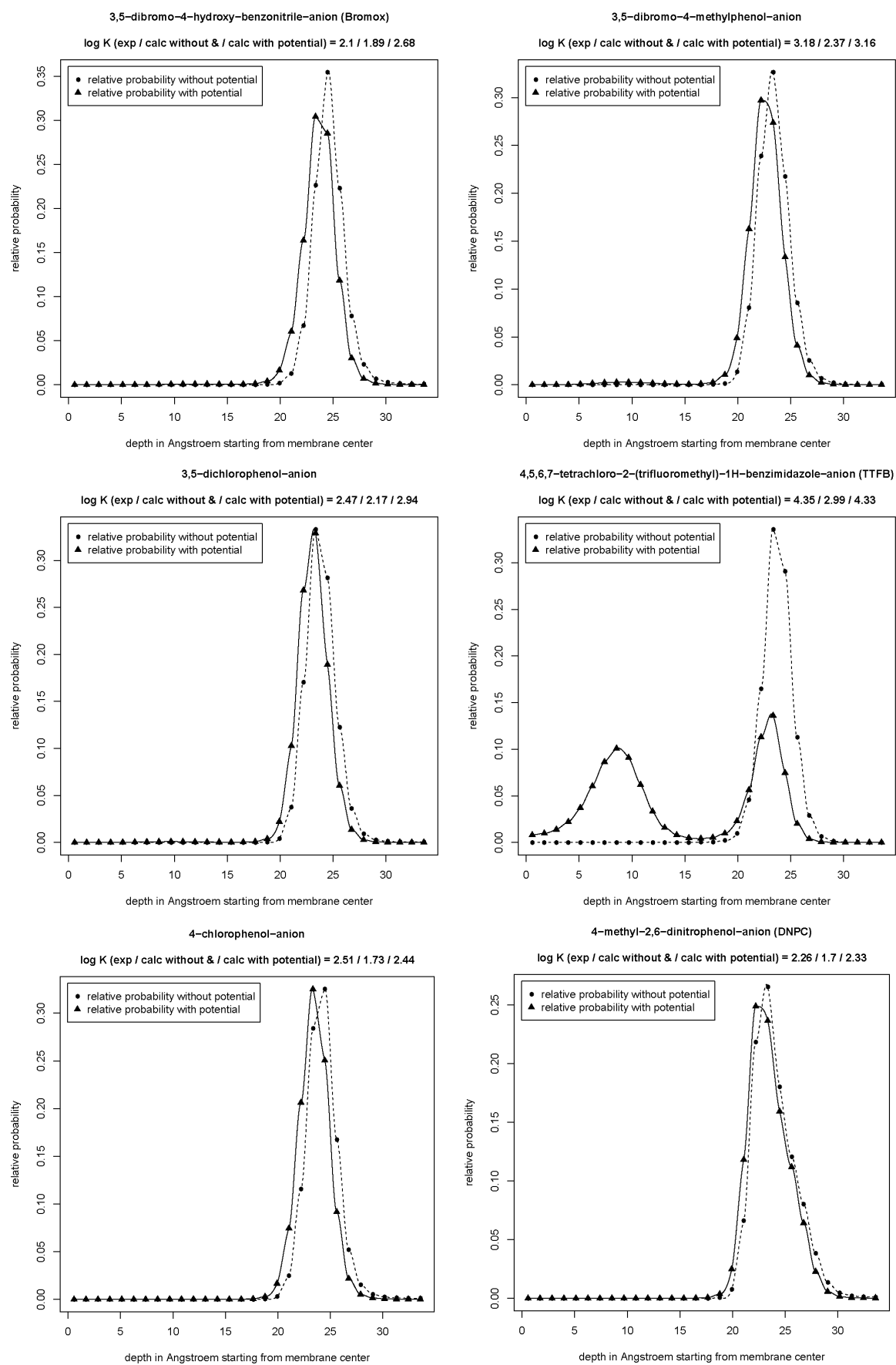


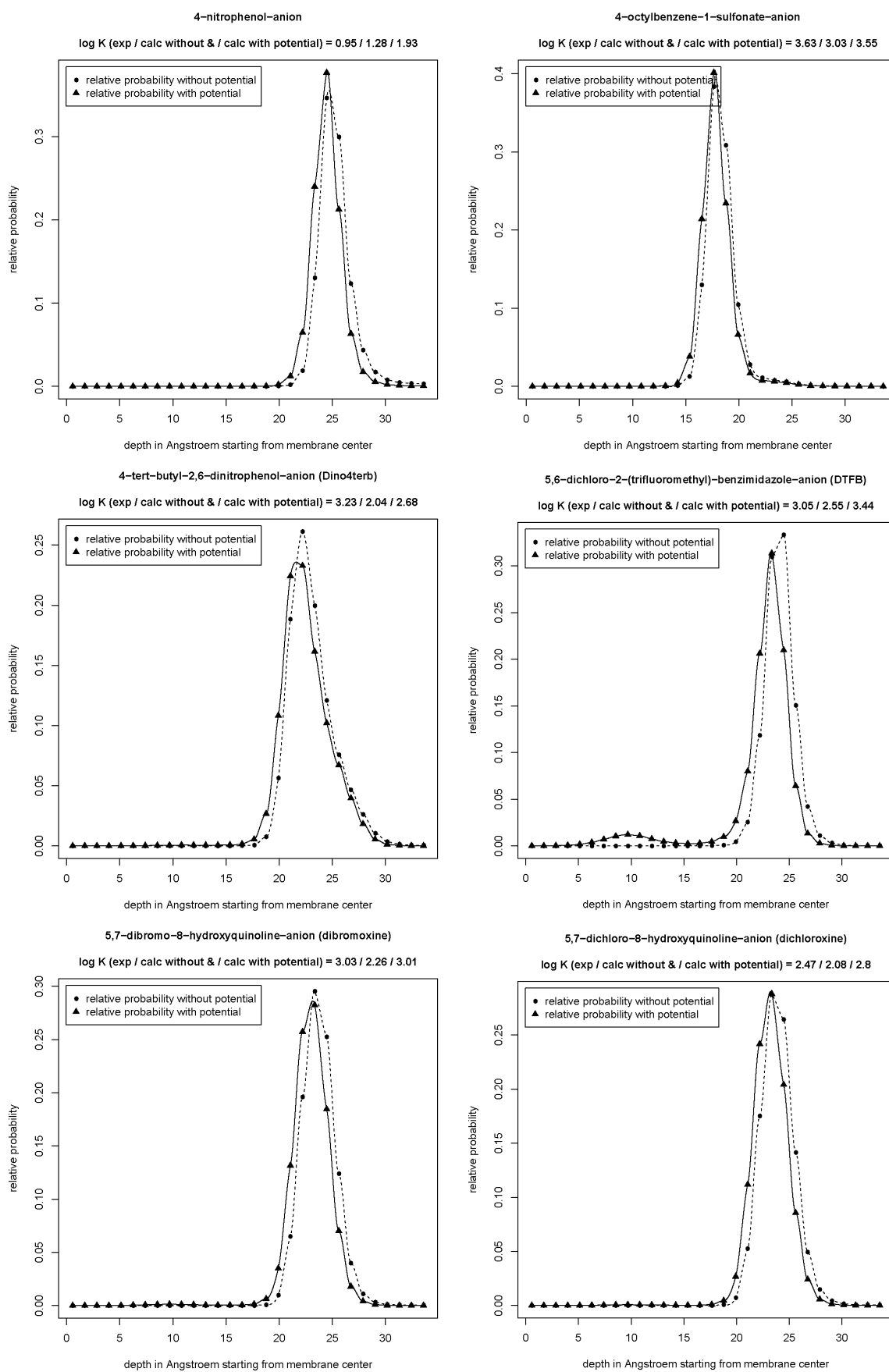
## 2.4 Influence of the membrane potential on the relative distribution of the anions

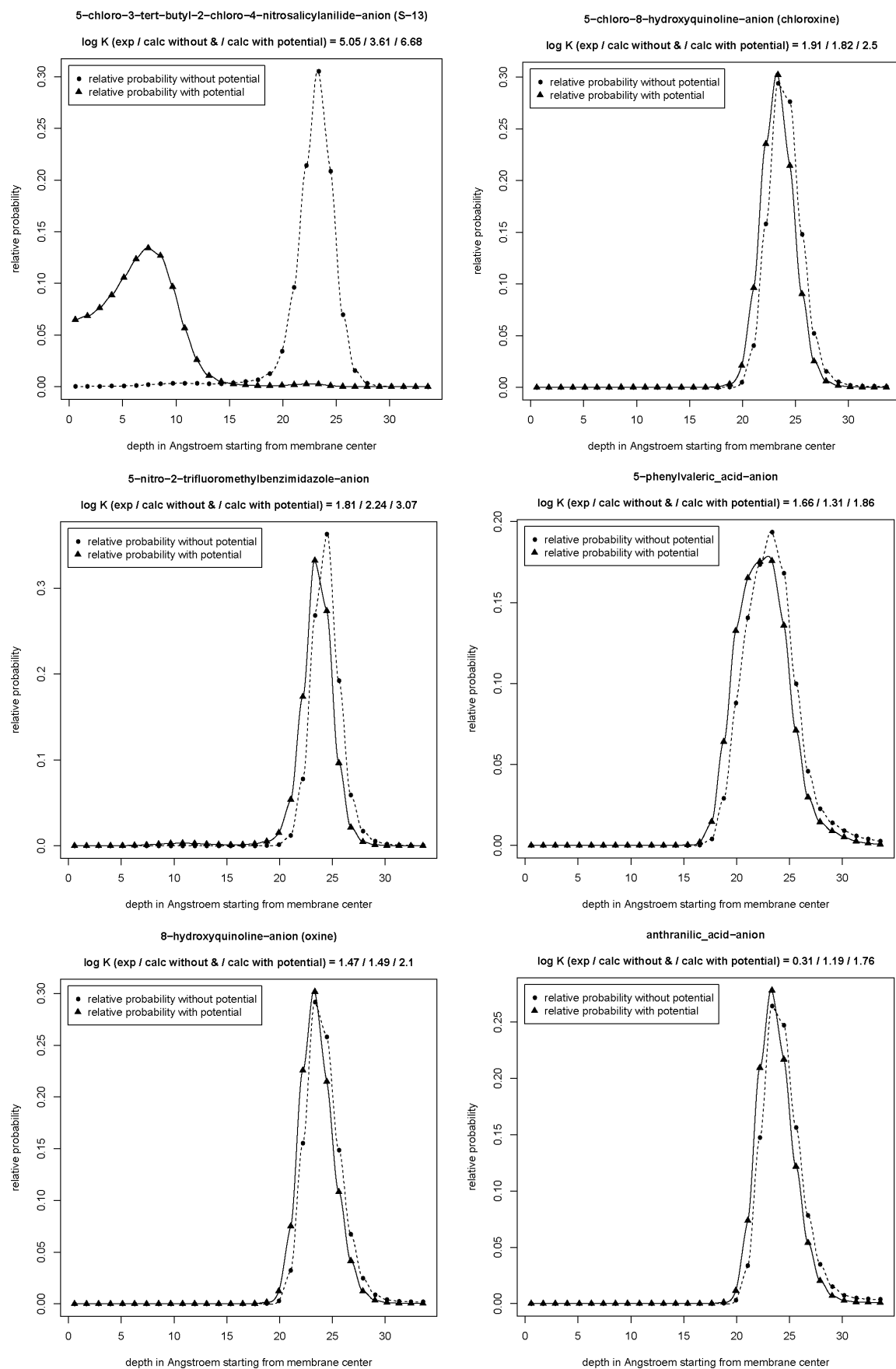




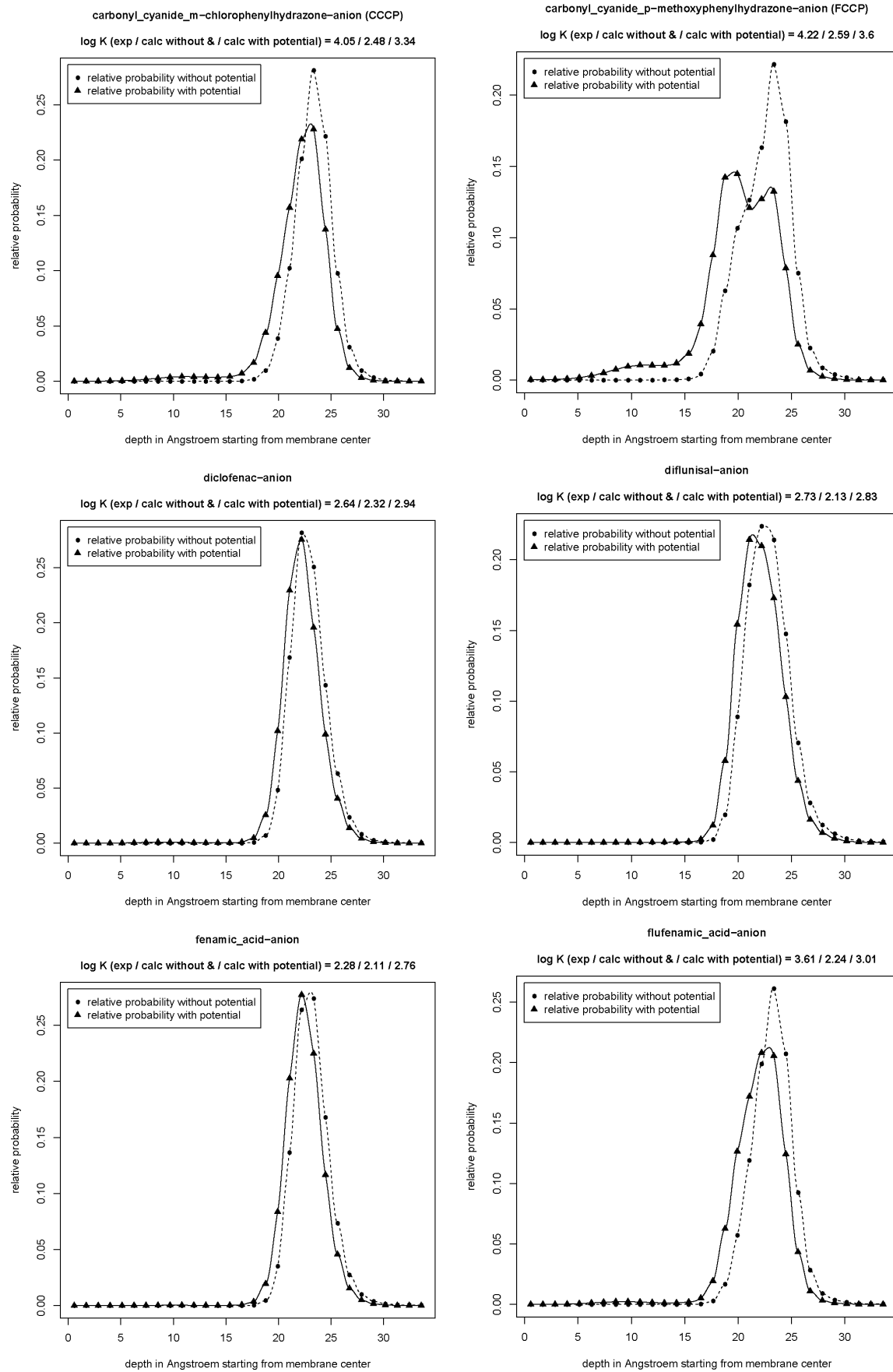


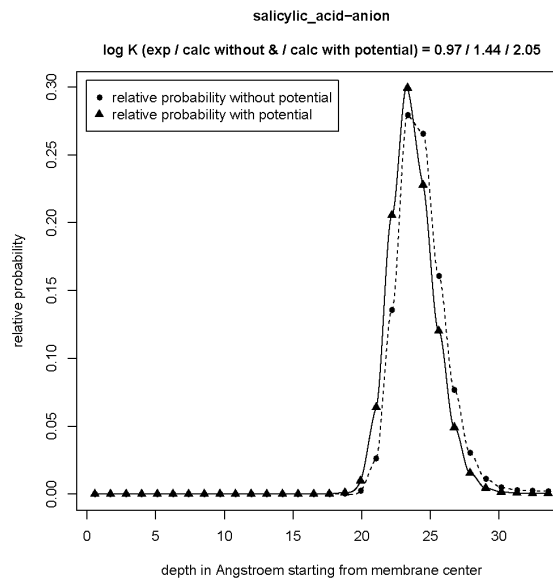
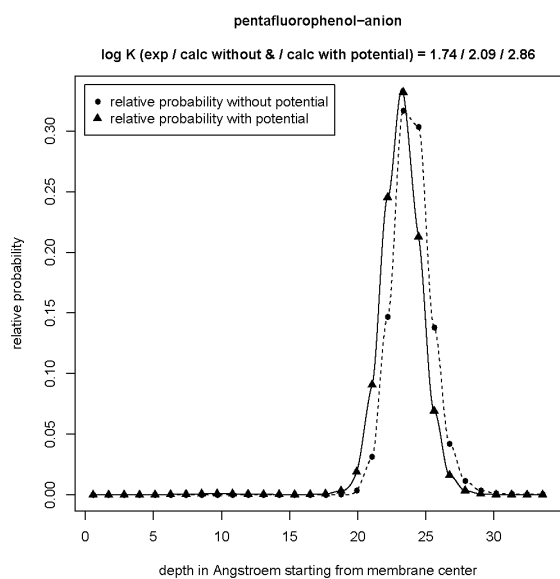
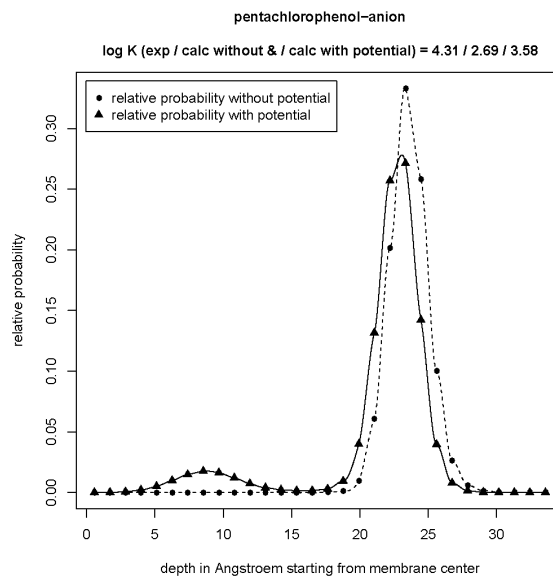
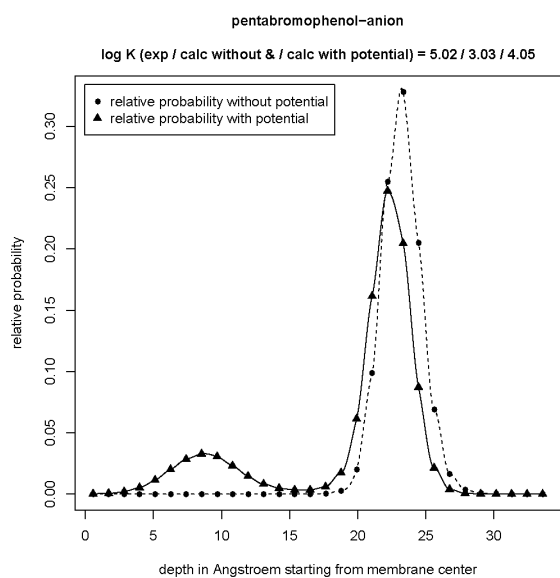
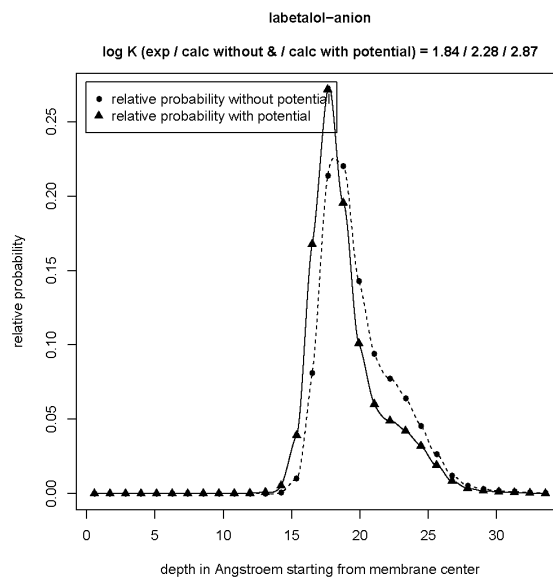
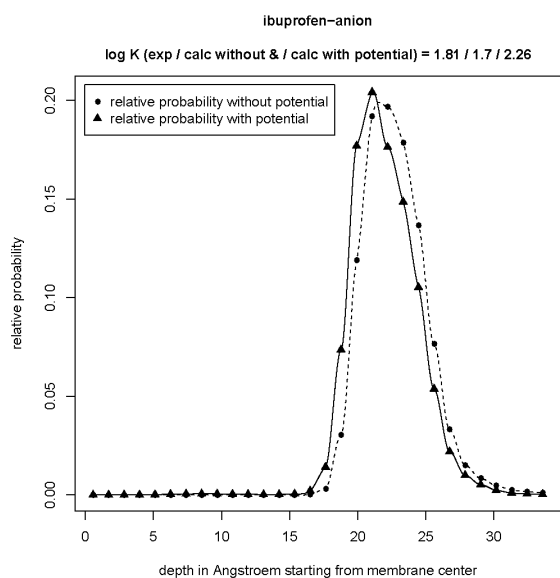


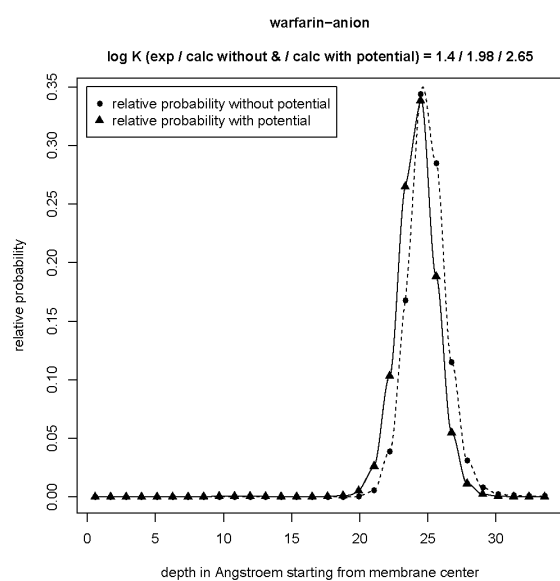
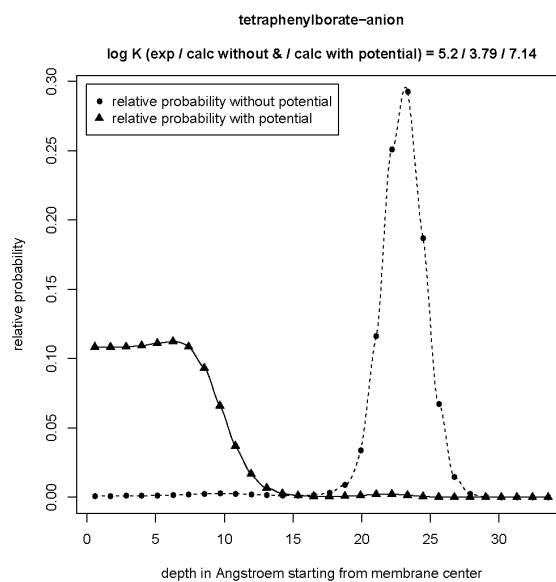
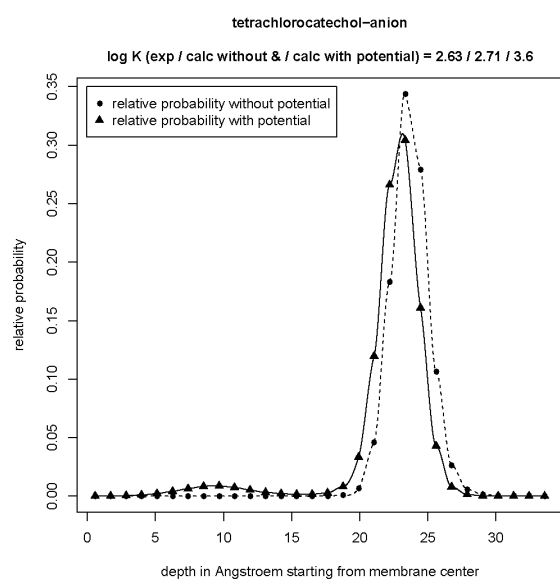




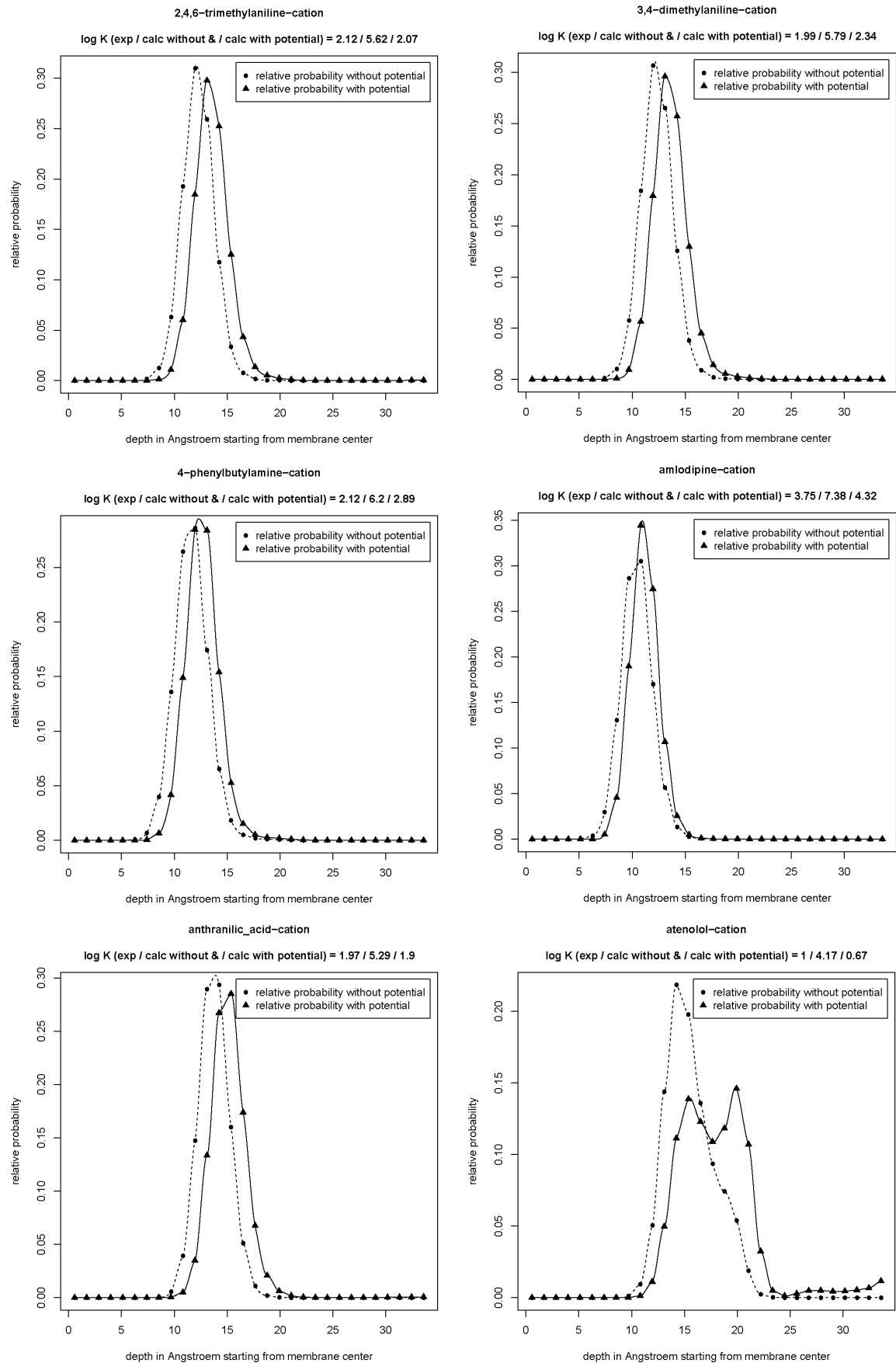


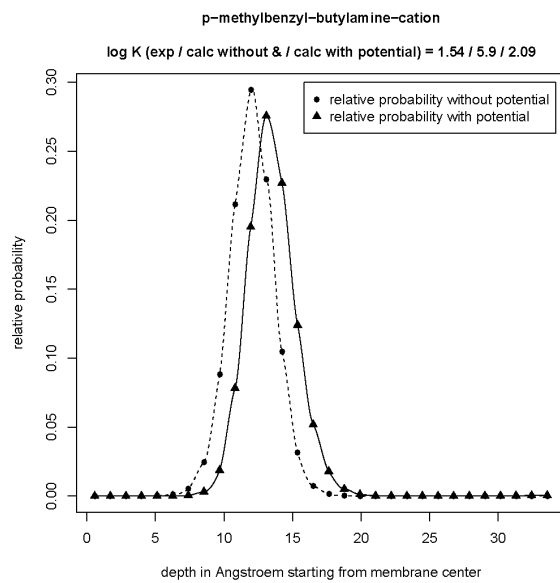
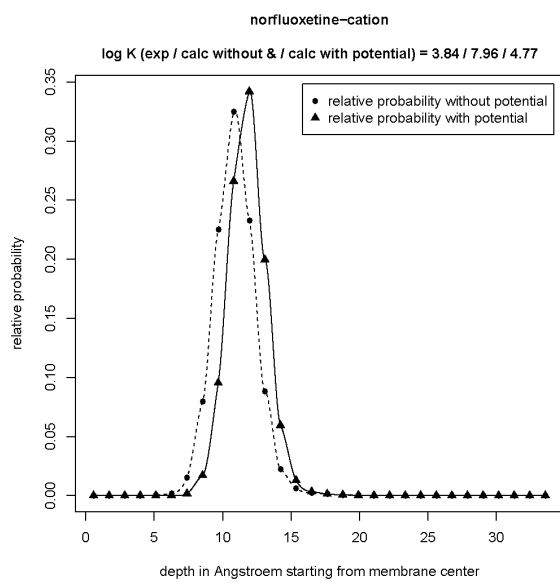
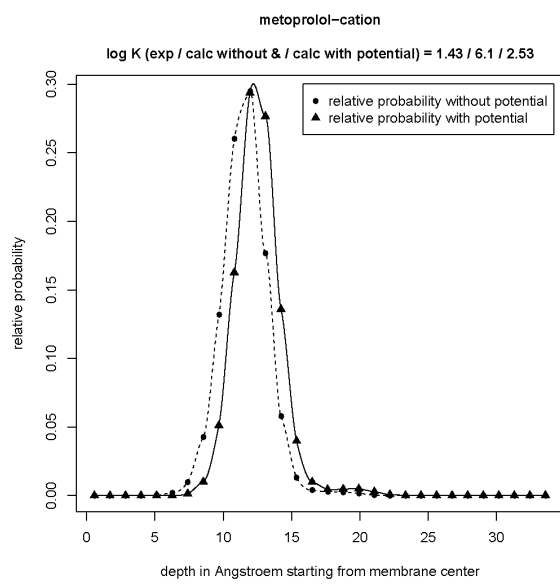
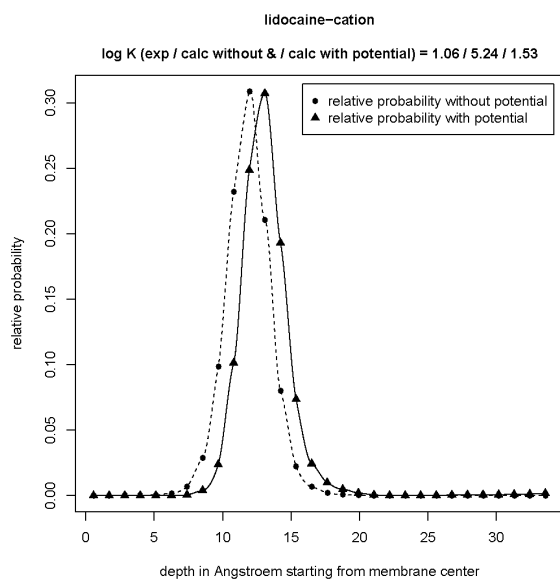
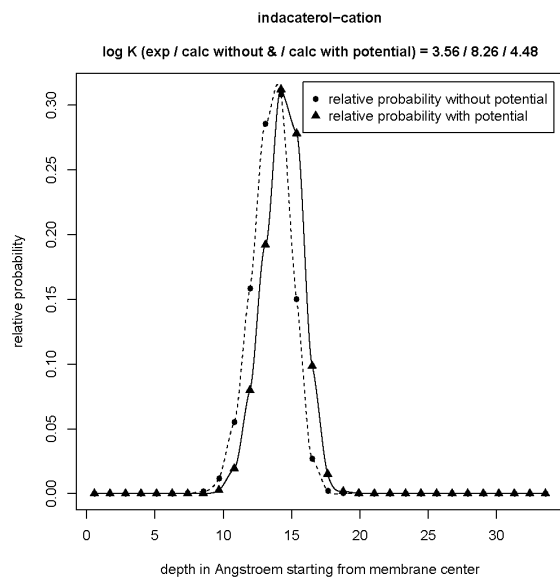
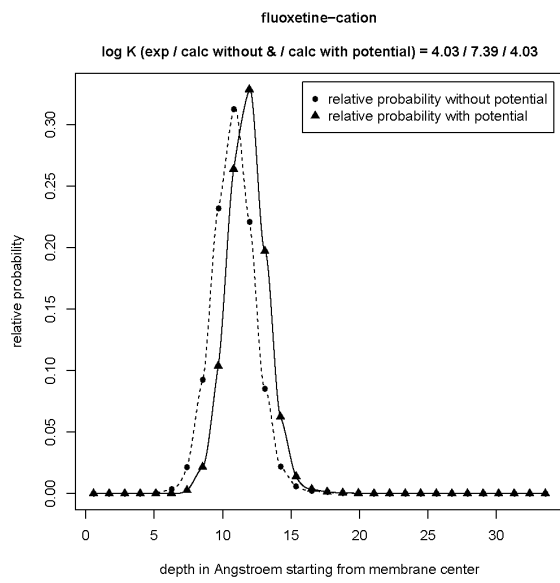


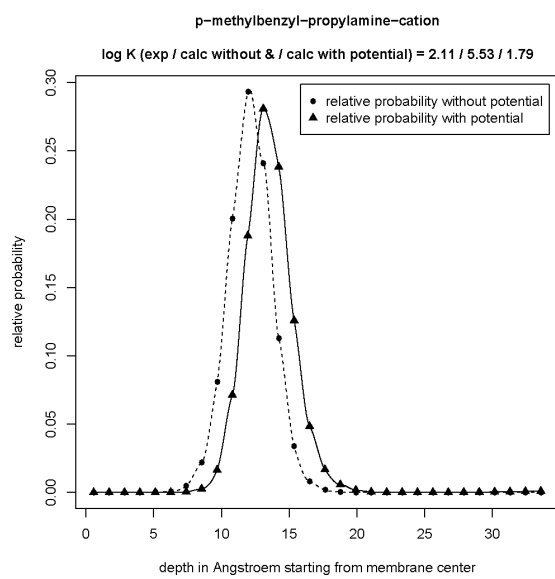
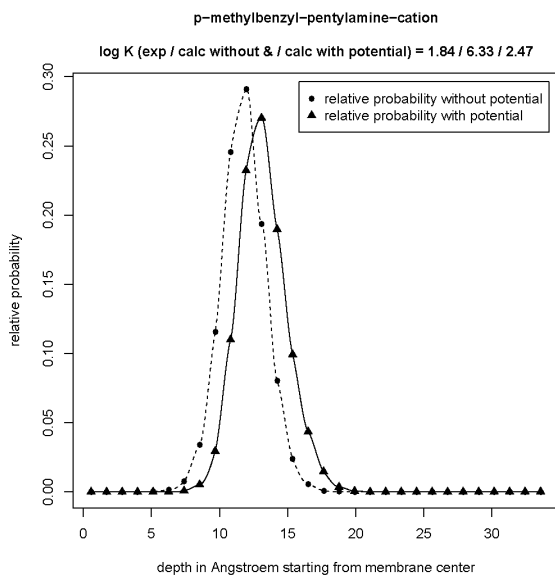
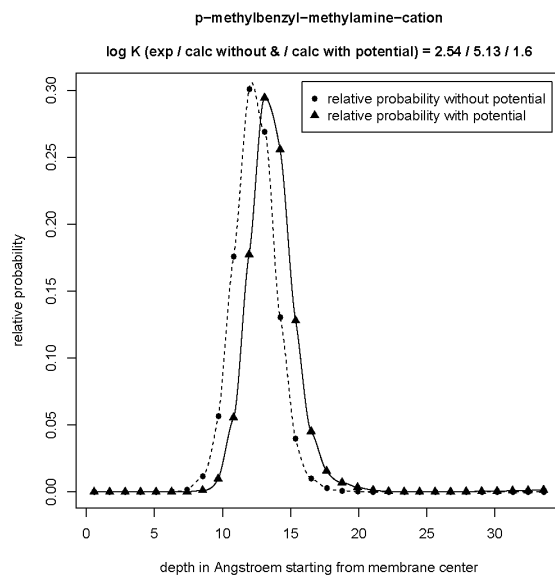
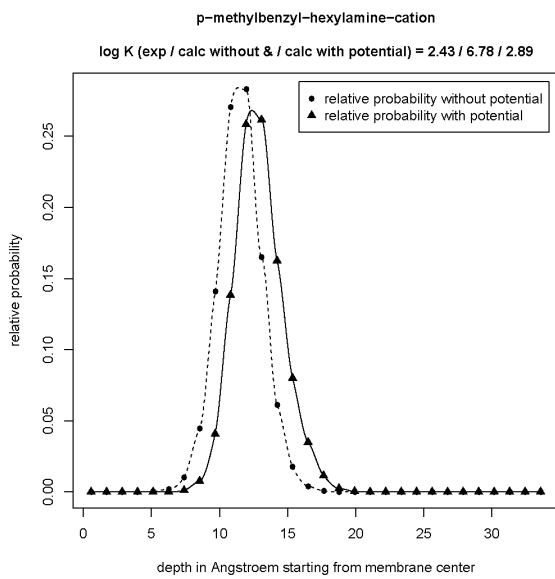
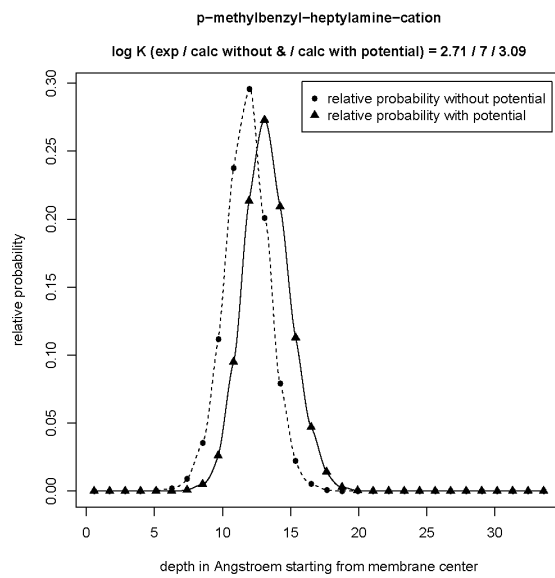
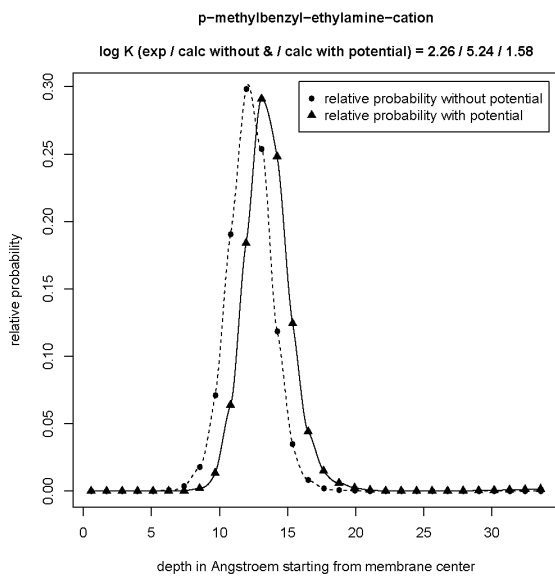


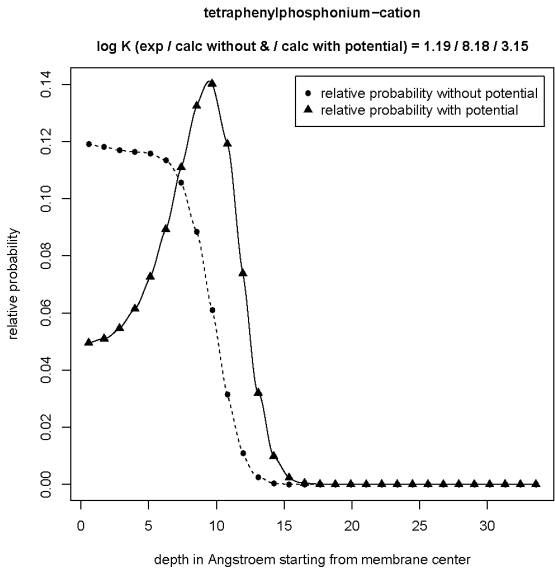
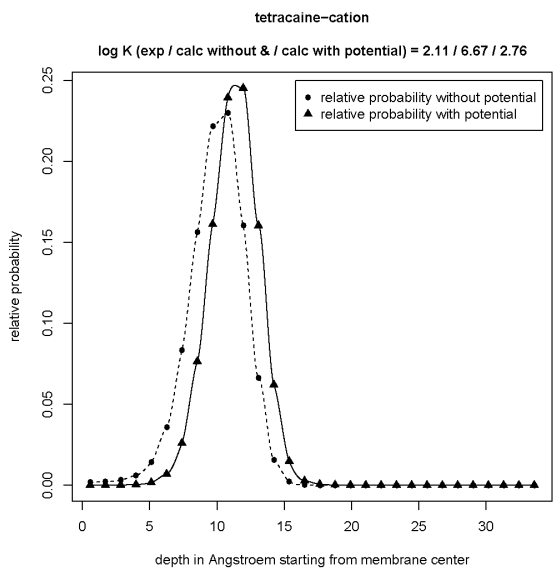
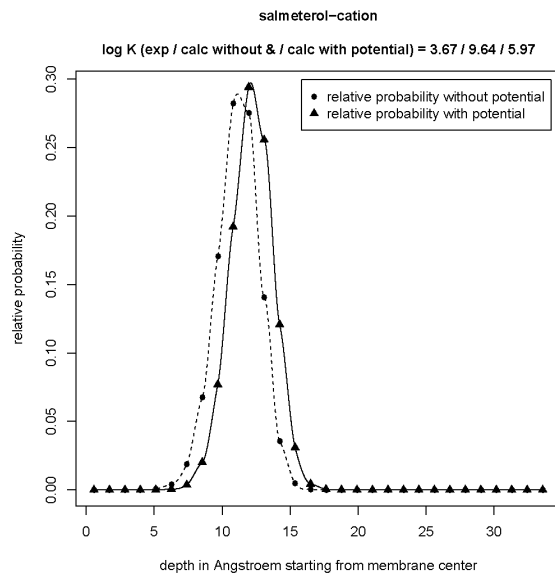
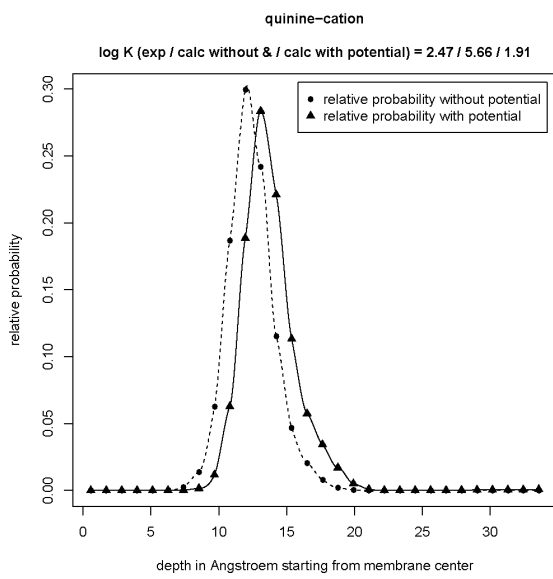
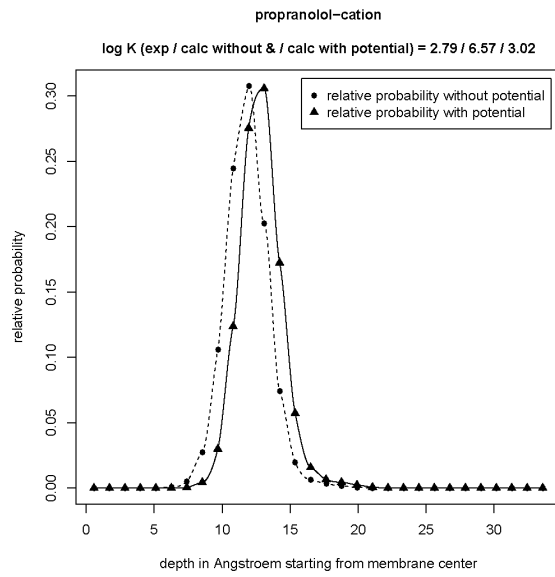
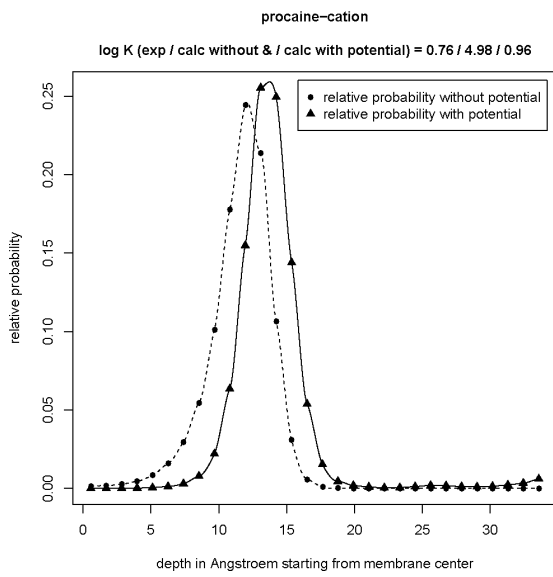


## 2.5 Influence of the membrane potential on the relative distribution of the cations









## 2.6 Predictions with the different models

SI-1, Table 3: Log K predictions obtained with the different models presented in the summary, section 1.3 are listed. For comparison with the experimental log  $K_{lipw}$  values, the values from Table 1 and 2 from SI-1 have been taken – if several values are listed, the arithmetic mean of the log  $K_{lipw}$  values were taken as given here. POPC stands for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine and DMPC for 1,2-Dimyristoyl-sn-glycero-3-phosphocholine.

abbreviation	charge	Training Set (Tr), Test Set (Va)	log $K_{lipw}$ (ion exp)	bulk POPC	bulk DMPC	POPC using COSMOmic without potential	DMPC: COSMOmic without pot. (Fig 1, summary)	POPC (1 Gauss potential) - model M1	DMPC (1 Gauss potential) - model M2	DMPC training set (1 Gauss potential) - model M2a	DMPC (2 Gauss potentials) - model M3	DMPC with 0.1 M KCl (1 Gauss potential) - model
2345Te CP	-	Tr	3.69	-5.45	-5.33	2.65	2.51	3.33	3.34	NA	3.59	3.37
2346Te CP	-	Tr	3.46	-5.04	-4.92	2.67	2.52	3.36	3.36	NA	3.65	3.39
2356Te CP	-	Tr	3.49	-4.79	-4.67	2.73	2.58	3.41	3.41	NA	3.67	3.44
245Tri CP	-	Tr	2.88	-6.42	-6.29	2.47	2.33	3.12	3.12	NA	3.31	3.15
246Tri BP	-	Tr	3.07	-5.08	-4.98	2.73	2.58	3.44	3.43	NA	3.79	3.45
246Tri CP	-	Va	2.52	-6.12	-5.98	2.45	2.32	3.10	3.12	3.04	3.33	3.14
24DCP	-	Tr	2.69	-7.89	-7.75	2.18	2.05	2.79	2.81	NA	2.91	2.84
24DNP	-	Va	1.90	-6.14	-5.98	1.80	1.66	2.40	2.40	2.33	2.70	2.43
26DCP	-	Tr	1.41	-7.36	-7.21	2.22	2.08	2.80	2.81	NA	2.91	2.84
26DNP	-	Tr	1.86	-7.02	-6.86	1.81	1.68	2.31	2.33	NA	2.30	2.35
2CP	-	Tr	0.92	-9.38	-9.23	1.83	1.72	2.38	2.41	NA	2.36	2.44
DNOC	-	Tr	2.35	-5.49	-5.34	1.91	1.78	2.52	2.52	NA	2.85	2.55
2NP	-	Tr	0.69	-9.66	-9.49	1.50	1.40	1.96	2.01	NA	1.70	2.05
Dinose b	-	Va	3.35	-3.92	-3.78	2.23	2.09	2.87	2.87	2.80	3.34	2.91
Dino2te rb	-	Va	3.59	-3.57	-3.43	2.32	2.18	2.98	2.98	2.92	3.48	3.02
345Tri CP	-	Va	3.16	-6.82	-6.69	2.45	2.32	3.10	3.11	3.04	3.30	3.15
34DCP	-	Va	2.85	-8.23	-8.09	2.19	2.06	2.80	2.82	2.75	2.94	2.85
34DNP	-	Tr	1.90	-5.83	-5.68	2.02	1.88	2.66	2.65	NA	3.22	2.68
Bromo	-	Tr	2.10	-5.61	-5.49	2.02	1.89	2.68	2.68	NA	3.12	2.72



x												
35DBC	-	Tr	3.18	-7.08	-6.97	2.51	2.37	3.15	3.16	NA	3.33	3.19
35DCP	-	Tr	2.47	-7.54	-7.40	2.30	2.17	2.93	2.94	NA	3.03	2.97
TTFB	-	Tr	4.35	-1.45	-1.35	3.16	2.99	4.27	4.33	NA	4.76	4.31
4CP	-	Tr	2.51	-9.95	-9.80	1.84	1.73	2.41	2.44	NA	2.45	2.47
DNPC	-	Va	2.26	-6.97	-6.81	1.83	1.70	2.30	2.33	2.26	2.28	2.36
4NP	-	Va	0.95	-8.84	-8.68	1.37	1.28	1.89	1.93	1.86	1.97	1.97
OBS	-	Tr	3.63	-8.70	-8.54	3.30	3.03	3.65	3.55	NA	3.56	3.60
Dino4te	-	Tr	3.23	-5.77	-5.62	2.16	2.04	2.64	2.68	NA	2.92	2.73
rb												
DTFB	-	Va	3.05	-3.67	-3.55	2.70	2.55	3.43	3.44	3.38	3.89	3.47
Dibrom	-	Tr	3.03	-7.59	-7.48	2.40	2.26	3.01	3.01	NA	3.13	3.05
ox												
Dichlor	-	Tr	2.47	-8.26	-8.13	2.20	2.08	2.77	2.80	NA	2.83	2.84
ox												
S-13	-	Va	5.05	1.71	1.78	3.81	3.61	6.72	6.68	6.96	6.97	6.67
Chloro	-	Va	1.91	-10.07	-9.93	1.93	1.82	2.46	2.50	2.43	2.44	2.54
x												
5-NB	-	Tr	1.81	-4.15	-4.01	2.40	2.24	3.08	3.07	NA	3.45	3.10
5-PA	-	Tr	1.66	-14.21	-14.07	1.38	1.31	1.77	1.86	NA	1.70	1.92
Oxine	-	Tr	1.47	-11.73	-11.59	1.60	1.49	2.06	2.10	NA	1.87	2.15
AA	-	Tr	0.31	-12.87	-12.69	1.28	1.19	1.69	1.76	NA	1.37	1.80
CCCP	-	Va	4.05	-3.19	-3.07	2.62	2.48	3.35	3.34	3.27	3.95	3.38
FCCP	-	Tr	4.22	-2.71	-2.60	2.73	2.59	3.65	3.60	NA	4.34	3.63
Dic	-	Tr	2.64	-8.37	-8.25	2.44	2.32	2.91	2.94	NA	3.07	3.01
Dif	-	Va	2.73	-7.48	-7.35	2.25	2.13	2.80	2.83	2.76	2.81	2.87
Fen	-	Va	2.28	-8.70	-8.57	2.23	2.11	2.73	2.76	2.70	2.73	2.81
Flu	-	Tr	3.61	-7.48	-7.36	2.35	2.24	2.99	3.01	NA	3.26	3.05
Ibu	-	Va	1.81	-12.07	-11.94	1.78	1.70	2.19	2.26	2.21	2.25	2.34
Lab	-	Tr	1.84	-11.38	-11.20	2.47	2.28	2.92	2.87	NA	3.16	2.94
PBrP	-	Tr	5.02	-2.96	-2.90	3.21	3.03	4.08	4.05	NA	4.56	4.06
PCP	-	Va	4.31	-4.15	-4.04	2.83	2.69	3.57	3.58	3.54	3.92	3.61
PFP	-	Va	1.74	-6.96	-6.83	2.22	2.09	2.85	2.86	2.79	2.99	2.90
SA	-	Tr	0.97	-10.10	-9.93	1.54	1.44	2.00	2.05	NA	1.72	2.10
TeCC	-	Tr	2.63	-3.64	-3.51	2.87	2.71	3.60	3.60	NA	3.95	3.62
TPB	-	Tr	5.20	2.71	2.80	3.94	3.79	7.27	7.14	NA	7.16	7.10
Warf	-	Tr	1.40	-5.84	-5.70	2.11	1.98	2.63	2.65	NA	2.92	2.70
246TM	+	Va	2.12	8.02	7.90	5.57	5.62	2.06	2.07	2.15	2.34	2.03
A												
34DM	+	Tr	1.99	8.41	8.28	5.75	5.79	2.38	2.34	NA	2.65	2.24
A												
4-	+	Va	2.12	8.63	8.55	6.23	6.20	3.00	2.89	3.01	2.99	2.78
PhenBu												
tA												
Amlodi	+	Tr	3.75	10.01	9.92	7.38	7.38	4.43	4.32	NA	4.27	4.23

P												
AA	+	Tr	1.97	8.49	8.32	5.18	5.29	1.84	1.90	NA	1.88	1.82
Aten	+	Tr	1.01	6.25	6.25	4.17	4.17	0.65	0.67	NA	0.87	0.67
Fluox	+	Va	4.03	9.56	9.47	7.47	7.39	4.14	4.03	4.16	3.97	4.02
Indac	+	Va	3.56	10.86	10.72	8.36	8.26	4.55	4.48	4.57	4.33	4.43
Lido	+	Va	1.07	6.78	6.77	5.26	5.24	1.50	1.53	1.61	1.53	1.65
Metro	+	Tr	1.43	8.36	8.31	6.05	6.10	2.48	2.53	NA	2.63	2.57
Norfluo	+	Tr	3.84	10.76	10.64	8.07	7.96	4.96	4.77	NA	4.87	4.65
x												
MBBut	+	Tr	1.54	7.99	7.93	5.85	5.90	2.02	2.09	NA	2.16	2.18
A												
MBEth	+	Va	2.26	7.03	6.97	5.23	5.24	1.55	1.58	1.66	1.61	1.62
A												
MBHep	+	Tr	2.71	9.43	9.35	6.88	7.00	2.96	3.09	NA	3.00	3.24
A												
MBHex	+	Tr	2.43	8.96	8.89	6.74	6.78	2.82	2.89	NA	2.96	2.99
A												
MBMet	+	Tr	2.54	6.90	6.85	5.13	5.13	1.60	1.60	NA	1.62	1.59
A												
MBPen	+	Va	1.84	8.47	8.40	6.29	6.33	2.41	2.47	2.53	2.56	2.57
tA												
MBPro	+	Tr	2.11	7.46	7.40	5.50	5.53	1.74	1.79	NA	1.86	1.85
pA												
Proc	+	Va	0.76	6.42	6.42	5.02	4.98	0.89	0.96	0.99	0.80	1.13
Prop	+	Va	2.79	8.98	8.92	6.58	6.57	3.04	3.02	3.11	3.16	3.03
Quinine	+	Tr	2.47	7.92	7.83	5.54	5.66	1.78	1.91	NA	2.04	2.01
Salmet	+	Tr	3.67	13.45	13.32	9.48	9.64	5.87	5.97	NA	6.09	6.00
Tetrac	+	Tr	2.11	7.60	7.60	6.80	6.67	2.80	2.76	NA	2.70	2.91
TPP	+	Va	1.19	7.76	7.62	8.37	8.18	3.28	3.15	2.94	3.19	3.39

**SI-1, Table 4: Smiles for all investigated anionic and cationic compounds.**

Compoundname	smiles (ion)
2,3,4,5-tetrachlorophenol-anion	[O-]C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1
2,3,4,6-tetrachlorophenol-anion	[O-]C1=C(Cl)C=C(Cl)C(Cl)=C1Cl
2,3,5,6-tetrachlorophenol-anion	[O-]C1=C(Cl)C(Cl)=CC(Cl)=C1Cl
2,4,5-trichlorophenol-anion	[O-]C1=CC(Cl)=C(Cl)C=C1Cl
2,4,6-tribromophenol-anion	[O-]c1c(Br)cc(Br)cc1Br
2,4,6-trichlorophenol-anion	[O-]C1=C(Cl)C=C(Cl)C=C1Cl
2,4-dichlorophenol-anion	[O-]C1=C(Cl)C=C(Cl)C=C1
2,4-dinitrophenol-anion	[O-]C1=CC=C(C=C1[N+](O-)=O)[N+](O-)=O
2,6-dichlorophenol-anion	[O-]C1=C(Cl)C=CC=C1Cl
2,6-dinitrophenol-anion	[O-]C1=C(C=CC=C1[N+](O-)=O)[N+](O-)=O
2-chlorophenol-anion	[O-]C1=C(Cl)C=CC=C1
2-methyl-4,6-dinitrophenol-anion	CC1=C([O-])C(=CC(=C1)[N+](O-)=O)[N+](O-)=O
2-nitrophenol-anion	[O-]C1=C(C=CC=C1)[N+](O-)=O
2-s-butyl-4,6-dinitrophenol-anion	CCC(C)C1=C([O-])C(=CC(=C1)[N+](O-)=O)[N+](O-)=O
2-tert-butyl-4,6-dinitrophenol-anion	CC(C)(C)C1=C([O-])C(=CC(=C1)[N+](O-)=O)[N+](O-)=O
3,4,5-trichlorophenol-anion	[O-]C1=CC(Cl)=C(Cl)C(Cl)=C1
3,4-dichlorophenol-anion	[O-]C1=CC(Cl)=C(Cl)C=C1
3,4-dinitrophenol-anion	[O-]C1=CC(=C(C=C1)[N+](O-)=O)[N+](O-)=O
3,5-dibromo-4-hydroxy-benzonitrile-anion	[O-]C1=C(Br)C=C(C=C1Br)C#N
3,5-dibromo-4-methylphenol-anion	CC1=C(Br)C=C([O-])C=C1Br
3,5-dichlorophenol-anion	[O-]C1=CC(Cl)=CC(Cl)=C1
4,5,6,7-tetrachloro-2-(trifluoromethyl)-1H-benzimidazole-anion	FC(F)(F)C1=NC2=C(Cl)C(Cl)=C(Cl)C(Cl)=C2[N-]1
4-chlorophenol-anion	[O-]C1=CC=C(Cl)C=C1
4-methyl-2,6-dinitrophenol-anion	CC1=CC(=C([O-])C(=C1)[N+](O-)=O)[N+](O-)=O
4-nitrophenol-anion	[O-]C1=CC=C(C=C1)[N+](O-)=O
4-octylbenzene-1-sulfonate-anion	CCCCCCCCC1=CC=C(C=C1)S([O-])(=O)=O
4-tert-butyl-2,6-dinitrophenol-anion	CC(C)(C)C1=CC(=C([O-])C(=C1)[N+](O-)=O)[N+](O-)=O
5,6-dichloro-2-(trifluoromethyl)-benzimidazole-anion	FC(F)(F)C1=NC2=CC(Cl)=C(Cl)C=C2[N-]1
5,7-dibromo-8-hydroxyquinoline-anion	[O-]C1=C(Br)C=C(Br)C2=CC=CN=C12
5,7-dichloro-8-hydroxyquinoline-anion	[O-]C1=C(Cl)C=C(Cl)C2=CC=CN=C12
5-chloro-3-tert-butyl-2'-chloro-4'-nitrosalicylanilide-anion	CC(C)(C)C1=CC(Cl)=CC(C(=O)[N-]C2=C(Cl)C=C(C=C2)[N+](O-)=O)=C1O
5-chloro-8-hydroxyquinoline-anion	[O-]C1=CC=C(Cl)C2=CC=CN=C12
5-nitro-2-trifluoromethylbenzimidazole-anion	[O-][N+](=O)C1=CC=C2[N-]C(=NC2=C1)C(F)(F)F
5-phenylvaleric acid-anion	[O-]C(=O)CCCCC1=CC=CC=C1
8-hydroxyquinoline-anion	[O-]C1=C2N=CC=CC2=CC=C1
anthranilic acid-anion	Nc1cccc1C([O-])=O
carbonyl cyanide m-chlorophenylhydrazone-anion	ClC1=CC=CC([N-]N=C(C#N)C#N)=C1
carbonyl cyanide	FC(F)(F)OC1=CC=C([N-]N=C(C#N)C#N)C=C1

methoxyphenylhydrazone-anion	
diclofenac-anion	<chem>[O-]C(=O)CC1=C(NC2=C(Cl)C=CC=C2Cl)C=CC=C1</chem>
diflunisal-anion	<chem>O=C([O-])c1cc(ccc1O)c2ccc(F)cc2F</chem>
fenamic acid-anion	<chem>[O-]C(=O)C1=C(NC2=CC=CC=C2)C=CC=C1</chem>
flufenamic acid-anion	<chem>[O-]C(=O)C1=CC=CC=C1NC1=CC(=CC=C1)C(F)(F)F</chem>
ibuprofen-anion	<chem>CC(C)CC1=CC=C(C=C1)C(C)C([O-])=O</chem>
labetalol-anion	<chem>CC(Cc1ccccc1)NCC(O)c1ccc([O-])c(c1)C(N)=O</chem>
pentabromophenol-anion	<chem>[O-]C1=C(Br)C(Br)=C(Br)C(Br)=C1Br</chem>
pentachlorophenol-anion	<chem>[O-]C1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1Cl</chem>
pentafluorophenol-anion	<chem>[O-]C1=C(F)C(F)=C(F)C(F)=C1F</chem>
salicylic acid-anion	<chem>OC1=C(C=CC=C1)C([O-])=O</chem>
tetrachlorocatechol-anion	<chem>OC1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1[O-]</chem>
tetraphenylborate-anion	<chem>C1=CC=C(C=C1)[B-](C1=CC=CC=C1)(C1=CC=CC=C1)C1=CC=CC=C1</chem>
warfarin-anion	<chem>CC(=O)CC(C1=CC=CC=C1)C1=C([O-])C2=C(OC1=O)C=CC=C2</chem>
2,4,6-trimethylaniline-cation	<chem>CC1=CC(C)=C([NH3+])C(C)=C1</chem>
3,4-dimethylaniline-cation	<chem>CC1=C(C)C=C([NH3+])C=C1</chem>
4-phenylbutylamine-cation	<chem>[NH3+]CCCC1=CC=CC=C1</chem>
amlodipine-cation	<chem>CCOC(=O)C1=C(COCC[NH3+])NC(C)=C(C1C1=CC=CC=C1Cl)C(=O)OC</chem>
anthranilic acid-cation	<chem>[NH3+]C1=C(C=CC=C1)C(O)=O</chem>
atenolol-cation	<chem>CC(C)[NH2+]CC(O)COC1=CC=C(CC(N)=O)C=C1</chem>
fluoxetine-cation	<chem>C[NH2+]CCC(OC1=CC=C(C=C1)C(F)(F)F)C1=CC=CC=C1</chem>
indacaterol-cation	<chem>CCC1=CC2=C(CC(C2)[NH2+]CC(O)C2=CC=C(O)C3=C2C=CC(=O)N3)C=C1CC</chem>
lidocaine-cation	<chem>CC[NH+](CC)CC(=O)NC1=C(C)C=CC=C1C</chem>
metoprolol-cation	<chem>COCCC1=CC=C(OCC(O)C[NH2+]C(C)C)C=C1</chem>
norfluoxetine-cation	<chem>[NH3+]CCC(OC1=CC=C(C=C1)C(F)(F)F)C1=CC=CC=C1</chem>
p-methylbenzyl-butylamine-cation	<chem>CCCC[NH2+]CC1=CC=C(C)C=C1</chem>
p-methylbenzyl-ethylamine-cation	<chem>CC[NH2+]CC1=CC=C(C)C=C1</chem>
p-methylbenzyl-heptylamine-cation	<chem>CCCCCCC[NH2+]Cc1ccc(C)cc1</chem>
p-methylbenzyl-hexylamine-cation	<chem>CCCCCC[NH2+]Cc1ccc(C)cc1</chem>
p-methylbenzyl-methylamine-cation	<chem>C[NH2+]CC1=CC=C(C)C=C1</chem>
p-methylbenzyl-pentylamine-cation	<chem>CCCCC[NH2+]Cc1ccc(C)cc1</chem>
p-methylbenzyl-propylamine-cation	<chem>CCC[NH2+]CC1=CC=C(C)C=C1</chem>
procaine-cation	<chem>CC[NH+](CC)CCOC(=O)C1=CC=C(N)C=C1</chem>
propranolol-cation	<chem>CC(C)[NH2+]CC(O)COC1=CC=CC2=C1C=CC=C2</chem>
quinine-cation	<chem>COC1=CC2=C(C=CN=C2C=C1)C(O)C1CC2CC[NH+]1CC2C=C</chem>
salmeterol-cation	<chem>OCC1=C(O)C=CC(=C1)C(O)C[NH2+]CCCCCOC1=CC=CC=C1</chem>
tetracaine-cation	<chem>CCCCNC1=CC=C(C=C1)C(=O)OCC[NH+](C)C</chem>
tetraphenylphosphonium-cation	<chem>C1=CC=C(C=C1)[P+](C1=CC=CC=C1)(C1=CC=CC=C1)C1=CC=CC=C1</chem>

---

**Supporting Information 2: Comparison of different models predicting the phospholipid-membrane water partition coefficients of neutral and charged compounds**

**1 Data selection**

## 1.1 Data collection cations

SI-2, Table 1. Data collection for cations based on a previously published data collection (Bittermann et al., 2014), with all new values marked in bold font. Multiple values for  $pK_a$ ,  $\log K_{lipw}$  (neutral) and  $\log K_{lipw}$  (ion) are marked in grey; in these cases the arithmetic mean was used for the calculation of  $\Delta mw$ . ‘P’ stands for values taken from the PhysProp-Database (<http://esc.syrres.com/fatepointer/search.asp>), egg-PC for egg-phosphatidylcholine, DOPC for dioleoylphosphatidylcholine, DPPC for 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine and POPC for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine.

CAS	compoundname	abbreviation	$pK_a$	$\log K_{lipw}$ (neutral) [L/kg]	$\log K_{lipw}$ (ion) [L/kg]	$\Delta mw$	lipid	method	T [°C]
88-05-1	2,4,6-trimethylaniline	246TMA	4.38 (Escher et al., 2000)	<b>2.38</b> (Escher et al., 2000)	2.12 (Escher et al., 2000)	<b>0.26</b>	DPPC/ DOPC	equilibrium dialysis	20
95-68-1	3,4-dimethylaniline	34DMA	5.23 (Escher et al., 2000)	<b>2.11</b> (Escher et al., 2000)	1.99 (Escher et al., 2000)	<b>0.12</b>	DPPC/ DOPC	equilibrium dialysis	20
13214-66-9	4-phenylbutylamine	4-PhenButA	10.54 (Austin et al., 1995)	<b>2.41</b> (Austin et al., 1995)	2.12 (Austin et al., 1995)	<b>0.29</b>	DMPC	ultrafiltration	37
<b>37517-30-9</b>	<b>acebutolol</b>	<b>ABL</b>	<b>9.67</b> (Betageri and Rogers, 1987)		<b>0.66</b> (Betageri and Rogers, 1987)		<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
<b>13655-52-2</b>	<b>alprenolol</b>	<b>APL</b>	<b>9.70</b> (Betageri and Rogers, 1987)		<b>2.17</b> (Betageri and Rogers, 1987)		<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
88150-42-9	amlodipine	Amlodip	9.02 (Austin et al., 1995)	<b>3.75</b> (Austin et al., 1995)	3.75 (Austin et al., 1995)	<b>0.00</b>	DMPC	ultrafiltration	37
118-92-3	anthranilic acid	AA	2.15 (Thomae et al., 2007)	<b>2.08</b> (Thomae et al., 2007)	1.97 (Thomae et al., 2007)	<b>0.11</b>	egg-PC	equilibrium dialysis	NA
29122-68-7	atenolol	Aten	9.55 (Betageri and Rogers, 1987)		0.51 (Escher et al., 2006)		POPC	equilibrium dialysis	NA
					1.50 (Yamamoto et al., 2005)		POPC	equilibrium dialysis	NA
					<b>1.03</b> (Betageri and Rogers, 1987)		<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>

23284-25-5	bupranolol	BPL	9.60 (Betageri and Rogers, 1987)		Rogers, 1987) 2.49 (Betageri and Rogers, 1987)		DMPC	ultrafiltration	30
83881-51-0	ceterizine	Cet_c	8.00 (Plemper van Balen et al., 2001)		3.20 (Plemper van Balen et al., 2001)		egg-PC	equilibrium dialysis	25
50-53-3	chlorpromazine	CLP	9.28 (Barzanti et al., 2007) 9.24 (Pallicer and Krämer, 2012)	5.10 (Barzanti et al., 2007)	3.10 (Barzanti et al., 2007) 3.69 (Pallicer and Krämer, 2012)	1.71	egg-PC	potentiometric titration	25
54910-89-3	fluoxetine	Fluox	10.06 (Brooks et al., 2003)		4.08 (Neuwoehner et al., 2009) 3.79 (Yamamoto et al., 2005) 4.23 (Nakamura et al., 2008)		POPC	equilibrium dialysis	NA
68-88-2	hydroxyzine	Hyd	7.49 (Plemper van Balen et al., 2001)	3.40 (Plemper van Balen et al., 2001)	2.80 (Plemper van Balen et al., 2001)	0.60	egg-PC	equilibrium dialysis	25
312753-06-3	indacaterol	Indac	6.7 (Lombardi et al., 2009)		3.56 (Lombardi et al., 2009)		DMPC	equilibrium dialysis	37
36894-69-6	labetalol	Lab_c	7.35 (Pallicer and Krämer, 2012)	2.73 (Pallicer and Krämer, 2012)	2.32 (Pallicer and Krämer, 2012)	0.40	egg-PC	equilibrium dialysis	25
137-58-6	lidocaine	Lido	7.86 (Ottiger and Wunderli-Allenspach, 1997) 7.96 (Avdeef et al., 1998)	2.06 (Ottiger and Wunderli-Allenspach, 1997) 2.39 (Avdeef et al., 1998)	0.91 (Ottiger and Wunderli-Allenspach, 1997) 1.22 (Avdeef et al., 1998)	1.16	egg-PC	equilibrium dialysis	37
51384-51-1	metoprolol	Metro	9.7 (Betageri and Rogers, 1987)		1.43 (Escher et al., 2006)		DOPC	potentiometric titration	25
							POPC	equilibrium dialysis	NA

					<b>1.13</b> (Betageri and Rogers, 1987)	<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
<b>20574-50-9</b>	<b>morantel</b>	<b>Mor</b>	<b>11.91</b> (Escher et al., 2008)		<b>2.00</b> (Escher et al., 2008)	<b>POPC</b>	<b>equilibrium dialysis</b>	<b>20</b>
<b>42200-33-9</b>	<b>nadolol</b>	<b>NDL</b>	<b>9.67</b> (Betageri and Rogers, 1987)		<b>0.95</b> (Betageri and Rogers, 1987)	<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
83891-03-6	norfluoxetine	Norfluox	9.05 (Brooks et al., 2003)		3.84 (Neuwoehner et al., 2009)	POPC	equilibrium dialysis	NA
<b>6452-71-7</b>	<b>oxprenolol</b>	<b>OPL</b>	<b>9.50</b> (Betageri and Rogers, 1987)		<b>1.51</b> (Betageri and Rogers, 1987)	<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
16183-21-4	p-methylbenzyl-butylamine	MBButA	9.98 (Fruttero et al., 1998)	<b>3.05</b> (Fruttero et al., 1998)	1.54 (Fruttero et al., 1998)	<b>1.51</b> egg-PC	potentiometric titration	NA
39099-13-3	p-methylbenzyl-ethylamine	MBEthA	10.04 (Fruttero et al., 1998)	<b>3.06</b> (Fruttero et al., 1998)	2.26 (Fruttero et al., 1998)	<b>0.80</b> egg-PC	potentiometric titration	NA
215177-24-5	p-methylbenzyl-hepotentiometric titrationylamine	MBHepA	10.02 (Fruttero et al., 1998)	<b>4.40</b> (Fruttero et al., 1998)	2.71 (Fruttero et al., 1998)	<b>1.69</b> egg-PC	potentiometric titration	NA
215177-23-4	p-methylbenzyl-hexylamine	MBHexA	10.17 (Fruttero et al., 1998)	<b>4.20</b> (Fruttero et al., 1998)	2.43 (Fruttero et al., 1998)	<b>1.77</b> egg-PC	potentiometric titration	NA
699-04-7	p-methylbenzyl-methylamine	MBMetA	9.93 (Fruttero et al., 1998)	<b>3.09</b> (Fruttero et al., 1998)	2.54 (Fruttero et al., 1998)	<b>0.55</b> egg-PC	potentiometric titration	NA
170303-38-5	p-methylbenzyl-pentylamine	MBPentA	10.08 (Fruttero et al., 1998)	<b>3.50</b> (Fruttero et al., 1998)	1.84 (Fruttero et al., 1998)	<b>1.66</b> egg-PC	potentiometric titration	NA
39190-96-0	p-methylbenzyl-propylamine	MBPropA	9.98 (Fruttero et al., 1998)	<b>3.07</b> (Fruttero et al., 1998)	2.11 (Fruttero et al., 1998)	<b>0.96</b> egg-PC	potentiometric titration	NA
<b>13523-86-9</b>	<b>pindolol</b>	<b>PDL</b>	<b>8.80</b> (Betageri and Rogers, 1987)		<b>1.40</b> (Betageri and Rogers, 1987)	<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
59-46-1	procaine	Proc	9.04 (Avdeef et al.,	<b>2.38</b> (Avdeef et al.,	0.76 (Avdeef et	<b>1.56</b> DOPC	potentiometric	25



			1998)	1998)	al., 1998)			titration	
				<b>2.20</b> (Barzanti et al., 2007)	<b>0.70</b> (Barzanti et al., 2007)		<b>egg-PC</b>	<b>potentiometric titration</b>	<b>25</b>
525-66-6	propranolol	Prop	9.24 (Ottiger and Wunderli-Allenspach, 1997)	<b>3.24</b> (Ottiger and Wunderli-Allenspach, 1997)	2.76 (Ottiger and Wunderli-Allenspach, 1997)	<b>0.58</b>	egg-PC	equilibrium dialysis	37
					3.06 (Escher et al., 2006)		POPC	equilibrium dialysis	NA
			9.45 (Pallicer and Krämer, 2012)		2.72 (Pallicer and Krämer, 2012)		egg-PC	equilibrium dialysis	25
			9.53 (Avdeef et al., 1998)		2.61 (Avdeef et al., 1998)		DOPC	potentiometric titration	25
					<b>2.68</b> (Betageri and Rogers, 1987)		<b>DMPC</b>	<b>ultrafiltration</b>	<b>30</b>
				<b>3.40</b> (Barzanti et al., 2007)	<b>2.60</b> (Barzanti et al., 2007)		<b>egg-PC</b>	<b>potentiometric titration</b>	<b>25</b>
130-95-0	quinine	Quinine	8.63 (Pallicer and Krämer, 2012)	<b>2.73</b> (Pallicer and Krämer, 2012)	2.47 (Pallicer and Krämer, 2012)	<b>0.53</b>	egg-PC	equilibrium dialysis	25
				<b>2.70</b> (Barzanti et al., 2007)	<b>1.90</b> (Barzanti et al., 2007)		<b>egg-PC</b>	<b>potentiometric titration</b>	<b>25</b>
89365-50-4	salmeterol	Salmet	8.8 (Lombardi et al., 2009)		3.67 (Lombardi et al., 2009)		DMPC	equilibrium dialysis	37
94-24-6	tetracaine	Tetrac	8.49 (Avdeef et al., 1998)	<b>3.23</b> (Avdeef et al., 1998)	2.11 (Avdeef et al., 1998)	<b>1.12</b>	DOPC	potentiometric titration	25
18198-39-5	Tetraphenyl-phosphonium	TPP			1.37 (Flewelling and Hubbell, 1986)		egg-PC	equilibrium dialysis	NA
					1.01 (Demura et		DPPC	electron	45

2933-94-0	toliprolol	TPL	9.60 (Betageri and Rogers, 1987)	al., 1987)	1.49 (Betageri and Rogers, 1987)	DMPC	paramagnetic resonance ultrafiltration	30
-----------	------------	-----	----------------------------------	------------	----------------------------------	------	---	----

## 1.2 Data collection anions

SI-2, Table 2. Data collection for anions based on a previously published data collection (Bittermann et al., 2014), with all new values marked in bold font. Multiple values for  $pK_a$ ,  $\log K_{lipw}$  (neutral) and  $\log K_{lipw}$  (ion) are marked in grey; in these cases the arithmetic mean was used for the calculation of  $\Delta mw$ . 'P' stands for values taken from the PhysProp-Database (<http://esc.syrres.com/fatepointer/search.asp>), egg-PC for egg-phosphatidylcholine, DOPC for dioleoylphosphatidylcholine, DPPC for 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine and POPC for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine.

CAS	compoundname	abbreviation	$pK_a$	$\log K_{lipw}$ (neutral) [L/kg]	$\log K_{lipw}$ (ion) [L/kg]	$\Delta mw$	lipid	method	T [°C]
4901-51-3	2,3,4,5-tetrachlorophenol	2345TeCP	6.35 (Schellenberg et al., 1984)	4.76 (Escher et al., 2000)	3.90 (Escher et al., 2000)	1.07	DPPC/	equilibrium	20
					3.48 (Smejtek et al., 1996)		DOPC	dialysis	
58-90-2	2,3,4,6-tetrachlorophenol	2346TeCP	5.40 (Schellenberg et al., 1984)	4.46 (Escher et al., 2000)	3.46 (Escher et al., 2000)	1.00	DPPC/	equilibrium	20
							DOPC	dialysis	
935-95-5	2,3,5,6-tetrachlorophenol	2356TeCP	5.14 <sup>P</sup>		3.49 (Smejtek et al., 1996)		egg-PC	electrophoretic mobility measurements	25
95-95-4	2,4,5-trichlorophenol	245TriCP	6.94 (Schellenberg et al., 1984)	4.46 (Escher et al., 2000)	2.98 (Escher et al., 2000)	1.58	DPPC/	equilibrium	20
					2.79 (Smejtek et al., 1996)		DOPC	dialysis	
							egg-PC	electrophoretic mobility measurements	25

118-79-6	2,4,6-tribromo-phenol	246TriBP	6.80 <sup>P</sup>		3.07 (Bittermann et al., 2014)	popc	equilibrium dialysis	25	
88-06-2	2,4,6-trichlorophenol	246TriCP	6.15 (Schellenberg et al., 1984)	<b>3.99</b> (Escher et al., 2000)	2.50 (Escher et al., 2000) 2.54 (Smejtek et al., 1996)	<b>1.47</b>	DPPC/ DOPC egg-PC	equilibrium dialysis electrophoretic mobility measurements	20 25
120-83-2	2,4-dichloro-phenol	24DCP	7.85 (Schellenberg et al., 1984)	<b>3.59</b> (Escher et al., 2000)	2.69 (Escher et al., 2000)	<b>0.90</b>	DPPC/ DOPC	equilibrium dialysis	20
51-28-5	2,4-dinitrophenol	24DNP	3.94 (Schwarzenbach et al., 1988)	<b>2.64</b> (Escher et al., 2000)	1.90 (Escher et al., 2000)	<b>0.74</b>	DPPC/ DOPC	equilibrium dialysis	20
<b>94-75-7</b>	<b>2,4-dichloro-phenoxyacetic acid</b>	<b>2,4-D</b>	<b>2.58</b> (Barzanti et al., 2007)	<b>3.60</b> (Barzanti et al., 2007)	<b>1.70</b> (Barzanti et al., 2007)	<b>1.90</b>	<b>egg-PC</b>	<b>potentiometric titration</b>	<b>25</b>
87-65-0	2,6-dichloro-phenol	26DCP	6.97 (Escher and Schwarzenbach, 1996)	<b>2.87</b> (Escher et al., 2000)	1.43 (Escher et al., 2000) 1.40 (Smejtek et al., 1996)	<b>1.46</b>	DPPC/ DOPC egg-PC	equilibrium dialysis electrophoretic mobility measurements	20 25
573-56-8	2,6-dinitrophenol	26DNP	3.70 (Escher and Schwarzenbach, 1996)	<b>2.03</b> (Escher et al., 2000)	1.86 (Escher et al., 2000)	<b>0.17</b>	DPPC/ DOPC	equilibrium dialysis	20
95-57-8	2-chlorophenol	2CP	8.56 (Escher and Schwarzenbach, 1996)	<b>2.79</b> (Escher et al., 2000)	0.92 (Escher et al., 2000)	<b>1.87</b>	DPPC/ DOPC	equilibrium dialysis	20
534-52-1	2-methyl-4,6-dinitrophenol	DNOC	4.31 (Schwarzenbach et al., 2000)	<b>2.76</b> (Escher et al., 2000)	2.35 (Escher et al., 2000)	<b>0.41</b>	DPPC/ DOPC	equilibrium dialysis	20

88-75-5	2-nitrophenol	2NP	7.23 (Schwarzenbach et al., 1988)	<b>1.89</b> (Escher et al., 2000)	0.69 (Escher et al., 2000)	<b>1.20</b>	DPPC/ DOPC	equilibrium dialysis	20
88-85-7	2-s-butyl-4,6-dinitrophenol	Dinoseb	4.62 (Schwarzenbach et al., 1988)	<b>3.96</b> (Escher et al., 2000)	3.35 (Escher et al., 2000)	<b>0.61</b>	DPPC/ DOPC	equilibrium dialysis	20
1420-07-1	2-tert-butyl-4,6-dinitrophenol	Dino2terb	4.80 (Miyoshi et al., 1987)	<b>4.10</b> (Escher et al., 2000)	3.59 (Escher et al., 2000)	<b>0.51</b>	DPPC/ DOPC	equilibrium dialysis	20
609-19-8	3,4,5-trichlorophenol	345TriCP	7.73 (Schellenberg et al., 1984)	<b>4.71</b> (Escher et al., 2000)	3.16 (Escher et al., 2000)	<b>1.55</b>	DPPC/ DOPC	equilibrium dialysis	20
95-77-2	3,4-dichlorophenol	34DCP	8.59 (Escher and Schwarzenbach, 1996)	<b>3.76</b> (Escher et al., 2000)	2.85 (Escher et al., 2000)	<b>0.91</b>	DPPC/ DOPC	equilibrium dialysis	20
577-71-9	3,4-dinitrophenol	34DNP	5.48 (Schwarzenbach et al., 2003)	<b>3.17</b> (Escher et al., 2000)	1.90 (Escher and Schwarzenbach, 1996)	<b>1.27</b>	DPPC/ DOPC	equilibrium dialysis	20
1689-84-5	3,5-dibromo-4-hydroxy-benzonitrile	Bromox	4.09 (Escher et al., 2001)	<b>3.16</b> (Escher et al., 2001)	2.10 (Escher et al., 2001)	<b>1.06</b>	POPC	TRANSIL	NA
13979-81-2	3,5-dibromo-4-methylphenol	35DBC	8.28 (Escher et al., 2001)	<b>4.51</b> (Escher et al., 2001)	3.18 (Escher et al., 2001)	<b>1.33</b>	POPC	TRANSIL	NA
591-35-5	3,5-dichlorophenol	35DCP	8.26 (Schwarzenbach et al., 2003)	<b>3.76</b> (Escher et al., 2000)	2.09 (Smejtek et al., 1996)	<b>1.67</b>	egg-PC	electrophoretic mobility measurements	25
2338-29-6	4,5,6,7-tetrachloro-2-(trifluoromethyl)-1H-benzimidazole	TTFB	5.30 (Dilger and McLaughlin, 1979)	<b>4.35</b> (Dilger and McLaughlin, 1979)	4.35 (Dilger and McLaughlin, 1979)	<b>0.00</b>	egg-PC	equilibrium dialysis	22.5

106-48-9	4-chlorophenol	4CP	9.38 (Escher and Schwarzenbach, 1996)	<b>2.96</b> (Escher et al., 2000)	2.51 (Escher et al., 2000)	<b>0.45</b>	DPPC/ DOPC	equilibrium dialysis	20
609-93-8	4-methyl-2,6-dinitrophenol	DNPC	4.06 (Schwarzenbach et al., 1988)	<b>2.34</b> (Escher et al., 2000)	2.26 (Escher et al., 2000)	<b>0.08</b>	DPPC/ DOPC	equilibrium dialysis	20
100-02-7	4-nitrophenol	4NP	7.08 (Schwarzenbach et al., 1988)	<b>2.72</b> (Escher and Schwarzenbach, 1996)	0.95 (Escher and Schwarzenbach, 1996)	<b>1.77</b>	DPPC/ DOPC	equilibrium dialysis	20
6149-03-7	4-octylbenzene-1-sulfonate	OBS			3.63 (Bittermann et al., 2014)		POPC	equilibrium dialysis	25
4097-49-8	4-tert-butyl-2,6-dinitrophenol	Dino4terb	4.11 (Schwarzenbach et al., 1988)	<b>3.81</b> (Escher et al., 2000)	3.23 (Escher et al., 2000)	<b>0.58</b>	DPPC/ DOPC	equilibrium dialysis	20
2338-25-2	5,6-dichloro-2-(trifluoromethyl)-benzimidazole	DTFB	7.30 (Cohen et al., 1977)	<b>3.05</b> (Cohen et al., 1977)	3.05 (Cohen et al., 1977)	<b>0.00</b>	egg-PC	equilibrium dialysis	NA
521-74-4	5,7-dibromo-8-hydroxyquinoline	Dibromox	2.90 (Kaiser and Escher, 2006)	<b>3.94</b> (Kaiser and Escher, 2006)	3.03 (Kaiser and Escher, 2006)	<b>0.91</b>	POPC	equilibrium dialysis	25
773-76-2	5,7-dichloro-8-hydroxyquinoline	Dichlorox	2.60 (Kaiser and Escher, 2006)	<b>3.35</b> (Kaiser and Escher, 2006)	2.47 (Kaiser and Escher, 2006)	<b>0.88</b>	POPC	equilibrium dialysis	25
16128-96-4	5-chloro-3-tert-butyl-2'-chloro-4'-nitrosalicylanilide	S-13	5.80 (Kasianowicz et al., 1987)	<b>6.44</b> (Kasianowicz et al., 1987)	5.05 (Kasianowicz et al., 1987)	<b>1.40</b>	egg-PC	equilibrium dialysis	21
130-16-5	5-chloro-8-hydroxyquinoline	Chlorox	3.71 (Kaiser and Escher, 2006)	<b>3.29</b> (Kaiser and Escher, 2006)	1.91 (Kaiser and Escher, 2006)	<b>1.38</b>	POPC	equilibrium dialysis	25
327-19-5	5-nitro-2-trifluoromethyl-	5-NB			1.81 (Bittermann et al., 2014)		POPC	equilibrium dialysis	25

	benzimidazole										
2270-20-4	5-phenylvaleric acid	5-PA	4.88 <sup>P</sup>	<b>3.06</b> (Austin et al., 1995)	1.66 (Avdeef et al., 1998)	<b>1.40</b>	DOPC	pH metric technique	NA		
148-24-3	8-hydroxy-quinoline	Oxine	4.89 (Kaiser and Escher, 2006)	<b>2.17</b> (Kaiser and Escher, 2006)	1.47 (Kaiser and Escher, 2006)	<b>0.70</b>	POPC	equilibrium dialysis	25		
<b>87848-99-5</b>	<b>acrivastine</b>	<b>Acr_a</b>	<b>2.20</b> (Plemper van Balen et al., 2001)		<b>2.60</b> (Plemper van Balen et al., 2001)		<b>egg-PC</b>	<b>equilibrium dialysis</b>	<b>25</b>		
118-92-3	anthranilic acid	AA	4.76 (Thomae et al., 2007)	<b>2.08</b> (Thomae et al., 2007)	0.13 (Thomae et al., 2007)	<b>1.95</b>	egg-PC	equilibrium dialysis	26		
555-60-2	carbonyl cyanide m-chlorophenyl-hydrazone	CCCP	5.95 (Kasianowicz et al., 1987)	<b>4.05</b> (Kasianowicz et al., 1987)	4.05 (Kasianowicz et al., 1987)	<b>0.00</b>	egg-PC	equilibrium dialysis	21		
370-86-5	carbonyl cyanide p-methoxyphenylhydrazone	FCCP	6.20 (Kasianowicz et al., 1987)	<b>4.22</b> (Kasianowicz et al., 1987)	4.22 (Kasianowicz et al., 1987)	<b>0.00</b>	egg-PC	equilibrium dialysis	21		
15307-86-5	diclofenac	Dic	3.99 (Avdeef et al., 1998)	<b>4.45</b> (Avdeef et al., 1998)	2.64 (Avdeef et al., 1998)	<b>1.81</b>	DOPC	potentiometric titration	25		
22494-42-4	diflunisal	Dif	3.00 (Pallicer and Krämer, 2012)		2.73 (Pallicer and Krämer, 2012)		egg-PC	equilibrium dialysis	25		
91-40-7	fenamic acid	Fen	3.99 <sup>P</sup>		2.28 (Bittermann et al., 2014)		POPC	equilibrium dialysis	25		
530-78-9	flufenamic acid	Flu			3.61 (Bittermann et al., 2014)		POPC	equilibrium dialysis	25		
15687-27-1	ibuprofen	Ibu	4.45 (Avdeef et al., 1998)	<b>3.80</b> (Avdeef et al., 1998)	1.81 (Avdeef et al., 1998)	<b>1.99</b>	DOPC	potentiometric titration	25		
36894-69-6	labetalol	Lab	7.35 (Pallicer and Krämer, 2012)	<b>2.73</b> (Pallicer and Krämer, 2012)	1.84 (Pallicer and Krämer, 2012)	<b>0.89</b>	egg-PC	equilibrium dialysis	25		

124-07-2	octanoic acid	Oct	4.89 <sup>P</sup>	2.91 (Inoue et al., 1988)	0.52 (Inoue et al., 1988)	2.39	DPPC	depression of the phase transition temperature	37
608-71-9	pentabromo-phenol	PBrP	4.62 <sup>P</sup>		5.02 (Smejtek et al., 1996)		egg-PC	electrophoretic mobility measurements	25
87-86-5	pentachloro-phenol	PCP	4.75 (Schellenberg et al., 1984)	5.10 (Escher et al., 2000)	4.35 (Escher et al., 2000) 4.28 (Smejtek et al., 1996)	0.79	DPPC/ DOPC egg-PC	equilibrium dialysis electrophoretic mobility measurements	20 25
771-61-9	pentafluoro-phenol	PFP	5.53 <sup>P</sup>		1.74 (Smejtek et al., 1996)		egg-PC	electrophoretic mobility measurements	25
1763-23-1	perfluorooctane-1-sulfonic acid	PFOS	0.14 <sup>P</sup>		3.15 (Lehmler et al., 2006)		DPPC	depression of the phase transition temperature	37
335-67-1	perfluorooctanoic acid	PFOA	2.80 <sup>P</sup>		2.34 (Inoue et al., 1988)		DPPC	depression of the phase transition temperature	37
69-72-7	salicylic acid	SA	2.75 (Thomae et al., 2005) 3.00 (Ottiger and Wunderli-Allenspach, 1997)	2.66 (Thomae et al., 2005) 2.59 (Thomae et al., 2007)	1.03 (Thomae et al., 2005) 0.85 (Thomae et al., 2007)	1.61	DPPC egg-PC	equilibrium dialysis equilibrium dialysis	37 25

				<b>2.50</b> (Ottiger and Wunderli-Allenspach, 1997)	1.04 (Ottiger and Wunderli-Allenspach, 1997)		egg-PC	equilibrium dialysis	37
1198-55-6	tetrachloro-catechol	TeCC	5.97 (Schweigert et al., 2001)	<b>4.41</b> (Schweigert et al., 2001)	2.63 (Schweigert et al., 2001)	<b>1.78</b>	DOPC	potentiometric titration	25
4358-26-3	tetraphenylborate	TPB			5.05 (Flewelling and Hubbell, 1986)		egg-PC	electron paramagnetic resonance	25
					5.35 (Flewelling and Hubbell, 1986)		egg-PC	electron paramagnetic resonance	25
81-81-2	warfarin	Warf	4.90 (Ottiger and Wunderli-Allenspach, 1997)	<b>3.39</b> (Ottiger and Wunderli-Allenspach, 1997)	1.40 (Ottiger and Wunderli-Allenspach, 1997)	<b>1.99</b>	egg-PC	equilibrium dialysis	37

### 1.3 Data collection zwitterions

SI-2, Table 3. Data collection for anions based on a previously published data collection (Bittermann et al., 2014), with all new values marked in bold font. Multiple values for  $pK_a$ ,  $\log K_{lipw}$  (neutral) and  $\log K_{lipw}$  (ion) are marked in grey; in these cases the arithmetic mean was used for the calculation of  $\Delta mw$ . 'P' stands for values taken from the PhysProp-Database (<http://esc.syrres.com/fatepointer/search.asp>), egg-PC for egg-phosphatidylcholine, DOPC for dioleoylphosphatidylcholine, DPPC for 1,2-Dipalmitoyl-sn-glycero-3-phosphocholine and POPC for 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine.

CAS	compoundname	abbreviation	$pK_a$	$\log K_{lipw}$ (zwitterion) [L/kg]	lipid	method	T [°C]
83881-51-0	ceterizine	Cet_zw	<b>2.93/8.00</b> (Plemper van Balen et al., 2001)	<b>2.30</b> (Plemper van Balen et al., 2001)	egg-PC	equilibrium dialysis	25
87848-99-5	acrivastine	Acr_zw	<b>2.20/9.55</b> (Plemper van Balen et al., 2001)	<b>1.50</b> (Plemper van Balen et al., 2001)	egg-PC	equilibrium dialysis	25



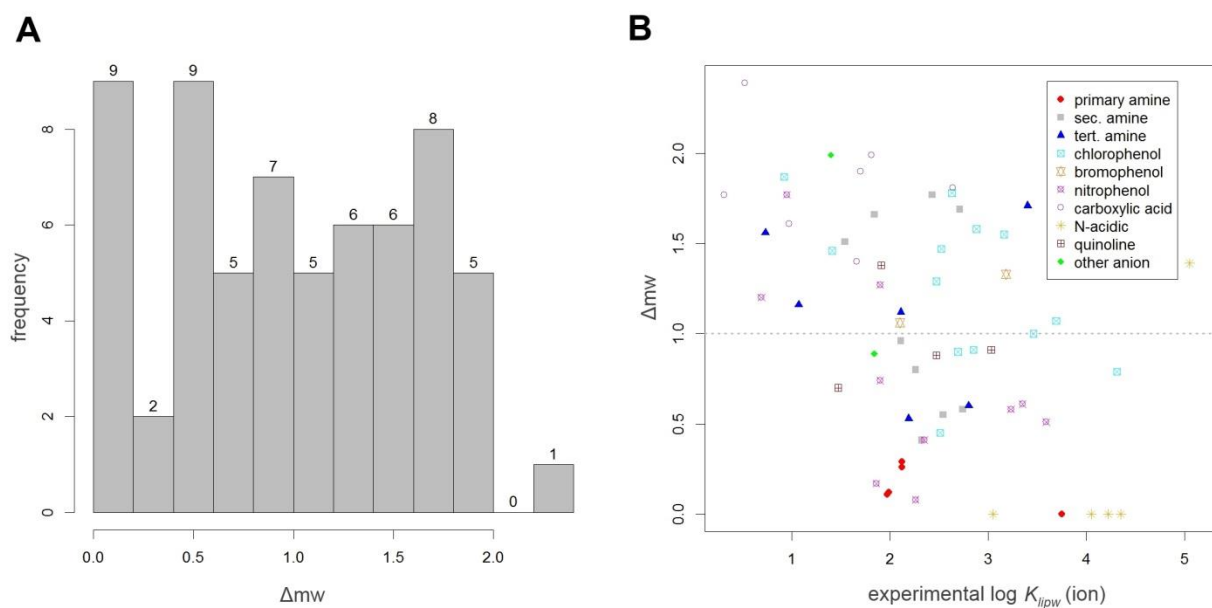
## 1.4 Omitted data

The experimental  $K_{lipw}$  data published by (Inoue et al., 1986) comprise a homologous series of four linear quaternary amines and three linear sulfates. These data have been omitted in our data compilation because they contradict unpublished LCMS measurements conducted by Droge et al. Droge used the 10 cm IAM.PC.DD2 column from Regis Technologies, with a buffer of 10 mM ammonium acetate at pH 5. Using a flow rate of 1 mL/min, and creating a ~95% split, the remaining 5% was injected in the ESI-LC-MS/MS (AB/Sciex (Applied Biosystems) 3000), and scanned for the mass of the cationic species. These HPLC conditions render similar ( $< \sim 0.2$  log units) sorption affinities of organic cations to the phospholipid material on the HPLC column as the sorption affinities determined with phospholipid liposomes (yet unpublished results).

## 2 Log $K_{ow}$ based prediction

### 2.1 Differences in $K_{lipw}$ between neutral and corresponding ionic species ( $\Delta m_w$ )

The experimental  $\Delta m_w$  values for 43 anionic and 20 cationic compounds (see Table 1 and 2, SI-2) scatter from 0 to 2.39 log units as shown in the histogram in SI-1, Fig. 1A. SI-1, Fig. 1B shows the corresponding experimental  $\log K_{lipw}$  values for these 63  $\Delta m_w$  values according to the different subclasses. All  $\Delta m_w$  values of the different species are summarized in SI-2, Table 4. Despite the limitation of the data, SI-1, Fig. 1B clearly underlines that  $\Delta m_w$  scatters widely over the whole range of  $\log K_{lipw}$  values. Although it seems to be safe to state that e.g. primary amines tend to have lower  $\Delta m_w$  values than carboxylic acids, there is no clear trend for most of the chemical classes listed.



SI-2, Figure 1. Histogram of 63 experimental  $\Delta mw$  values (A) and plot of the experimental  $\Delta mw$  values against their experimental  $\log K_{lipw}$  values (B). The dotted grey line indicates the generic  $\Delta mw$  value of 1 log unit.

SI-2, Table 4. Summary of  $\Delta mw$  values  $\pm$  standard deviation of the different classes as depicted in the SI-2, Fig. 1B and the Summary, section 1.4.2.1, Fig. 7B.

class	Charge	number	$\Delta mw$	SD
primary amine	+	5	0.16	0.12
secondary amine	+	9	1.10	0.55
tertiary amine	+	6	1.11	0.48
chlorophenol	-	13	1.24	0.42
bromophenol	-	2	1.20	0.19
nitrophenol	-	10	0.73	0.53
carboxylic acid	-	7	1.84	0.31
N-acidic	-	5	0.28	0.62
quinoline	-	4	0.97	0.29
other anion	-	2	1.44	0.78

It has already been discussed that  $\Delta mw$  values for phenols are closer to 1 log unit (Escher et al., 2000), while  $\Delta mw$  values of carboxylic acids are usually higher (Escher and Sigg, 2004) (see also SI-2, Fig. 1). It has been hypothesized that this is due to the charge delocalization being more effective in the case of phenols than in the case of carboxylic acids (Escher and

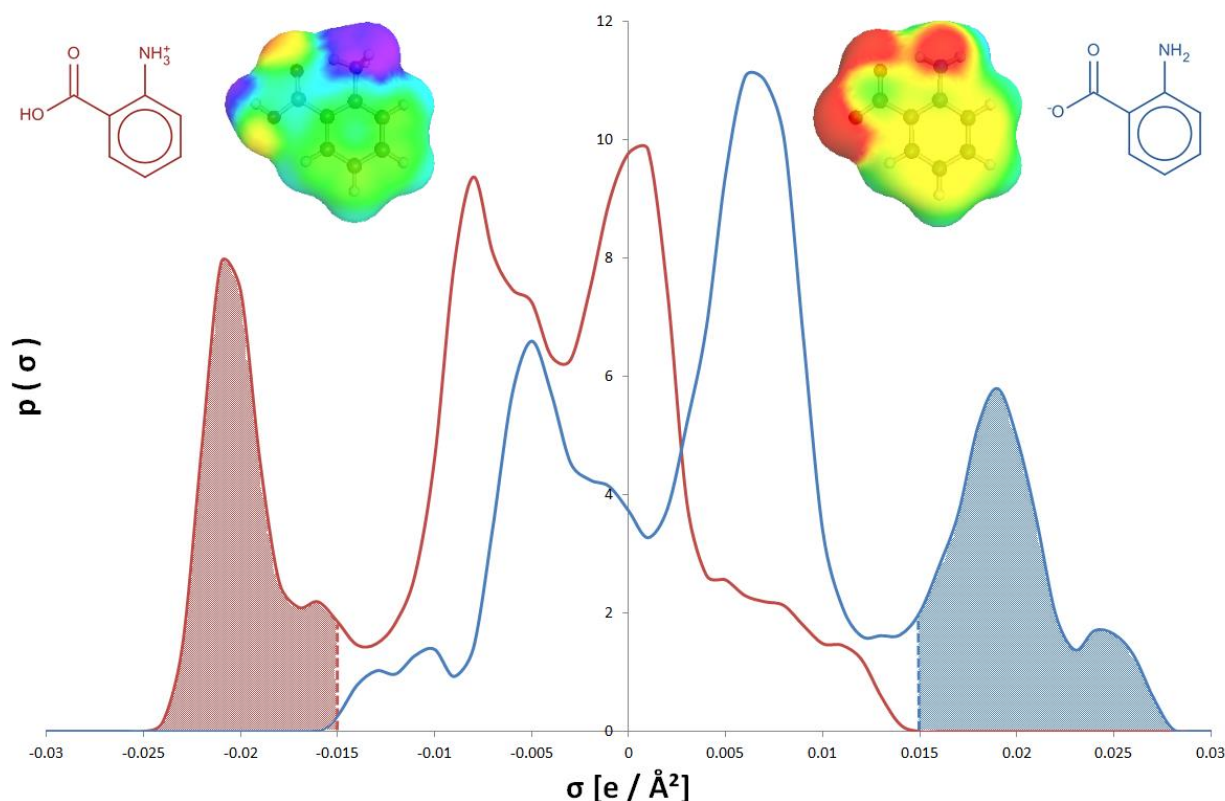
Sigg, 2004). For cationic compounds on the other hand an increase in  $\Delta mw$  values has been observed from primary to secondary to tertiary amines (Neuwoehner et al., 2009). It has been reasoned that the positively charged  $\text{NH}_2$ -group of primary amines might interact more favorably with the polar parts of the membrane than secondary or tertiary amines (Neuwoehner et al., 2009). This seems like a contradictory finding, because a higher charge density increases  $\Delta mw$  values for the anionic phenol/carboxylic acid pair, while a higher charge density decreases the  $\Delta mw$  values in the case of cationic primary/secondary/tertiary amines. In order to investigate whether this is a generalizable feature for anionic and cationic compounds independent of the chemical classes, we correlated the charge densities of the ions with their respective  $\Delta mw$  values (see SI-2, Fig. 2 and SI-2, Fig.3). But before discussing SI-2, Fig. 3 it has to be explained how the charge densities of the ions have been derived:

From the SMILES code of the ions, a 3D structure has been generated with CORINA (Sadowski et al., 1994) (available from Molecular Networks GmbH, Erlangen, Germany; <http://www.molecular-networks.com>). Subsequently, full energy minimization and conformer generation of TZVP (Becke, 1988; Eichkorn et al., 1995; Perdew, 1986; Schäfer et al., 1994) cosmo files are calculated with COSMOconfX13 (version 3.0, COSMOlogic) templates (Vainio and Johnson, 2007). This quantum chemical calculation is based on Turbomole version 6.5<sup>5</sup> and yields at least one cosmo file for every molecule (with up to 10 cosmo files for 10 conformers being possible). Within every cosmo file, the charge density profile of the respective molecule is stored; the so called  $\sigma$ -profile (detailed explanations can be found in the literature: see the publications of Klamt (Klamt, 2005, 1995) for a thorough mathematical derivation and the reviews (Eckert and Klamt, 2002; Klamt et al., 2010) for practical hands-on examples). As an example, SI-2, Fig. 2 shows the  $\sigma$ -profiles of the anthranilic acid cation (in red) and the anthranilic acid anion (in blue). It is important to note, that the charge is given with respect to the perfect conductor; i.e. negative partial charges on molecules are represented by positive screening charge densities and vice versa (Eckert and Klamt, 2002; Klamt, 1995). To determine the influence of the charge densities on  $\Delta mw$  we chose the energetically most favorable conformer of every compound and integrated the charge densities above the threshold of  $0.015 \text{ e}/\text{Å}^2$  in the case of anions and below  $-0.015 \text{ e}/\text{Å}^2$  in the case of cations (shown as dotted blue and red area under the curve in SI-2, Fig. 2). This

---

<sup>5</sup> TURBOMOLE V6.5 2013, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>

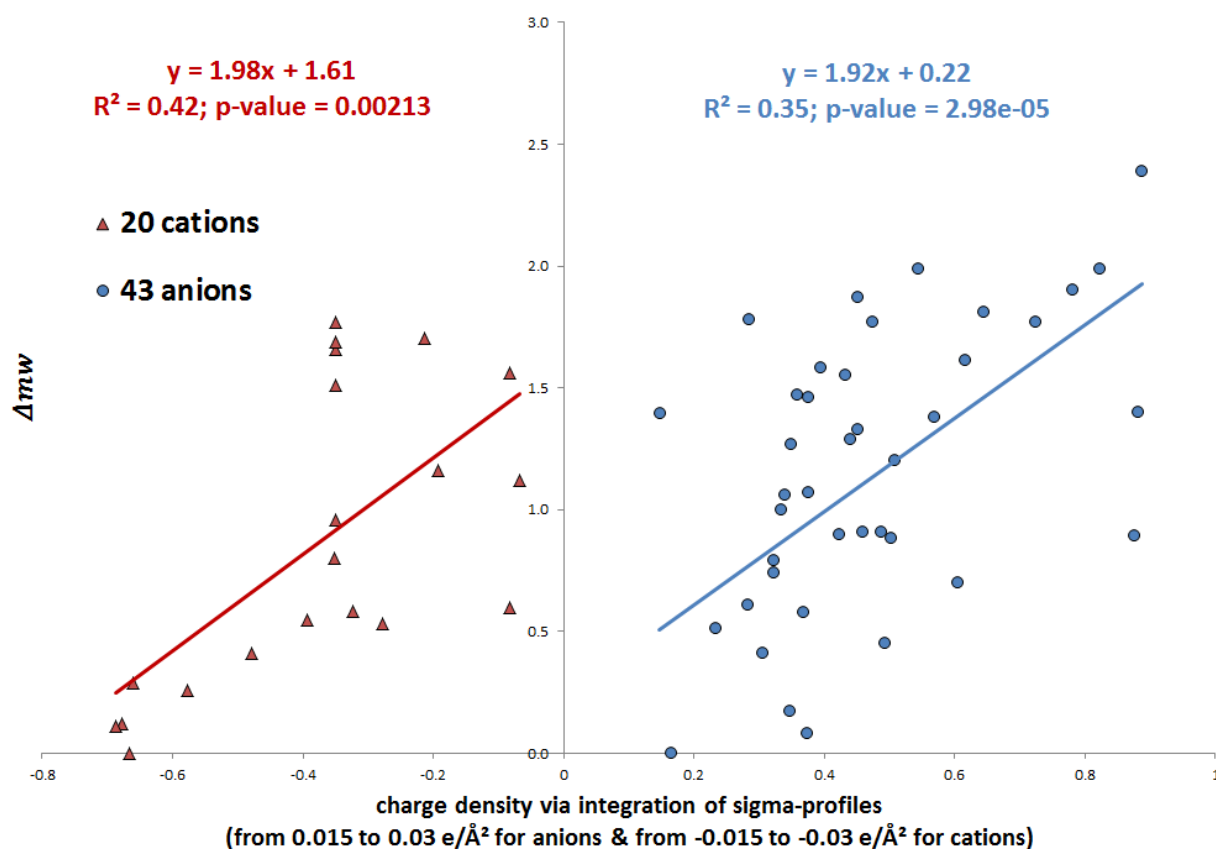
threshold determined by visual examination is somewhat arbitrary, but gave the best correlations with the  $\Delta mw$  values as depicted in SI-2, Fig. 3.



SI-2, Figure 2.  $\sigma$ -profiles of anthranilic acid cation (red) and anthranilic acid anion (blue). The area under the curve has been integrated from  $-0.015$  to  $-0.003$   $e/\text{\AA}^2$  (indicated by the red dotted area) and  $0.015$  to  $0.0031$   $e/\text{\AA}^2$  (indicated by the blue dotted area) for cations and anions, respectively. Representations of the  $\sigma$ -profiles in combination with the 3-D structures of the molecules (as stored in the cosmo-files) are depicted above with the color coding according to the charge densities.

SI-2, Fig. 3 shows the correlation between  $\Delta mw$  values of 20 cationic ( $R^2 = 0.42$ ) and 43 ( $R^2 = 0.35$ ) anionic compounds against their integrated charge densities. Although the proportion of variance of the  $\Delta mw$  values explained by the charge densities is not striking (as shown by the low  $R^2$ ), the p-values of both coefficients are significant. A least squares regression gives positive values for the slope of both cationic and anionic compounds. Following the cosmo notation this indicates a contrary influence of the charge densities of cationic and anionic compounds on  $\Delta mw$ : for the cations higher charge densities are correlated with lower  $\Delta mw$  values, while for the anions higher charge densities are correlated with higher  $\Delta mw$  values. This is in good agreement with the previous findings for the anionic phenol/carboxylic acid pair (Escher et al., 2000; Escher and Sigg, 2004) and the cationic primary/secondary/tertiary amines (Neuwoehner et al., 2009). However, the scatter within these correlations is so high,

that we think the charge densities should be seen rather as a qualitative than a quantitative indicator of  $\Delta mw$  values.

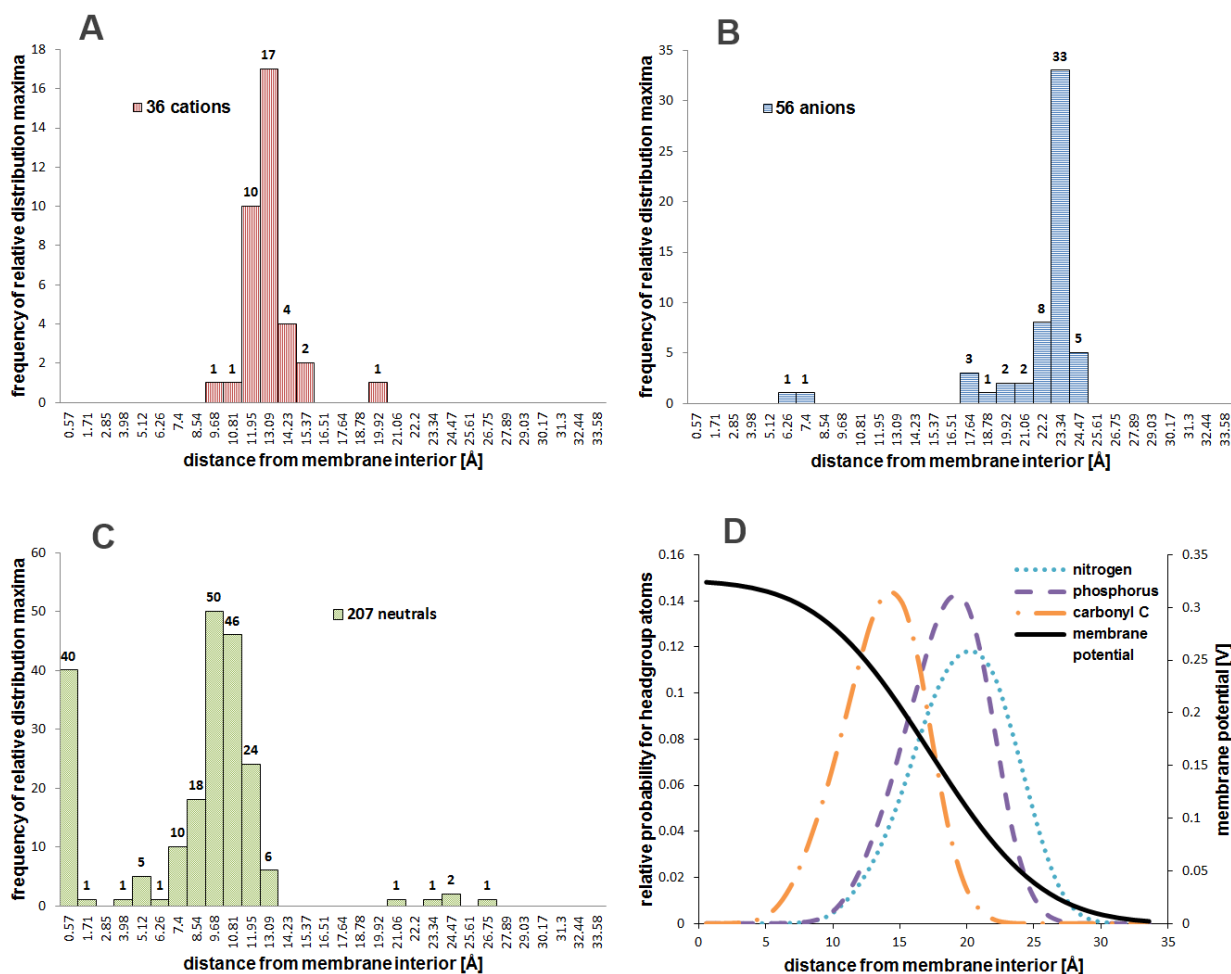


SI-2, Fig. 3. Experimental  $\Delta mw$  values of 20 cations and 43 anions against their integrated sigma profile from 0.015 to 0.031  $e/\text{\AA}^2$  and -0.02 to -0.031  $e/\text{\AA}^2$ , respectively.

## 2.2 Why does the $\log K_{ow}$ approach predict $K_{lipw}$ for anions better than for cations

Fig. 6 in the summary, section 1.4 shows that the  $\log K_{ow}$  approach predicts the  $\log K_{lipw}$  values of anions (RMSE = 0.79,  $R^2 = 0.61$ ,  $n = 56$ ) better than the  $\log K_{lipw}$  of cations (RMSE = 1.14,  $R^2 = 0.23$ ,  $n = 36$ ), while the  $\log K_{lipw}$  of neutral compounds are predicted most accurately (RMSE = 0.52,  $R^2 = 0.93$ ,  $n = 207$ ). We investigated this finding by examining the distribution of the molecules in the membrane, according to the relative distribution profiles calculated for every molecule with COSMOmic. To this end we plotted the frequency of the relative distribution maxima (i.e. the membrane layer with the maximum probability to find the center of mass of a given molecule) in SI-2, Fig. 4: neutral compounds (C) seem to sorb mainly either to the membrane center or to a distance of 9.7  $\text{\AA}$  from the membrane center (where the probability of carbonyl carbons levels off and the alkane-like interior of the membrane begins (D)). Most of the cations in the dataset sorb deeper in the membrane (with a

peak in the frequency distribution at 13.1 Å (A)) than the anions (with a peak in the frequency distribution at 23.3 Å (B)).

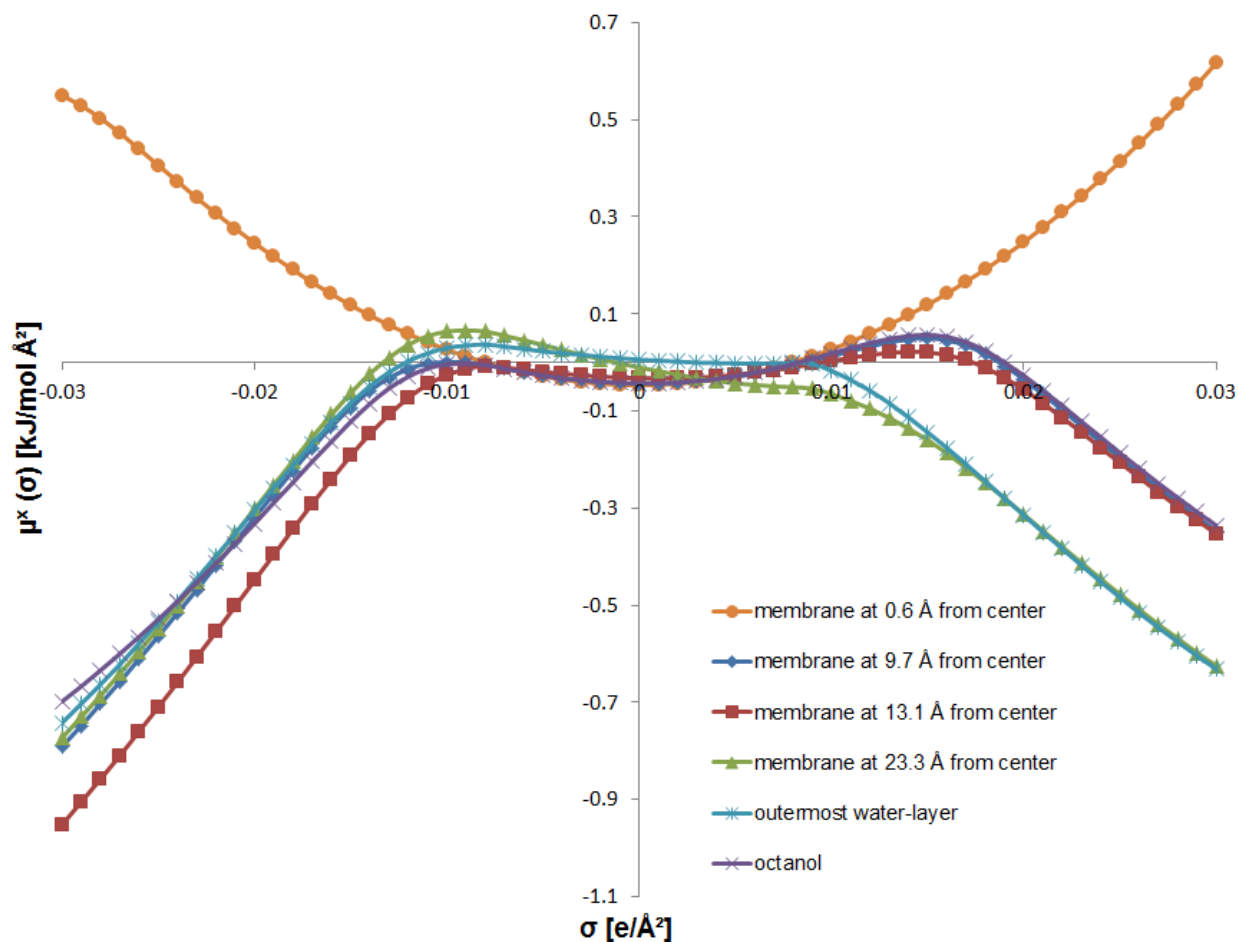


SI-2, Figure 4. Histograms of the relative distribution maxima of 36 cationic (A), 56 anionic (B) and 207 neutral compounds (C) as calculated with COSMOmic; D shows the corresponding relative probability of the membrane headgroup atoms and the shape of the membrane potential.

According to the relative distribution maxima in the histograms of the differently charged compounds in SI-2, Fig. 4, SI-2, Fig. 5 shows the corresponding  $\sigma$ -potentials of the respective layers in the membrane and of octanol as a comparison. The  $\sigma$ -potentials can be seen as a characteristic function of a solvent that describes possible interactions with a solutes' surface area of polarity  $\sigma$  (Eckert and Klamt, 2002; Klamt, 2005, 1995; Klamt et al., 2010) (in the case of the different membrane layers, the  $\sigma$ -potentials refer to theoretical homogeneous solvents with the specific properties of the respective layer) (Bittermann et al., 2014; Klamt et al., 2008). The  $\sigma$ -potential of the membrane center (which is the preferred sorption environment for around 20% of neutral compounds; SI-2, Fig. 4C) has an almost parabolic shape, which is equivalent to purely dielectric behavior and indicates the lack of hydrogen bonding capacities (Eckert and Klamt, 2002) (i.e., not surprising, the membrane interior has

properties equivalent to hexane). The dominating sorption depth for neutral compounds at around 9.7 Å from the membrane center, however, shows properties that resemble very closely the properties of octanol – which is a satisfying explanation why most of the  $K_{lipw}$  values of neutral compounds agree well with their respective  $K_{ow}$  values. For the ions, the comparison to the  $\sigma$ -potential of octanol is not as clear: while the potential at around 13.1 Å from the membrane center (being the preferred sorption depth of cations) fits well with the corresponding positive  $\sigma$ -values of octanol, the potential at around 23.3 Å from the membrane center (being the preferred sorption depth of anions) fits well with the corresponding negative  $\sigma$ -values of octanol. In other words, the preferred sorption depth of cations has H-bond donor properties comparable to those of octanol, while the preferred sorption depth of anions has H-bond acceptor properties comparable to those of octanol. This cannot satisfyingly explain the different modeling performance of the  $\log K_{ow}$  approach with respect to anions and cations. Also the neighboring layers will be of importance and the specific orientation of a given ion in the membrane.

Stepping back from the comparison to the  $\sigma$ -potential of octanol, the  $\sigma$ -potentials of the preferred layers of anionic and cationic compounds intuitively do make sense: Most cations sorb mainly to the membrane depth with the most pronounced H-bond acceptor properties, i.e. the  $\sigma$ -potential has the most negative  $\mu$ -values for negative  $\sigma$ -values (red curve). This is because the cations are good H-bond donors. The anions on the other hand are good H-bond acceptors. Therefore they mainly sorb to the layer with the most negative  $\mu$ -values for positive  $\sigma$ -values (green curve), i.e. the layer with the best H-bond donor properties. Interestingly, the preferred sorption depth of anions has very similar properties as the outermost water layer.

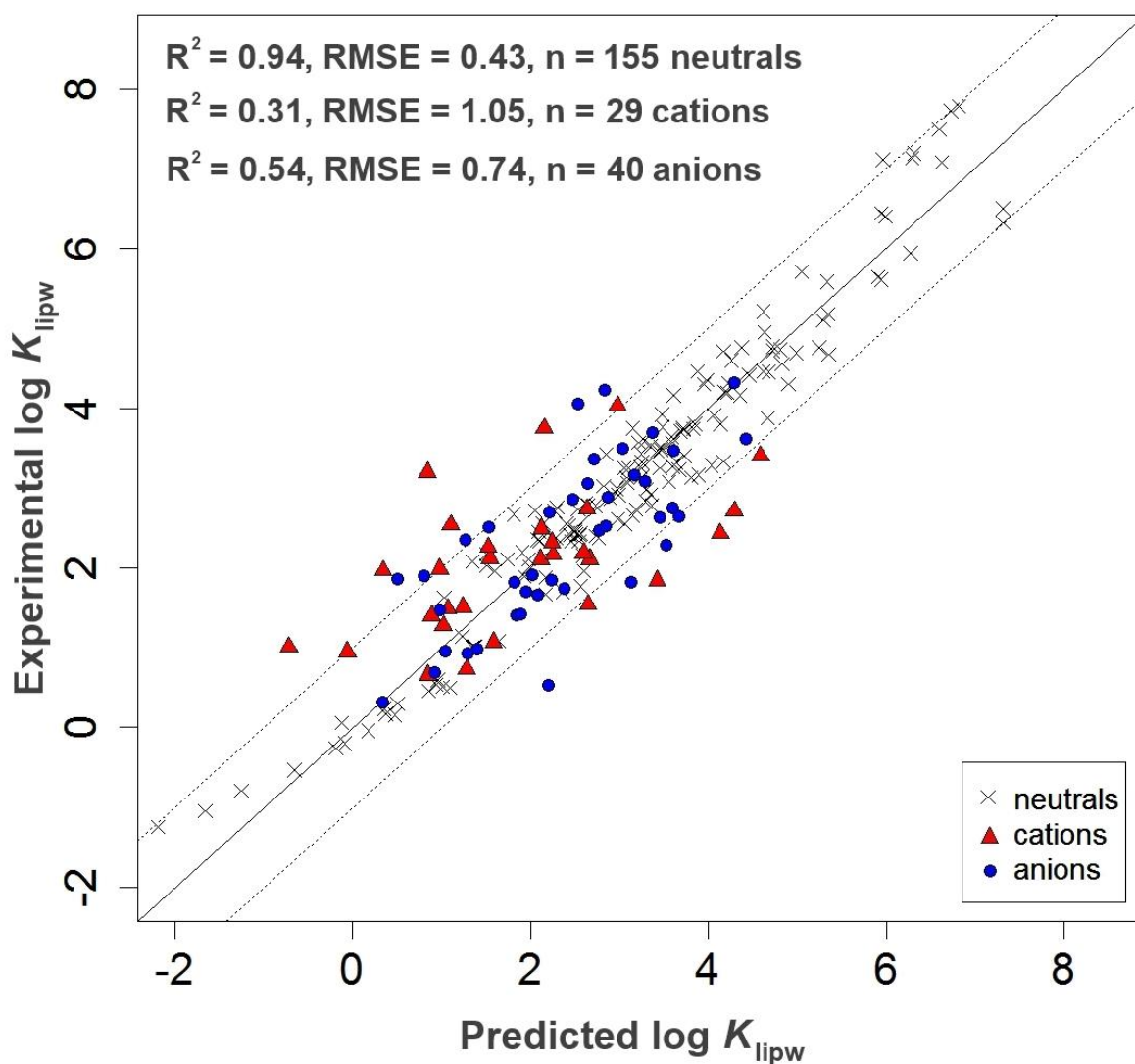


SI-2, Figure 5.  $\sigma$ -potentials of different membrane layers used in the COSMOmic calculation as well as  $\sigma$ -potential of octanol as comparison (at  $T = 298.15$  K).

### 2.3 Ruling out artefacts from KowWIN estimation errors

The predictability of KowWIN is likely to be poor for compounds with functional groups that were not included with sufficient entries (or not included at all) in the parametrization dataset. In order to test whether the limitations of the empirical correlation approach with  $\log K_{ow}$  are due to the limited applicability domain of KowWIN, we reevaluated the model using only experimental  $\log K_{ow}$  values (i.e. 155 out of 207 neutral compounds, 29 out of 36 cationic compounds and 40 out of 56 anionic compounds). SI-2, Fig. 6 below shows that the model based exclusively on experimental  $\log K_{ow}$  values does not yield substantially better results than the model using also  $\log K_{ow}$  values estimated with KowWIN, as presented in the summary above (compare with SI-2, Fig. 1 in the summary).





SI-2, Figure 6. Comparison between the experimental log  $K_{lipw}$  values of 155 neutral, 29 cationic and 40 anionic compounds and the predicted values according to the empirical correlation approach with log  $K_{ow}$  using only experimental log  $K_{ow}$  values, simple regression and  $\Delta mw$  as outlined in the summary, section 1.4. Deviations of 1 log unit from the straight identity line are shown as dotted lines.

## 2.4 Predicted data

### 2.4.1 Cations

SI-2, Table 5. Log  $K_{ow}$  values derived with EpiSuite<sup>6</sup> (estimated and experimental) and resulting log  $K_{lipw}$  values of the corresponding ions according to the Endo QSAR ((Endo et al., 2011) Eq. 1) minus 1 log unit. For the log  $K_{lipw}$  calculation experimental log  $K_{ow}$  values were preferred over estimated ones.

CAS	abbreviation	SMILES of corresponding neutral form (as used in EpiSuite)	log $K_{ow}$	log $K_{ow}$	log $K_{lipw}$
			neutral (EpiSuite estimation)	neutral (EpiSuite experimental)	ion (Endo QSAR – 1 log unit)
88-05-1	246TMA	<chem>CC1=CC(C)=C(N)C(C)=C1</chem>	2.72	-	1.87
95-64-7	34DMA	<chem>CC1=C(C)C=C(N)C=C1</chem>	2.17	1.84	0.98
13214-66-9	4-PhenButA	<chem>NCCCCC1=CC=CC=C1</chem>	2.54	2.4	1.54
118-92-3	AA_cation	<chem>NC1=C(C=CC=C1)C(O)=O</chem>	1.36	1.21	0.34
37517-30-9	ABL	<chem>O=C(Nc1ccc(OCC(O)CNC(C)C)c(c1)C(=O)C)CCC</chem>	1.19	1.71	0.85
88150-42-9	Amlodip	<chem>CCOC(=O)C1=C(COCCN)NC(C)=C(C1C1=CC=CC=C1Cl)C(=O)OC</chem>	2.07	3	2.15
13655-52-2	APL	<chem>O(c1ccccc1C\C=C)CC(O)CNC(C)C</chem>	2.81	3.1	2.25
29122-68-7	Aten	<chem>CC(C)NCC(O)COC1=CC=C(CC(N)=O)C=C1</chem>	-0.03	0.16	-0.72
23284-25-5	BPL	<chem>CC1=CC(=C(C=C1)Cl)OCC(CNC(C)C)CO</chem>	3.07	2.97	2.12

<sup>6</sup> EPI Suite Exposure Assessment Tools and Models.

<http://www.epa.gov/opptintr/exposure/pubs/episuite.htm>.

83881-51-0	Cet_c	<chem>Clc1ccc(cc1)C(c2ccccc2)N3CCN(CC3)CCOCC(=O)O</chem>	-0.61	1.7	0.84
50-53-3	CLP	<chem>CN(C)CCCN1c2ccccc2Sc2ccc(Cl)cc12</chem>	5.2	5.41	4.58
54910-89-3	Fluox	<chem>CNCCC(OC1=CC=C(C=C1)C(F)(F)F)C1=CC=CC=C1</chem>	4.65	3.82	2.98
68-88-2	Hyd	<chem>OCCOCCN1CCN(CC1)C(c1ccccc1)c1ccc(Cl)cc1</chem>	2.36	-	1.50
312753-06-3	Indac	<chem>CCC1=CC2=C(CC(C2)NCC(O)C2=CC=C(O)C3=C2C=CC(=O)N3)C=C1CC</chem>	3.3	-	2.45
36894-69-6	Lab_c	<chem>CC(CCCc1ccccc1)NCC(O)c1ccc(O)c(c1)C(N)=O</chem>	2.41	3.09	2.24
137-58-6	Lido	<chem>CCN(CC)CC(=O)NC1=C(C)C=CC=C1C</chem>	1.66	2.44	1.58
16183-21-4	MBButA	<chem>CCCCNCC1=CC=C(C=C1)C</chem>	3.56	3.49	2.64
39099-13-3	MBEthA	<chem>CCNCC1=CC=C(C)C=C1</chem>	2.57	2.38	1.52
215177-24-5	MBHepA	<chem>CCCCCCCNCc1ccc(C)cc1</chem>	5.03	5.12	4.29
215177-23-4	MBHexA	<chem>CCCCCCNCc1ccc(C)cc1</chem>	4.54	4.96	4.13
699-04-7	MBMetA	<chem>CNCC1=CC=C(C)C=C1</chem>	2.08	1.96	1.10
170303-38-5	MBPentA	<chem>CCCCCNCc1ccc(C)cc1</chem>	4.05	4.26	3.42
39190-96-0	MBPropA	<chem>CCCNCC1=CC=C(C)C=C1</chem>	3.06	2.96	2.11
51384-51-1	Metro	<chem>COCCC1=CC=C(OCC(O)CNC(C)C)C=C1</chem>	1.69	1.88	1.02
20574-50-9	Mor	<chem>CN1CCCN=C1\C=C\C1=C(C)C=CS1</chem>	3.69	-	2.85
42200-33-9	NDL	<chem>OC(CNC(C)(C)C)COc1ccccc2c1C[C@H](O)[C@H](O)C2</chem>	1.17	0.81	-0.06
83891-03-6	Norfluox	<chem>NCCC(OC1=CC=C(C=C1)C(F)(F)F)C1=CC=CC=C1</chem>	4.18	-	3.34
6452-71-7	OPL	<chem>O(c1ccccc1OC\C=C)CC(O)CNC(C)C</chem>	1.83	2.1	1.24
13523-86-9	PDL	<chem>CC(C)NCC(O)COc2ccccc1nccc12</chem>	1.48	1.75	0.89
59-46-1	Proc	<chem>CCN(CC)CCOC(=O)C1=CC=C(N)C=C1</chem>	1.99	2.14	1.28
525-66-6	Prop	<chem>CC(C)NCC(O)COC1=CC=CC2=C1C=CC=C2</chem>	2.6	3.48	2.63
130-95-0	Quinine	<chem>COC1=CC2=C(C=CN=C2C=C1)C(O)C1CC2CCN1CC2C=C</chem>	3.29	3.44	2.59

89365-50-4	Salmet	<chem>OCC1=C(O)C=CC(=C1)C(O)CNCCCCCOCCCCC1=CC=CC=C1</chem>	4.15	-	3.31
94-24-6	Tetrac	<chem>CCCCNC1=CC=C(C=C1)C(=O)OCCN(C)C</chem>	3.02	3.51	2.67
2933-94-0	TPL	<chem>CC1=CC(=CC=C1)OCC(CNC(C)C)O</chem>	1.97	1.93	1.07
18198-39-5	TPP	<chem>C1=CC=C(C=C1)P(C1=CC=CC=C1)(C1=CC=CC=C1)C1=CC=CC=C1</chem>	5.28	-	4.45

## 2.4.2 Anions

SI-2, Table 6. Log  $K_{ow}$  values derived with EpiSuite<sup>7</sup> (estimated and experimental) and resulting log  $K_{lipw}$  values of the corresponding ions according to the Endo QSAR ((Endo et al., 2011) Eq. 1) minus 1 log unit. For the log  $K_{lipw}$  calculation experimental log  $K_{ow}$  values were preferred over estimated ones.

CAS	Abbr- viation	SMILES of corresponding neutral form (as used in EpiSuite)	log $K_{ow}$	log $K_{ow}$	log $K_{lipw}$
			neutral (EpiSuite estimation)	neutral (EpiSuite exp.)	ion (Endo QSAR – 1 log unit)
94-75-7	2,4-D	<chem>Clc1cc(Cl)ccc1OCC(=O)O</chem>	2.62	2.81	1.96
4901-51-3	2345TeCP	<chem>OC1=C(Cl)C(Cl)=C(Cl)C(Cl)=C1</chem>	4.09	4.21	3.37
58-90-2	2346TeCP	<chem>ClC1=C(Cl)C(=C(Cl)C=C1Cl)O</chem>	4.09	4.45	3.61
935-95-5	2356TeCP	<chem>Clc1c(O)c(Cl)c(Cl)cc1Cl</chem>	4.09	3.88	3.04
95-95-4	245TriCP	<chem>ClC1=C(Cl)C=C(O)C(=C1)Cl</chem>	3.45	3.72	2.88
89365-49-1	246TriBP	<chem>Oc1c(Br)cc(Br)cc1Br</chem>	4.18	4.13	3.29
88-06-2	246TriCP	<chem>OC1=C(Cl)C=C(Cl)C=C1Cl</chem>	3.45	3.69	2.85

<sup>7</sup> EPI Suite Exposure Assessment Tools and Models.

<http://www.epa.gov/opptintr/exposure/pubs/episuite.htm>.

120-83-2	24DCP	<chem>OC1=C(Cl)C=C(C=C1)Cl</chem>	2.8	3.06	2.21
51-28-5	24DNP	<chem>[O-][N+](=O)C1=CC(=CC=C1O)[N+](=O)[O-]</chem>	1.73	1.67	0.81
87-65-0	26DCP	<chem>OC1=C(Cl)C=CC=C1Cl</chem>	2.8	2.75	1.90
573-56-8	26DNP	<chem>[O-][N+](=O)C1=C(O)C(=CC=C1)[N+](=O)[O-]</chem>	1.73	1.37	0.50
95-57-8	2CP	<chem>OC1=C(Cl)C=CC=C1</chem>	2.16	2.15	1.29
88-75-5	2NP	<chem>[O-][N+](=O)C1=C(O)C=CC=C1</chem>	1.91	1.79	0.93
609-19-8	345TriCP	<chem>C1=C(C=C(C(=C1Cl)Cl)Cl)O</chem>	3.45	4.01	3.17
95-77-2	34DCP	<chem>ClC1=C(Cl)C=CC(=C1)O</chem>	2.8	3.33	2.48
577-71-9	34DNP	<chem>OC1=CC([N+](=O)[O-])=C([N+](=O)[O-])C=C1</chem>	1.73	-	0.87
13979-81-2	35DBC	<chem>OC1=CC(Br)=C(C)C(Br)=C1</chem>	3.84	-	3.00
591-35-5	35DCP	<chem>ClC1=CC(=CC(=C1)O)Cl</chem>	2.8	3.62	2.78
106-48-9	4CP	<chem>OC1=CC=C(Cl)C=C1</chem>	2.16	2.39	1.53
100-02-7	4NP	<chem>[O-][N+](=O)C1=CC=C(O)C=C1</chem>	1.91	1.91	1.05
327-19-5	5-NB	<chem>[O-][N+](=O)C1=CC=C2NC(=NC2=C1)C(F)(F)F</chem>	2.02	2.68	1.83
2270-20-4	5-PA	<chem>C1(=CC=CC=C1)CCCCC(=O)O</chem>	3.27	2.94	2.09
118-92-3	AA_anion	<chem>c1ccc(c(c1)C(=O)O)N</chem>	1.36	1.21	0.34
87848-99-5	Acr_a	<chem>O=C(O)\C=C\c3nc(\C(=C\CN1CCCC1)c2ccc(cc2)C)ccc3</chem>	2.08	-	1.22
1689-84-5	Bromox	<chem>OC1=C(Br)C=C(C#N)C=C1(Br)</chem>	3.39	-	2.54
555-60-2	CCCP	<chem>N#C\C(=N\NC1=CC(=CC=C1)Cl)C#N</chem>	3.15	3.38	2.53
130-16-5	Chlorox	<chem>Clc1ccc(O)c2ncccc12</chem>	2.31	2.88	2.03
521-74-4	Dibromox	<chem>Brc1c(O)c2ncccc2c(Br)c1</chem>	3.44	-	2.59
15307-86-5	Dic	<chem>ClC1=C(NC2=C(CC(=O)O)C=CC=C2)C(Cl)=CC=C1</chem>	4.02	4.51	3.68
773-76-2	Dichlorox	<chem>Clc1c(O)c2ncccc2c(Cl)c1</chem>	2.95	-	2.10

22494-42-4	Dif	<chem>O=C(O)c1cc(ccc1O)c2ccc(F)cc2F</chem>	4.41	4.44	3.60
1420-07-1	Dino2terb	<chem>CC(C)(C)C1=C(O)C(=CC(=C1)[N+](=O)[O-])[N+](=O)[O-]</chem>	3.64	-	2.80
4097-49-8	Dino4terb	<chem>[O-][N+](=O)C1=C(O)C(=CC(=C1)C(C)(C)C)[N+](=O)[O-]</chem>	3.64	-	2.80
88-85-7	Dinoseb	<chem>CC(CC)C1=C(O)C(=CC(=C1)[N+](=O)[O-])[N+](=O)[O-]</chem>	3.67	3.56	2.72
534-52-1	DNOC	<chem>[O-][N+](=O)C1=CC(=CC(=C1O)C)[N+](=O)[O-]</chem>	2.27	2.13	1.27
609-93-8	DNPC	<chem>[O-][N+](=O)C1=C(O)C(=CC(=C1)C)[N+](=O)[O-]</chem>	2.27	-	1.41
2338-25-2	DTFB	<chem>C1=C2C(=CC(=C1Cl)Cl)N=C([NH]2)C(F)(F)F</chem>	3.49	3.49	2.64
370-86-5	FCCP	<chem>N#CC(C#N)=NNC1=CC=C(OC(F)(F)F)C=C1</chem>	3.55	3.68	2.84
91-40-7	Fen	<chem>OC(=O)C1=C(NC2=CC=CC=C2)C=CC=C1</chem>	4.18	4.36	3.52
530-78-9	Flu	<chem>OC(=O)C1=CC=CC=C1NC1=CC(=CC=C1)C(F)(F)F</chem>	5.15	5.25	4.42
15687-27-1	Ibu	<chem>CC(C(=O)O)C1=CC=C(CC(C)C)C=C1</chem>	3.79	3.97	3.13
36894-69-6	Lab_a	<chem>O=C(c1cc(ccc1O)C(O)CNC(C)CCc2ccccc2)N</chem>	2.41	3.09	2.24
6149-03-7	OBS	<chem>CCCCCCCCC1=CC=C(C=C1)S(O)(=O)=O</chem>	2.82	-	1.97
124-07-2	Oct	<chem>CCCCCCCC(=O)O</chem>	3.03	3.05	2.20
148-24-3	Oxine	<chem>c1cc2ccnc2c(c1)O</chem>	1.66	1.85	0.99
608-71-9	PBrP	<chem>BrC1c(O)c(Br)c(Br)c(Br)c1Br</chem>	5.96	-	5.14
87-86-5	PCP	<chem>ClC1=C(Cl)C(=C(O)C(=C1Cl)Cl)Cl</chem>	4.74	5.12	4.29
335-67-1	PFOA	<chem>FC(F)(C(F)(F)C(=O)O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F</chem>	4.81	-	3.98
1763-23-1	PFOS	<chem>FC(F)(C(F)(F)S(=O)(=O)O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F</chem>	4.49	-	3.65
771-61-9	PFP	<chem>Fc1c(F)c(F)c(F)c(O)c1F</chem>	2.51	3.23	2.38
16128-96-4	S-13	<chem>[O-][N+](=O)C1=CC(Cl)=C(NC(=O)C2=C(O)C(C(C)(C)C)=CC(Cl)=C2)C=C1</chem>	6.47	-	5.65
69-72-7	SA	<chem>OC(=O)C1=C(O)C=CC=C1</chem>	2.24	2.26	1.40
1198-55-6	TeCC	<chem>ClC1=C(Cl)C(=C(O)C(=C1Cl)O)Cl</chem>	3.61	4.29	3.45

4358-26-3	TPB	<chem>C1=CC=C(C=C1)B(C1=CC=CC=C1)(C1=CC=CC=C1)C1=CC=CC=C1</chem>	7.28	-	6.47
2338-29-6	TTFB	<chem>ClC1=C2NC(C(F)(F)(F))=NC2=C(Cl)C(Cl)=C1(Cl)</chem>	4.78	-	3.95
81-81-2	Warf	<chem>CC(CC(C1=CC=CC=C1)C3=C(O)C2=C(C=CC=C2)OC3=O)=O</chem>	2.23	2.7	1.85

### 3 pp-LFER based prediction

#### 3.1 Derivation of the solute descriptors of ions from the solute descriptors of neutral compounds

As outlined in the summary, section 1.4, solute descriptors of certain ions can be recalculated based on the solute descriptors of neutral compounds ( $E$ ,  $S$ ,  $A$ ,  $B$  and  $V$ ). For carboxylic acid anions this recalculation is as follows (Abraham and Acree, Jr, 2010a):

$$\begin{aligned}
 E(\text{ion}) &= 0.15 + E \\
 S(\text{ion}) &= 1.224 + 0.908 * E + 0.827 * S + 0.453 * V \\
 A(\text{ion}) &= -0.208 - 0.058 * S + 0.354 * A + 0.076 * V \\
 B(\text{ion}) &= 2.150 - 0.204 * S + 1.217 * B + 0.314 * V \\
 V(\text{ion}) &= -0.0215 + V \\
 J^-(\text{ion}) &= 1.793 + 0.267 * E - 0.195 * S + 0.350 * V \\
 J^+(\text{ion}) &= 0
 \end{aligned}$$

For phenoxides the solute descriptors of ions can be recalculated by (Abraham and Acree, Jr, 2010b):

$$\begin{aligned}
 E(\text{ion}) &= 0.15 + E \\
 S(\text{ion}) &= 4.692 + 4.639 * E - 2.9 * S + 5.326 * A + 5.218 * B - 0.776 \\
 &\quad * pK_a(\text{water}) \\
 A(\text{ion}) &= 0 \\
 B(\text{ion}) &= 1.7 + 1.103 * E - 0.732S + 0.728 * A + 0.564 * B - 0.0255 \\
 &\quad * pK_a(\text{water}) \\
 V(\text{ion}) &= -0.0215 + V \\
 J^-(\text{ion}) &= 2.165 + 2.579 * E - 1.504 * S + 1.708 * A + 0.045 * B - 0.217 \\
 &\quad * pK_a(\text{water}) \\
 J^+(\text{ion}) &= 0
 \end{aligned}$$

The recalculation of the solute descriptors for amine cations (with  $NA$  being number of hydrogen atoms attached to charged nitrogen) is as follows (Abraham and Acree, Jr, 2010b):

$$\begin{aligned}
 E(\text{ion}) &= -0.15 + E \\
 S(\text{ion}) &= 0.463 + 0.473 * S + 2.419 * B \\
 A(\text{ion}) &= -0.052 - 0.35 * E + 1.48 * S + 0.327 * NA \\
 B(\text{ion}) &= 0 \\
 V(\text{ion}) &= 0.0215 + V
 \end{aligned}$$



$$J^-(ion) = 0$$

$$J^+(ion) = 0.628 + 1.002 * E - .794 * S + 1.128 * B - 0.191 * NA$$

There are also formulas available for the recalculation of pyridinium cations (Abraham and Acree, Jr, 2010c), but there are no pyridinium cations in the experimental dataset.

### 3.2 Discussion of possible artefacts from Absolv predicted solute descriptors

In order to rule out artefacts that might arise from errors in Absolv<sup>8</sup> predicted solute descriptors, we conducted the fitting procedure for the pp-LFER of Eq. 9 from the summary, section 1.4 again – but this time we limited ourselves strictly to the solute descriptors of those ions, whose corresponding neutral compounds do have experimental solute descriptors (i.e. 11 cations and 25 anions).

In analogy to Eq. 9 in the summary, section 1.4, we took the published equation (Endo et al., 2011) for the neutral system descriptors (c, s, a, b, v), fixed these values and fitted  $j^+$  and  $j^-$  with the 36 remaining ions (instead of 74 ions in the summary, section 1.4) to yield the following Eq. I:

$$\log K_{lipw} = 0.26(\pm 0.08) + 0.85(\pm 0.05)E - 0.75(\pm 0.08)S + 0.29(\pm 0.09)A - \\ 3.84(\pm 0.10)B + 3.35(\pm 0.09)V - 1.66(\pm 0.12)J^+ + 3.92(\pm 0.06)J^-; SD = 0.974, \\ n(ion) = 36, R^2 = 0.993$$

(I)

Eq. 9 in the summary, section 1.4 and Eq. I are essentially identical;  $j^+$  and  $j^-$  only change 0.06 and 0.02 units, which is within the standard error of these parameters. Therefore it seemed justified to take advantage of solute descriptors for ions that are derived from Absolv predicted values in order to have the maximum amount of descriptors for fitting the system parameters.

---

<sup>8</sup> Advanced Chemistry Development, Inc. (ACD/Labs). Absolv prediction module data sheet. Toronto, ON (Canada). <http://www.acdlabs.com/products/percepta/predictors/absolv/>

### 3.3 pp-LFER solute descriptors

SI-2, Table 7. Solute descriptors of the neutral compounds that correspond to the 32 cations and 42 anions, which can be recalculated to solute descriptors of ions with the formulas given above (Abraham and Acree, Jr, 2010a, 2010b, 2010c). Values in bold font are taken from the 'UFZ-LSER database'<sup>9</sup> and refer to experimentally derived values, letters in normal font are Absolv predicted values.

CAS	Compoundname	A	B	S	E	V	reference
118-92-3	anthranilic acid	0.75	0.71	1.26	1.09	1.0315	Absolv
69-72-7	salicylic acid	<b>0.73</b>	<b>0.37</b>	<b>0.85</b>	<b>0.9</b>	<b>0.99</b>	<b>Abraham, M. H., Acree, W. E., Leo, A. J., Hoekman, D. (2009) New J. Chem., 33, 1685-1692.</b>
2270-20-4	5-phenylvaleric acid	0.57	0.46	1.09	0.74	1.50	Absolv
15687-27-1	ibuprofen	<b>0.56</b>	<b>0.79</b>	<b>0.7</b>	<b>0.73</b>	<b>1.777</b>	<b>Abraham, M. H., Acree, W. E., Leo, A. J., Hoekman, D. (2009) New J. Chem., 33, 1685-1692.</b>
124-07-2	octanoic acid	<b>0.62</b>	<b>0.45</b>	<b>0.65</b>	<b>0.15</b>	<b>1.31</b>	<b>Abraham, M. H. (2003) J. Environ. Monit., 5, 747-752.</b>
91-40-7	fenamic acid	0.65	0.70	1.58	1.60	1.64	Absolv
15307-86-5	diclofenac	<b>0.55</b>	<b>0.77</b>	<b>1.85</b>	<b>1.81</b>	<b>2.025</b>	<b>Abraham, M. H., Acree, W. E., Leo, A. J., Hoekman, D. (2009) New J. Chem., 33, 1685-1692.</b>
530-78-9	flufenamic acid	0.72	0.59	1.36	1.26	1.83	Absolv
335-67-1	perfluorooctanoic acid	0.84	0.29	-0.34	-0.90	1.57	Absolv
22494-42-4	diflunisal	0.70	0.44	1.50	1.55	1.63	Absolv
94-75-7	2,4-dichlorophenoxyacetic acid	0.57	0.58	1.41	1.04	1.38	Absolv
87848-99-5	acrivastine	0.57	1.45	2.00	2.07	2.81	Absolv

<sup>9</sup> Endo, S., Watanabe, N., Ulrich, N., Bronner, G., Goss, K.-U., UFZ-LSER database v 2.1, Leipzig, Germany, UFZ - Helmholtz Centre for Environmental Research. 2015 [https://www.ufz.de/index.php?en=31698&contentonly=1&lserd\\_data\[mvc\]=Public/start](https://www.ufz.de/index.php?en=31698&contentonly=1&lserd_data[mvc]=Public/start)

88-75-5	2-nitrophenol	0.05	0.37	1.05	1.02	0.949	Abraham, M. H., Andonian-Haftvan, J., Whiting, G. S., Leo, A., Taft, R. S. (1994) <i>J. Chem. Soc. Perkin Trans.</i> , 2, 1777-1791.
95-57-8	2-chlorophenol	0.32	0.31	0.88	0.85	0.898	Sprunger, L., Proctor, A., Acree, W. E., Abraham, M. H. (2007) <i>J. Chromatogr. A</i> , 1175, 162-173.
100-02-7	4-nitrophenol	0.82	0.26	1.72	1.07	0.949	Abraham, M. H., Andonian-Haftvan, J., Whiting, G. S., Leo, A., Taft, R. S. (1994) <i>J. Chem. Soc. Perkin Trans.</i> , 2, 1777-1791.
87-65-0	2,6-dichlorophenol	0.38	0.24	0.9	0.9	1.02	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , 18, 1497-1505.
771-61-9	pentafluorophenol	0.79	0.09	0.83	0.36	0.864	Zissimos, A. M., Abraham, M. H., Du, C. M., Valko, K., Bevan, C., Reynolds, D., Wood, J., Tam, K. Y. (2002) <i>J. Chem. Soc. Perkin Trans. 2</i> , 2001-2010.
573-56-8	2,6-dinitrophenol	0.17	0.48	2.04	1.22	1.124	Abraham, M. H., Du, C. M., Platts, J. A. (2000) <i>J. Org. Chem.</i> , 65, 7114-7118.
577-71-9	3,4-dinitrophenol	1.14	0.16	2.25	1.32	1.124	Abraham, M. H., Du, C. M., Platts, J. A. (2000) <i>J. Org. Chem.</i> , 65, 7114-7118.
51-28-5	2,4-dinitrophenol	0.09	0.56	1.49	1.2	1.124	Abraham, M. H., Acree, W. E., Leo, A. J., Hoekman, D. (2009) <i>New J. Chem.</i> , 33, 1685-1692.
1689-84-5	3,5-dibromo-4-hydroxy-benzonitrile	0.42	0.34	1.48	1.47	1.28	Absolv
609-93-8	4-methyl-2,6-dinitrophenol	0.00	0.36	1.59	1.18	1.26	Absolv
534-52-1	2-methyl-4,6-dinitrophenol	0.04	0.52	1.59	1.2	1.264	Abraham, M. H., Acree, W. E. (2010) <i>J. Org. Chem.</i> , 75, 3021-3026.
591-35-5	3,5-dichlorophenol	0.77	0	1.17	1.02	1.02	Abraham, M. H., Chadha, H. S., Whiting, G. S., Mitchell, R. C. (1994) <i>J. Pharm. Sci.</i> , 83, 1085-1100.

106-48-9	4-chlorophenol	<b>0.67</b>	<b>0.21</b>	<b>1.08</b>	<b>0.92</b>	<b>0.898</b>	Abraham, M. H., Andonian-Haftvan, J., Whiting, G. S., Leo, A., Taft, R. S. (1994) <i>J. Chem. Soc. Perkin Trans.</i> , <b>2</b> , 1777-1791.
88-06-2	2,4,6-trichlorophenol	<b>0.82</b>	<b>0.08</b>	<b>1.01</b>	<b>1.01</b>	<b>1.142</b>	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , <b>18</b> , 1497-1505.
1198-55-6	tetrachlorocatechol	1.35	0.01	1.14	1.22	1.32	Absolv
120-83-2	2,4-dichlorophenol	<b>0.54</b>	<b>0.17</b>	<b>0.82</b>	<b>0.96</b>	<b>1.02</b>	Abraham, M. H., Acree, W. E., Leo, A. J., Hoekman, D. (2009) <i>New J. Chem.</i> , <b>33</b> , 1685-1692.
95-77-2	3,4-dichlorophenol	<b>0.74</b>	<b>0</b>	<b>1.2</b>	<b>1.02</b>	<b>1.02</b>	Abraham, M. H., Chadha, H. S., Whiting, G. S., Mitchell, R. C. (1994) <i>J. Pharm. Sci.</i> , <b>83</b> , 1085-1100.
95-95-4	2,4,5-trichlorophenol	<b>0.73</b>	<b>0.1</b>	<b>0.92</b>	<b>1.07</b>	<b>1.142</b>	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , <b>18</b> , 1497-1505.
89365-49-1	2,4,6-tribromophenolate	0.42	0.15	1.18	1.62	1.30	Absolv
609-19-8	3,4,5-trichlorophenol	<b>0.99</b>	<b>0</b>	<b>0.92</b>	<b>1.13</b>	<b>1.142</b>	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , <b>18</b> , 1497-1505.
13979-81-2	3,5-dibromo-4-methylphenol	0.79	0.24	1.13	1.42	1.27	Absolv
4097-49-8	4-tert-butyl-2,6-dinitrophenol	0.00	0.41	1.54	1.16	1.69	Absolv
88-85-7	2-s-butyl-4,6-dinitrophenol	<b>0.17</b>	<b>0.35</b>	<b>1.95</b>	1.25	<b>1.687</b>	Bronner, G., Goss, K.-U. (2010) <i>Fluid Phase Equilibria</i> , <b>299</b> , 207-215.
58-90-2	2,3,4,6-tetrachlorophenol	0.61	0.07	1.04	1.17	1.26	Absolv
935-95-5	2,3,5,6-tetrachlorophenol	<b>0.46</b>	<b>0.22</b>	<b>0.86</b>	<b>1.11</b>	<b>1.265</b>	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , <b>18</b> , 1497-1505.
1420-07-1	2-tert-butyl-4,6-dinitrophenol	0.28	0.49	1.70	1.23	1.69	Absolv
4901-51-3	2,3,4,5-tetrachlorophenol	<b>0.7</b>	<b>0.13</b>	<b>0.88</b>	<b>1.17</b>	<b>1.265</b>	Hoover, K. R., Acree Jr, W. E., Abraham, M. H. (2005) <i>Chem. Res. Toxicol.</i> , <b>18</b> , 1497-1505.

87-86-5	pentachlorophenol	<b>0.61</b>	<b>0.09</b>	<b>0.86</b>	<b>1.22</b>	<b>1.387</b>	<b>Sprunger, L., Proctor, A., Acree, W. E., Abraham, M. H. (2007) J. Chromatogr. A, 1175, 162-173.</b>
608-71-9	pentabromophenol	<b>0.64</b>	<b>0.59</b>	<b>1.02</b>	2.19	<b>1.65</b>	<b>Stenzel, A., Goss, K.-U., Endo, S. (2013) Environ. Sci. Technol., 47, 1399-1406.</b>
36894-69-6	labetalol	1.00	1.72	2.30	2.15	2.64	Absolv
13214-66-9	4-phenylbutylamine	0.21	0.67	0.97	0.77	1.38	Absolv
88150-42-9	amlodipine	0.36	2.19	2.26	1.65	3.02	Absolv
83891-03-6	norfluoxetine	0.21	0.88	1.34	1.06	2.10	Absolv
525-66-6	propranolol	<b>0.44</b>	<b>1.31</b>	<b>1.43</b>	<b>1.84</b>	<b>2.148</b>	<b>Abraham, M. H., Ibrahim, A. (2007) Int. J. Pharma., 329, 129.</b>
699-04-7	(p-methylbenzyl)methylamine	0.13	0.56	0.75	0.75	1.24	Absolv
39099-13-3	(p-methylbenzyl)ethylamine	0.13	0.56	0.76	0.75	1.38	Absolv
39190-96-0	(p-methylbenzyl)propylamine	0.13	0.57	0.76	0.75	1.52	Absolv
16183-21-4	(p-methylbenzyl)buthylamine	0.13	0.57	0.77	0.75	1.66	Absolv
170303-38-5	(p-methylbenzyl)pentylamine	0.13	0.57	0.77	0.75	1.80	Absolv
215177-23-4	(p-methylbenzyl)heptylamine	0.13	0.58	0.77	0.74	1.94	Absolv
215177-24-5	(p-methylbenzyl)heptylamine	0.13	0.58	0.78	0.74	2.08	Absolv
51384-51-1	1-[4-(2-methoxyethyl)phenoxy]-3-[(1-methylethyl)amino]-2-propanol	0.29	1.52	1.22	1.10	2.26	Absolv
29122-68-7	4-[2-hydroxy-3-[(1-methylethyl)amino]propoxy]-benzeneacetamide	<b>0.69</b>	<b>2</b>	<b>1.88</b>	<b>1.45</b>	<b>2.176</b>	<b>Abraham, M. H., Ibrahim, A., Acree, W. E. (2008) Eur. J. Med. Chem., 43, 478-485.</b>
54910-89-3	fluoxetine	<b>0.1</b>	<b>0.93</b>	<b>1.3</b>	<b>1</b>	<b>2.24</b>	<b>Abraham, M. H., Ibrahim, A., Acree, W. E. (2008) Eur. J. Med. Chem., 43, 478-485.</b>
312753-06-3	indacaterol	1.08	1.91	2.32	2.76	3.09	Absolv

89365-50-4	salmeterol	1.19	2.11	1.97	2.05	3.49	Absolv
37517-30-9	acebutolol	<b>0.9</b>	<b>2.1</b>	<b>2.42</b>	<b>1.6</b>	<b>2.756</b>	<b>Abraham, M. H., Ibrahim, A. (2007) Int. J. Pharma., 329, 129.</b>
13655-52-2	alprenolol	<b>0.1</b>	<b>1.25</b>	<b>1.03</b>	<b>1.25</b>	<b>2.159</b>	<b>Sprunger, L., Blake-Taylor, B. H., Wairegi, A., Acree, W. E., Abraham, M. H. (2007) J. Chromatogr. A, 1160, 235-245.</b>
23284-25-5	bupranolol	0.29	1.30	1.09	1.16	2.18	Absolv
36894-69-6	labetalol	1.00	1.72	2.30	2.15	2.64	Absolv
42200-33-9	nadolol	0.83	1.90	1.56	1.68	2.49	Absolv
6452-71-7	oxprenolol	<b>0.17</b>	<b>1.62</b>	<b>1.49</b>	<b>1.31</b>	<b>2.217</b>	<b>Abraham, M. H., Ibrahim, A. (2007) Int. J. Pharma., 329, 129.</b>
13523-86-9	pindolol	<b>0.3</b>	<b>1.48</b>	<b>1.65</b>	<b>1.7</b>	<b>2.009</b>	<b>Abraham, M. H., Ibrahim, A. (2007) Int. J. Pharma., 329, 129.</b>
2933-94-0	toliprolol	0.29	1.30	1.06	1.06	1.92	Absolv
137-58-6	lidocaine	<b>0.06</b>	<b>1.24</b>	<b>1.47</b>	<b>1.11</b>	<b>2.059</b>	<b>Abraham, M. H., Ibrahim, A. (2007) Int. J. Pharma., 329, 129.</b>
94-24-6	tetracaine	0.13	1.25	1.45	1.02	2.26	Absolv
59-46-1	procaine	<b>0.23</b>	<b>1.47</b>	<b>1.26</b>	<b>1.14</b>	<b>1.977</b>	<b>Abraham, M. H., Ibrahim, A., Zhao, Y., Acree, W. E. (2006) J. Pharma. Sci., 95, 2091-2100.</b>
130-95-0	(R)-(6-Methoxyquinolin-4-yl)((2S,4S,8R)-8-vinylquinuclidin-2-yl)methanol	<b>0.37</b>	<b>1.97</b>	<b>1.23</b>	<b>2.47</b>	<b>2.551</b>	<b>Zissimos, A. M., Abraham, M. H., Barker, M. C., Box, K., Tam, Y. K. (2002) J. Chem. Soc. Perkin Trans. 2, 470-477.</b>
83881-51-0	ceterizine	0.57	1.76	2.24	2.05	2.94	Absolv
50-53-3	chlorpromazine	<b>0</b>	<b>1.01</b>	<b>1.57</b>	<b>2.2</b>	<b>2.406</b>	<b>Abraham, M. H., Acree, W. E. (2004) New J. Chem., 28, 1538-1543.</b>

68-88-2	hydroxyzine	<b>0.1</b>	<b>1.89</b>	<b>2.21</b>	<b>2</b>	<b>2.923</b>	<b>Abraham, M. H., Ibrahim, A., Acree, W. E. (2008) Eur. J. Med. Chem., 43, 478-485.</b>		
20574-50-9	morantel	0	0.76	0.76	1.25	1.7733	Absolv		

**SI-2, Table 8. Solute descriptors of 32 cationic and 42 anionic compounds. The values are based on the values in the previous Table 7 and have been recalculated using the formulas given above (Abraham and Acree, Jr, 2010a, 2010b, 2010c).**

Compoundname	A (ion)	B (ion)	S (ion)	E (ion)	V (ion)	J <sup>+</sup> (ion)	J <sup>-</sup> (ion)	class
anthranilic acid anion	0.06	3.08	3.72	1.24	1.01	0.00	2.20	carboxylic acid
salicylic acid anion	0.08	2.74	3.19	1.05	0.97	0.00	2.21	carboxylic acid
5-phenylvaleric acid anion	0.04	2.96	3.47	0.89	1.47	0.00	2.30	carboxylic acid
ibuprofen anion	0.08	3.53	3.27	0.88	1.76	0.00	2.47	carboxylic acid
octanoic acid anion	0.07	2.98	2.49	0.30	1.29	0.00	2.16	carboxylic acid
fenamic acid anion	0.06	3.19	4.73	1.75	1.62	0.00	2.49	carboxylic acid
diclofenac anion	0.03	3.35	5.31	1.96	2.00	0.00	2.62	carboxylic acid
flufenamic acid anion	0.11	3.17	4.32	1.41	1.81	0.00	2.51	carboxylic acid
perfluorooctanoic acid anion	0.23	3.07	0.84	-0.75	1.55	0.00	2.17	carboxylic acid
diflunisal anion	0.08	2.89	4.61	1.70	1.61	0.00	2.49	carboxylic acid
2,4-dichlorophenoxyacetic acid anion	0.02	3.00	3.96	1.19	1.35	0.00	2.28	carboxylic acid
acrivastine anion	0.09	4.39	6.03	2.22	2.79	0.00	2.94	carboxylic acid
2-nitrophenol anion	0.00	2.12	2.97	1.17	0.93	0.00	1.75	phenol

2-chlorophenol anion	0.00	2.18	2.76	1.00	0.88	0.00	1.74	phenol
4-nitrophenol anion	0.00	2.18	4.90	1.22	0.93	0.00	2.21	phenol
2,6-dichlorophenol anion	0.00	2.27	4.12	1.05	1.00	0.00	2.28	phenol
pentafluorophenol anion	0.00	1.97	4.34	0.51	0.84	0.00	2.00	phenol
2,6-dinitrophenol anion	0.00	1.85	4.97	1.37	1.10	0.00	1.75	phenol
3,4-dinitrophenol anion	0.00	2.29	6.94	1.47	1.10	0.00	2.95	phenol
2,4-dinitrophenol anion	0.00	2.21	6.28	1.35	1.10	0.00	2.34	phenol
3,5-dibromo-4-hydroxy-benzonitrile anion	0.00	2.63	8.06	1.62	1.26	0.00	3.58	phenol
4-methyl-2,6-dinitrophenol anion	0.00	1.94	4.28	1.33	1.24	0.00	1.95	phenol
2-methyl-4,6-dinitrophenol anion	0.00	2.07	5.23	1.35	1.24	0.00	2.02	phenol
3,5-dichlorophenol anion	0.00	2.32	3.72	1.17	1.00	0.00	2.56	phenol
4-chlorophenol anion	0.00	2.29	3.21	1.07	0.88	0.00	2.03	phenol
2,4,6-trichlorophenol anion	0.00	2.56	6.46	1.16	1.12	0.00	3.32	phenol
tetrachlorocatechol anion	0.00	3.05	9.66	1.37	1.30	0.00	4.61	phenol



2,4-dichlorophenol anion	0.00	2.45	4.44	1.11	1.00	0.00	2.63	phenol
3,4-dichlorophenol anion	0.00	2.27	3.22	1.17	1.00	0.00	2.39	phenol
2,4,5-trichlorophenol anion	0.00	2.62	6.01	1.22	1.12	0.00	3.29	phenol
2,4,6-tribromophenolate anion	0.00	2.84	6.53	1.77	1.28	0.00	3.82	phenol
3,4,5-trichlorophenol anion	0.00	2.80	6.54	1.28	1.12	0.00	3.71	phenol
3,5-dibromo-4-methylphenol anion	0.00	2.94	7.04	1.57	1.24	0.00	3.69	phenol
4-tert-butyl-2,6-dinitrophenol anion	0.00	1.98	4.56	1.31	1.67	0.00	1.97	phenol
2-s-butyl-4,6-dinitrophenol anion	0.00	1.85	3.98	1.40	1.67	0.00	1.76	phenol
2,3,4,6-tetrachlorophenol anion	0.00	2.58	6.53	1.32	1.24	0.00	3.49	phenol
2,3,5,6-tetrachlorophenol anion	0.00	2.62	6.96	1.26	1.24	0.00	3.41	phenol
2-tert-butyl-4,6-dinitrophenol anion	0.00	2.17	5.79	1.38	1.67	0.00	2.24	phenol
2,3,4,5-tetrachlorophenol anion	0.00	2.77	7.05	1.32	1.24	0.00	3.68	phenol
pentachlorophenol anion	0.00	2.79	7.89	1.37	1.37	0.00	4.03	phenol
pentabromophenol anion	0.00	4.05	14.80	2.34	1.63	0.00	6.40	phenol
labetalol anion	0.00	3.90	16.59	2.30	2.62	0.00	4.44	phenol
4-phenylbutylamine cation	2.10	0.00	2.54	0.62	1.40	0.81	0.00	primary amine
amlodipine cation	3.70	0.00	6.83	1.50	3.05	2.38	0.00	primary amine
norfluoxetine cation	2.54	0.00	3.23	0.91	2.12	1.05	0.00	primary amine

propranolol cation	2.07	0.00	4.31	1.69	2.17	2.43	0.00	secondary amine
(p-methylbenzyl)methylamine cation	1.45	0.00	2.17	0.60	1.26	1.03	0.00	secondary amine
(p-methylbenzyl)ethylamine cation	1.46	0.00	2.18	0.60	1.40	1.03	0.00	secondary amine
(p-methylbenzyl)propylamine cation	1.46	0.00	2.20	0.60	1.54	1.04	0.00	secondary amine
(p-methylbenzyl)butylamine cation	1.48	0.00	2.21	0.60	1.68	1.03	0.00	secondary amine
(p-methylbenzyl)pentylamine cation	1.48	0.00	2.21	0.60	1.82	1.03	0.00	secondary amine
(p-methylbenzyl)heptylamine cation	1.48	0.00	2.23	0.59	1.96	1.03	0.00	secondary amine
(p-methylbenzyl)heptylamine cation	1.50	0.00	2.23	0.59	2.11	1.02	0.00	secondary amine
1-[4-(2-methoxyethyl)phenoxy]-3-[(1-methylethyl)amino]-2-propanol cation	2.02	0.00	4.72	0.95	2.28	2.09	0.00	secondary amine
4-[2-hydroxy-3-[(1-methylethyl)amino]propoxy]-benzeneacetamide cation	2.88	0.00	6.19	1.30	2.20	2.46	0.00	secondary amine
fluoxetine cation	2.18	0.00	3.33	0.85	2.26	1.26	0.00	secondary amine
indacaterol cation	3.07	0.00	6.18	2.61	3.11	3.32	0.00	secondary amine
salmeterol cation	2.80	0.00	6.50	1.90	3.51	3.12	0.00	secondary amine
acebutolol cation	3.62	0.00	6.69	1.45	2.78	2.30	0.00	secondary amine
alprenolol cation	1.69	0.00	3.97	1.10	2.18	2.09	0.00	secondary amine
bupranolol cation	1.81	0.00	4.12	1.01	2.20	2.01	0.00	secondary amine
labetalol cation	3.25	0.00	5.71	2.00	2.66	2.51	0.00	secondary amine
nadolol cation	2.32	0.00	5.80	1.53	2.51	2.83	0.00	secondary amine

oxprenolol cation	2.35	0.00	5.09	1.16	2.24	2.20	0.00	secondary amine
pindolol cation	2.45	0.00	4.82	1.55	2.03	2.31	0.00	secondary amine
toliprolol cation	1.80	0.00	4.11	0.91	1.94	1.93	0.00	secondary amine
lidocaine cation	2.06	0.00	4.16	0.96	2.08	1.78	0.00	tertiary amine
tetracaine cation	2.06	0.00	4.17	0.87	2.28	1.72	0.00	tertiary amine
procaine cation	1.74	0.00	4.61	0.99	2.00	2.24	0.00	tertiary amine
(R)-(6-Methoxyquinolin-4-yl)((2S,4S,8R)-8-vinylquinuclidin-2-yl)methanol cation	1.23	0.00	5.81	2.32	2.57	4.16	0.00	tertiary amine
ceterizine cation	2.87	0.00	5.78	1.90	2.96	2.70	0.00	tertiary amine
chlorpromazine cation	1.83	0.00	3.65	2.05	2.43	2.53	0.00	tertiary amine
hydroxyzine cation	2.85	0.00	6.08	1.85	2.94	2.82	0.00	tertiary amine
morantel cation	0.96	0.00	2.66	1.10	1.79	1.94	0.00	tertiary amine

## 4 COSMOmic

### 4.1 Modelling details

As outlined in detail in (Endo et al., 2011) for neutral compounds, the main factor influencing  $K_{lipw}$  is the membrane fluidity. Liposomes below the main phase transition temperature are in a ‘gel phase’ state with low fluidity and exhibit roughly 20 to 100 times lower values of  $K_{lipw}$  than liposomes in the ‘liquid-crystalline’ phase. Because the ‘liquid-crystalline’ phase is considered to be the natural condition, care has been taken in this study that only experimental data above the main phase transition temperature have been used.

In contrast,  $K_{lipw}$  values measured with different lipids above the main phase transition temperature exhibit mostly only up to +/- 0.2 log units variations for the different compounds (Endo et al., 2011). These small differences are superimposed by differences in the experimental method and interlaboratory differences (which might very well be higher than 0.2 log units, see SI, section 1) and as in agreement with these findings, these differences cannot be distinguished by using either DMPC or POPC within COSMOmic as shown previously (Bittermann et al., 2014). But also are these differences superimposed by differences in the experimental method and interlaboratory differences (which might very well be higher than 0.2 log units, see SI, section 1). Unfortunately there is only a very limited amount of  $K_{lipw}$  values for organic ions measured with different lipid types – but the data listed in the SI (section 1, marked in grey) give the same picture as (Endo et al., 2011) reported for neutral compounds.

Therefore, a DMPC membrane can - to the best of our knowledge - considered to be a reasonable model for an ‘average’ phospholipid-membrane above the main phase transition temperature. Using a POPC membrane instead would have yielded equally good results, requiring the same effort, as shown previously (Bittermann et al., 2014).

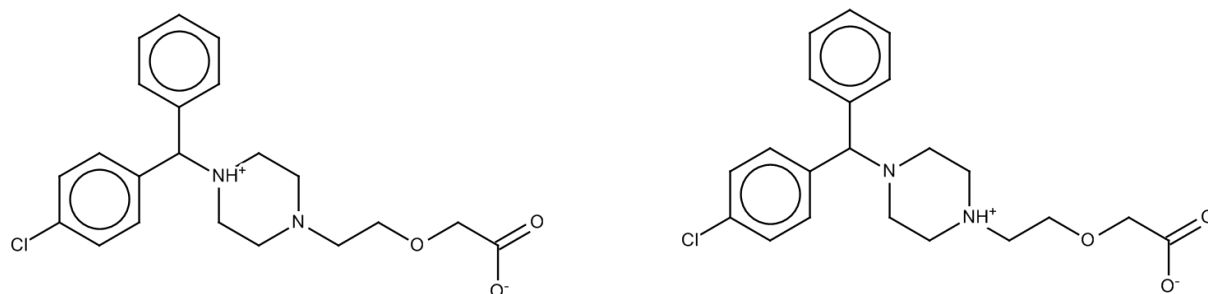
Similarly, the temperature plays only a minor role for  $K_{lipw}$  as long as it does not go below the main phase transition temperature – artefacts from different experimental methods and different laboratories are a much higher concern. While the empirical correlation approach with  $\log K_{ow}$  and the pp-LFER extension for ionic compounds do not account for temperature differences, the temperature for COSMOmic calculations is set to 25 °C (despite a real DMPC membrane being in the ‘gel phase’ state at 25 °C). COSMOmic is parametrized only for the liquid crystalline state and does not account for a gel-phase: the 3 dimensional structure of the membrane which is needed for the calculation in COSMOmic is derived from an MD-

simulation. This membrane is in the liquid crystalline state (Jakobtorweihen et al., 2013) – after virtually slicing the membrane and putting it into the COSMO-RS based part of the calculation, no changes in the 3 dimensional structure and hence the fluidity of the membrane can be considered (Bittermann et al., 2014; Klamt et al., 2008).

In theory, *COSMOmic* can also describe different kinds of phospholipids – but up to now only the membrane potentials of DMPC and POPC have been parametrized (based on the same experimental  $K_{lipw}$  values and yielding equal predictive power!) (Bittermann et al., 2014). In the authors opinion, there is not enough experimental data for the different kinds of phospholipid membranes (e.g. membranes containing high amounts of cholesterol) to re-parametrize *COSMOmic*. Looking at the data gathered in this work, DMPC and POPC (or egg-PC) membranes cannot be distinguished.

#### 4.2 Treatment of cetirizine

The cosmo files have been derived as drafted in the summary, section 1.4 and outlined in greater detail above. For the ceterizine cation as well as for the ceterizine zwitterion there are two equivalent structures possible, as shown in SI-2, Fig. 7. For the *COSMOmic* calculation, cosmo files of both possible structures have been calculated and have been treated as different conformers of the same species.



SI-2, Figure 7. Two equivalent structures for the ceterizine zwitterion which have been treated as different conformers in the *COSMOmic* calculation.

### 4.3 Calculated data using COSMOmic

SI-2, Table 9. Calculated  $\log K_{lipw}$  values of cations, anions and zwitterions using a dmPC membrane (Jakobtorweihen et al., 2013) with COSMOmic (Bittermann et al., 2014; Klamt et al., 2008).

CAS	Compoundname	abbreviation	class detail	log (exp)	$K_{lipw}$	log (calcd)	$K_{lipw}$	Charge
18198-39-5	tetraphenylphosphonium	TPP	other	1.19		3.10		K+
95-64-7	3,4-dimethylaniline	34DMA	primary amine	1.99		2.45		K+
88-05-1	2,4,6-trimethylaniline	246TMA	primary amine	2.12		2.17		K+
13214-66-9	4-phenylbutylamine	4-PhenButA	primary amine	2.12		3.01		K+
88150-42-9	amlodipine	Amlodip	primary amine	3.75		4.41		K+
83891-03-6	norfluoxetine	Norfluox	primary amine	3.84		4.86		K+
118-92-3	anthranilic_acid	AA_cation	primary amine	1.97		2.04		K+
525-66-6	propranolol	Prop	secondary amine	2.74		3.11		K+
699-04-7	(p-methylbenzyl)methylamine	MBMetA	secondary amine	2.54		1.69		K+
39099-13-3	(p-methylbenzyl)ethylamine	MBEthA	secondary amine	2.26		1.66		K+
39190-96-	(p-methylbenzyl)propylamine	MBPropA	secondary amine	2.11		1.86		K+

0							
16183-21-4	(p-methylbenzyl)buthylamine	MBButA	secondary amine	1.54	2.17		K+
170303-38-5	(p-methylbenzyl)pentylamine	MBPentA	secondary amine	1.84	2.55		K+
215177-23-4	(p-methylbenzyl)hexylamine	MBHexA	secondary amine	2.43	2.96		K+
215177-24-5	(p-methylbenzyl)heptylamine	MBHepA	secondary amine	2.71	3.16		K+
51384-51-1	1-[4-(2-methoxyethyl)phenoxy]-3-[(1-methylethyl)amino]-2-propanol	Metro	secondary amine	1.28	2.62		K+
29122-68-7	4-[2-hydroxy-3-[(1-methylethyl)amino]propoxy]-benzeneacetamide	Aten	secondary amine	1.01	0.78		K+
54910-89-3	fluoxetine	Fluox	secondary amine	4.03	4.09		K+
312753-06-3	indacaterol	Indac	secondary amine	3.56	4.61		K+
89365-50-4	salmeterol	Salmet	secondary amine	3.67	6.10		K+
37517-30-	acebutolol	ABL	secondary amine	0.66	2.00		K+

9							
13655-52-2	alprenolol	APL	secondary amine	2.17	3.03		K+
23284-25-5	bupranolol	BPL	secondary amine	2.49	2.71		K+
36894-69-6	labetalol	Lab_c	secondary amine	2.32	3.31		K+
42200-33-9	nadolol	NDL	secondary amine	0.95	1.80		K+
6452-71-7	oxprenolol	OPL	secondary amine	1.51	2.21		K+
13523-86-9	pindolol	PDL	secondary amine	1.40	2.35		K+
2933-94-0	toliprolol	TPL	secondary amine	1.49	2.62		K+
137-58-6	lidocaine	Lido	tertiary amine	1.07	1.59		K+
94-24-6	tetracaine	Tetrac	tertiary amine	2.11	2.80		K+
59-46-1	procaine	Proc	tertiary amine	0.73	1.01		K+
130-95-0	(R)-(6-Methoxyquinolin-4-yl)((2S,4S,8R)-8-vinylquinuclidin-2-yl)methanol	Quinine	tertiary amine	2.19	2.00		K+
83881-51-0	ceterizine	Cet_c	tertiary amine	3.20	3.94		K+



50-53-3	chlorpromazine	CLP	tertiary amine	3.40	3.20	K+
68-88-2	hydroxyzine	Hyd	tertiary amine	2.80	3.01	K+
20574-50-9	morantel	Mor	tertiary amine	2.00	1.21	K+
1689-84-5	3,5-dibromo-4-hydroxy-benzonitrile	Bromox	bromophenol	2.10	2.72	A-
89365-49-1	2,4,6-tribromophenolate	246TriBP	bromophenol	3.07	3.48	A-
13979-81-2	3,5-dibromo-4-methylphenol	35DBC	bromophenol	3.18	3.20	A-
608-71-9	pentabromophenol	PBrP	bromophenol	5.02	4.09	A-
118-92-3	anthranilic acid	AA_anion	carboxylic acid	0.31	1.77	A-
69-72-7	salicylic acid	SA	carboxylic acid	0.97	2.07	A-
2270-20-4	5-phenylvaleric acid	5-PA	carboxylic acid	1.66	1.88	A-
15687-27-1	ibuprofen	Ibu	carboxylic acid	1.81	2.29	A-
124-07-2	octanoic acid	Oct	carboxylic acid	0.52	1.77	A-
91-40-7	fenamic acid	Fen	carboxylic acid	2.28	2.80	A-
15307-86-5	diclofenac	Dic	carboxylic acid	2.64	2.99	A-
530-78-9	flufenamic acid	Flu	carboxylic acid	3.61	3.05	A-
335-67-1	perfluorooctanoic acid	PFOA	carboxylic acid	2.34	2.88	A-

22494-42-4	diflunisal	Dif	carboxylic acid	2.75	2.86	A-
94-75-7	2,4-dichlorophenoxyacetic acid	2,4-D	carboxylic acid	1.70	2.04	A-
87848-99-5	acrivastine	Acr_a	carboxylic acid	2.60	3.31	A-
95-57-8	2-chlorophenol	2CP	chlorophenol	0.92	2.44	A-
87-65-0	2,6-dichlorophenol	26DCP	chlorophenol	1.41	2.85	A-
591-35-5	3,5-dichlorophenol	35DCP	chlorophenol	2.47	2.98	A-
106-48-9	4-chlorophenol	4CP	chlorophenol	2.51	2.47	A-
88-06-2	2,4,6-trichlorophenol	246TriCP	chlorophenol	2.52	3.16	A-
1198-55-6	tetrachlorocatechol	TeCC	chlorophenol	2.63	3.65	A-
120-83-2	2,4-dichlorophenol	24DCP	chlorophenol	2.69	2.85	A-
95-77-2	3,4-dichlorophenol	34DCP	chlorophenol	2.85	2.86	A-
95-95-4	2,4,5-trichlorophenol	245TriCP	chlorophenol	2.88	3.17	A-
609-19-8	3,4,5-trichlorophenol	345TriCP	chlorophenol	3.16	3.16	A-
58-90-2	2,3,4,6-tetrachlorophenol	2346TeCP	chlorophenol	3.46	3.41	A-
935-95-5	2,3,5,6-tetrachlorophenol	2356TeCP	chlorophenol	3.49	3.46	A-
4901-51-3	2,3,4,5-tetrachlorophenol	2345TeCP	chlorophenol	3.69	3.39	A-
87-86-5	pentachlorophenol	PCP	chlorophenol	4.31	3.62	A-
327-19-5	5-nitro-2-trifluoromethylbenzimidazole	5-NB	N-acidic	1.81	3.12	A-
2338-25-2	5,6-dichloro-2-(trifluoromethyl)-	DTFB	N-acidic	3.05	3.48	A-

	benzimidazole						
2338-29-6	4,5,6,7-tetrachloro-2-(trifluoromethyl)- 1H-benzimidazole		TTFB	N-acidic	4.35	4.32	A-
555-60-2	carbonyl cyanide chlorophenylhydrazone	m-	CCCP	N-acidic	4.05	3.40	A-
370-86-5	carbonyl cyanide methoxyphenylhydrazone	p-	FCCP	N-acidic	4.22	3.64	A-
16128-96-4	5-chloro-3-tert-butyl-2'-chloro-4'- nitrosalicylanilide		S-13	N-acidic	5.05	6.63	A-
88-75-5	2-nitrophenol		2NP	nitrophenol	0.69	2.03	A-
100-02-7	4-nitrophenol		4NP	nitrophenol	0.95	1.96	A-
573-56-8	2,6-dinitrophenol		26DNP	nitrophenol	1.86	2.38	A-
577-71-9	3,4-dinitrophenol		34DNP	nitrophenol	1.90	2.71	A-
51-28-5	2,4-dinitrophenol		24DNP	nitrophenol	1.90	2.45	A-
609-93-8	4-methyl-2,6-dinitrophenol		DNPC	nitrophenol	2.26	2.38	A-
534-52-1	2-methyl-4,6-dinitrophenol		DNOC	nitrophenol	2.35	2.58	A-
4097-49-8	4-tert-butyl-2,6-dinitrophenol		Dino4terb	nitrophenol	3.23	2.74	A-
88-85-7	2-s-butyl-4,6-dinitrophenol		Dinoseb	nitrophenol	3.35	2.94	A-
1420-07-1	2-tert-butyl-4,6-dinitrophenol		Dino2terb	nitrophenol	3.59	3.05	A-
81-81-2	warfarin		Warf	other anion	1.40	2.69	A-
4358-26-3	tetraphenylborat		TPB	other anion	5.20	7.05	A-

771-61-9	pentafluorophenol	PFP	other anion	1.74	2.89	A-
36894-69-6	labetalol	Lab_a	other anion	1.84	2.86	A-
148-24-3	8-hydroxyquinoline	Oxine	quinoline	1.47	2.13	A-
130-16-5	5-chloro-8-hydroxyquinoline	Chlorox	quinoline	1.91	2.53	A-
773-76-2	5,7-dichloro-8-hydroxyquinoline	Dichlorox	quinoline	2.47	2.84	A-
521-74-4	5,7-dibromo-8-hydroxyquinoline	Dibromox	quinoline	3.03	3.06	A-
6149-03-7	4-octylbenzene-1-sulfonate	OBS	sulfonate	3.63	3.54	A-
1763-23-1	perfluorooctane-1-sulfonic acid	PFOS	sulfonate	3.15	3.53	A-
83881-51-0	ceterizine	Cet_zw	NA	2.3	1.19	zwitter
87848-99-5	acrivastine	Acr_zw	NA	1.5	2.15	zwitter

## **Supporting Information 3: Assessing the toxicity of ionic liquids – Application of the Critical Membrane Concentration approach**

### **1 Toxicity data for neutral compounds – sorted out data**

**SI-3, Table 1: Seven chemicals that have been sorted out of the original dataset (Kipka and Di Toro, 2009) because their water solubility is below the respective reported LC50 [mmol/L].**

<b>name</b>	<b>water solubility exp (PhysProp database) [mmol/L]</b>	<b>LC50 [mmol/L]</b>
1,2,4,5-tetrachlorobenzene	2.76E-03	2.80E-03
1-methylphenanthrene	1.40E-03	4.84E-03
decahydronaphthalene	6.43E-03	1.08E-02
heptane	3.39E-02	2.45E+00
hexachlorobenzene	2.18E-05	1.35E-04
propylcyclopentane	1.82E-02	2.78E-02
tert-butylbenzene	2.20E-01	4.57E-01

Experimental data on water solubility were collected from the PhysProp database using EPISuite<sup>10</sup> in the smiles batch mode. Experimental values for water solubility were available for 291 chemicals; if more than one experimental value was given in the database, the arithmetic mean was calculated. For the 70 chemicals that were not included in the PhysProp database, water solubility was predicted with EPISuite for the sake of completeness. However, it turned out that experimental water solubility values were available for all of the above listed chemicals (SI-3, Table 1) that were excluded from the toxicity dataset for neutral chemicals because the reported experimental LC50's exceeded the water solubility.

---

<sup>10</sup> U.S. EPA, EPISuite Exposure Assessment Tools and Models, US Environmental Protection Agency, 2009, <https://www.epa.gov/>.

- 15 **SI-3, Table 2: 23 acidic chemicals that have been sorted out of the original dataset** (Kipka and Di Toro,  
16 2009) **because their pKa is smaller than 9. The pKa predictions have been done with JChem<sup>11</sup>.**

IUPAC	pKa (Marvin/JChem)
2,4,6-trichlorophenol	5.99
2,4-dichlorophenol	7.44
2-chlorophenol	7.97
pentachlorophenol	4.98
2-nitrophenol	6.63
4-chlorophenol	8.96
4-nitrophenol	7.07
2,3,4-trichlorophenol	6.95
2,3-dichlorophenol	7.36
2,4,5-trichlorophenol	6.83
2,5-dichlorophenol	7.23
2,6-dichlorophenol	6.48
3,4-dichlorophenol	8.36
3,5-dichlorophenol	8.06
3-chlorophenol	8.79
3-nitrophenol	7.89
1,1,1,3,3,3-hexafluoropropan-2-ol	7.97
2-hydroxybenzamide	8.21
2,3,4,5-tetrachlorophenol	6.33
2,3,5,6-tetrachlorophenol	5.25
2,3,5-trichlorophenol	6.62
2,3,6-trichlorophenol	5.86
3,4,5-trichlorophenol	7.75

17

---

<sup>11</sup> JChem for Excel, version 15.10.2600.341, Copyright 2008-2015 ChemAxon Ltd.  
<https://www.chemaxon.com/>.

18 **SI-3, Table 3: 5 basic chemicals that have been sorted out of the original dataset** (Kipka and Di Toro, 2009)  
19 **because the pKa values of the corresponding protonated acids are larger than 5. The pKa predictions**  
20 **have been done with JChem<sup>12</sup>.**

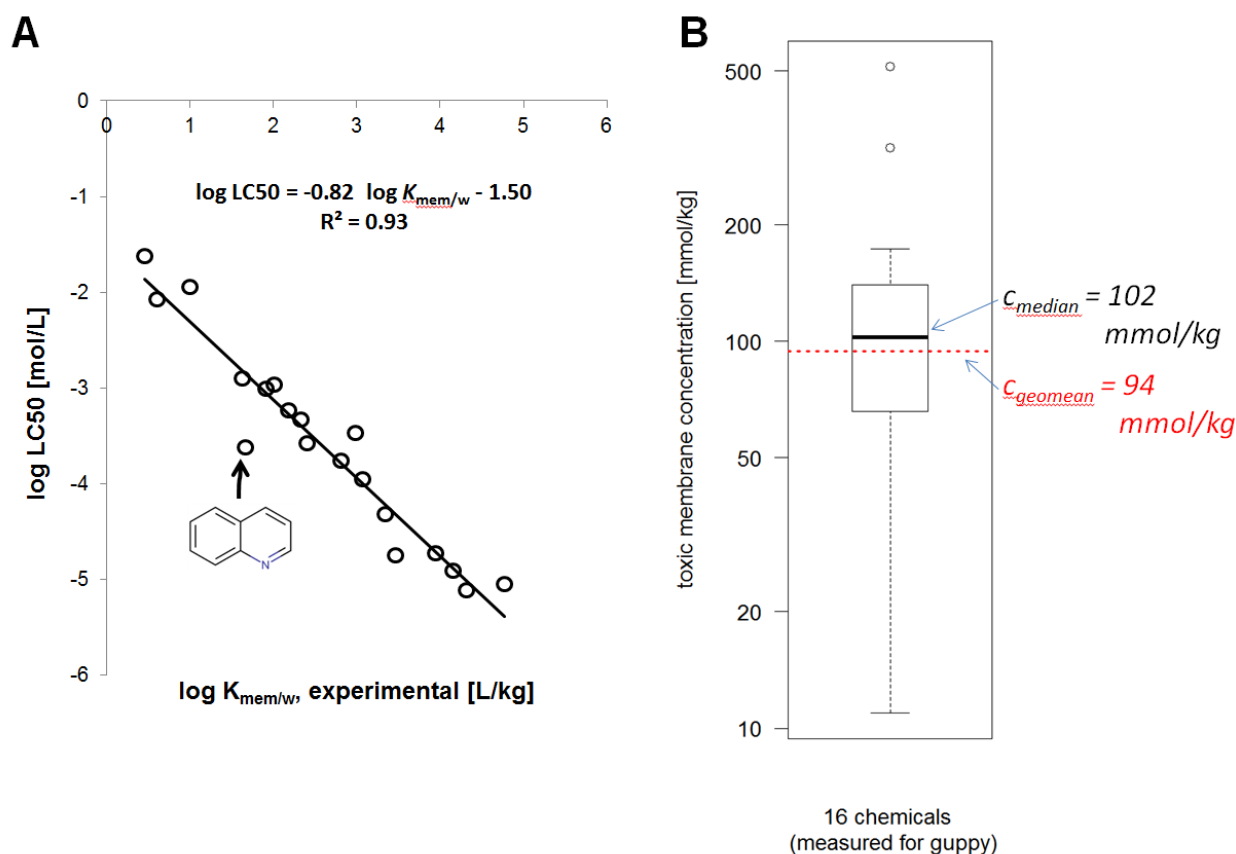
IUPAC	pKa of corresponding protonated acid (Marvin/JChem)
N,N-dimethylaniline	5.02
pyridine	5.12
N,N-diethylaniline	5.86
2-ethylpyridine	5.64
4-(hexyloxy)aniline	5.10

21  
22 As a side note it is worth to mention that the original data compilation (Kipka and Di Toro,  
23 2009) has some spelling errors in the chemical names and that in some cases different names  
24 have been used for the exact same chemical, so that a coherent re-naming of the dataset was  
25 necessary before a correct summary of the data was possible. Simply checking for  
26 unambiguous chemical names in the original list yields 399 unambiguous names – but they  
27 only represent 368 different chemicals.

---

<sup>12</sup> JChem for Excel, version 15.10.2600.341, Copyright 2008-2015 ChemAxon Ltd.  
<https://www.chemaxon.com/>.

## 28 2 Regression analyses of data from (Vaes et al., 1998)



29  
30 SI-3, Figure 1: A) log LC50 against log  $K_{mem/w}$  for 19 neutral chemicals with regression line; data taken  
31 from (Vaes et al., 1998). The regression analysis was made with Origin 2015. B) Tukey boxplot of the  
32 resulting membrane concentration calculated based on Eq. 11 of the summary, section 1.5 (the bottom  
33 and top of the box represent the first and third quartiles, the heavy line inside box represent the median;  
34 whiskers set at lowest/highest data point still within 1.5 interquartile range of the lower/upper quartile).  
35 The analysis was done with R version 2.14.2.

36 The slope of -0.82 for the data of (Vaes et al., 1998) (19 neutral chemicals) is close to the  
37 slope of -0.92 from the re-analyzed data set with 1687 LC50 for 361 neutral organic  
38 chemicals (Kipka and Di Toro, 2009), but has never been discussed in the original  
39 publication of (Vaes et al., 1998). The geometrical mean of the resulting toxic membrane  
40 concentrations (calculated according to Eq. 11 in the summary, section 1.5) is 94 mmol/kg.



### 41 **3 Comparison of pp-LFERs for $\log K_{mem/w}$ and $\log TLM$**

42  $\log K_{mem/w} = 0.26 + 0.85E - 0.75S + 0.29A - 3.84B + 3.35V; SD = 0.279,$

43  $n(neutral) = 131, R^2 = 0.979$  (SI-3, Eq. 1)

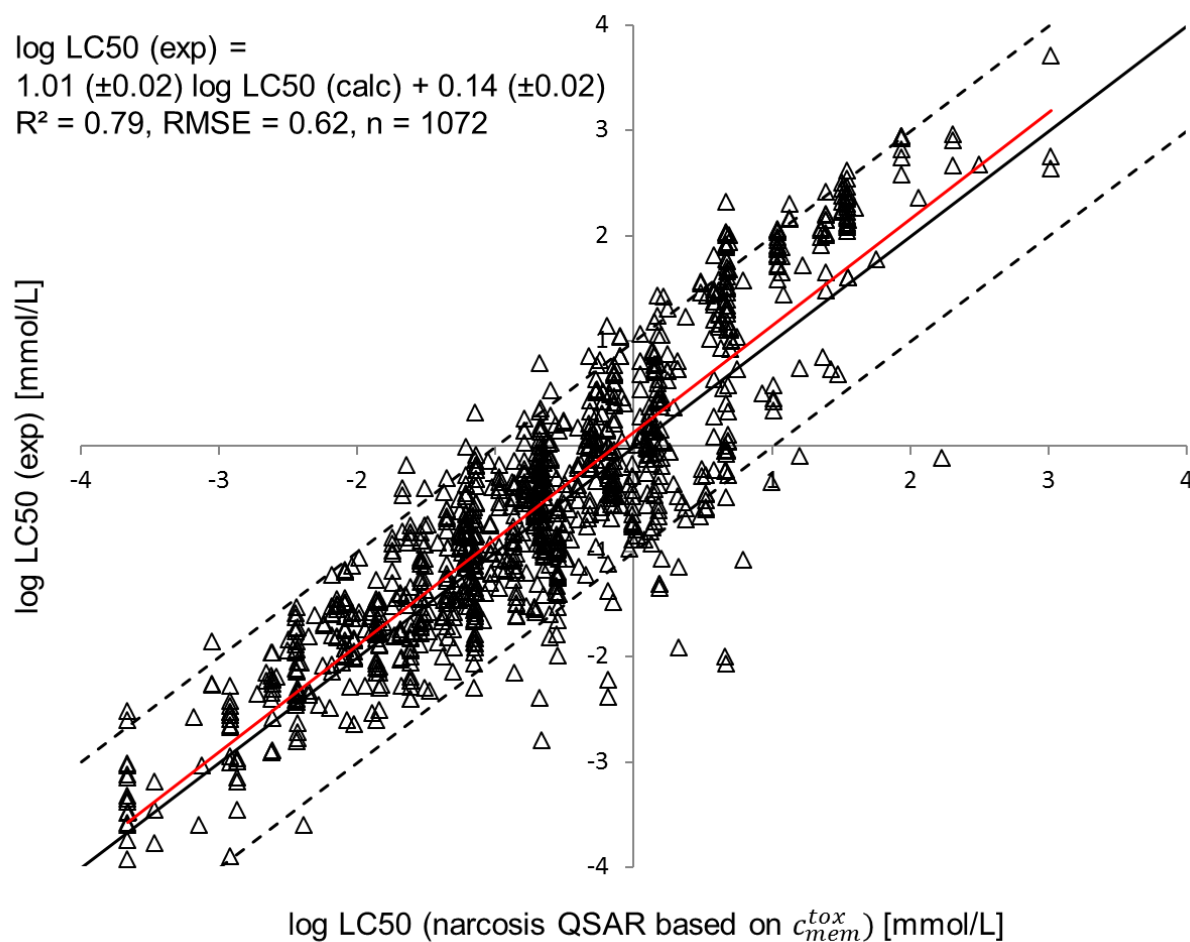
44  $\log K_{TLM} = -0.44 + 0.51E + 0.71S + 0.92A - 4.40B + 3.14V; n(neutral) = 1687,$

45  $R^2$  and  $SD$  not given (SI-3, Eq. 2)

46 pp-LFER SI-3, Eq.1 for  $\log K_{mem/w}$  is from (Endo et al., 2011), while SI-3, Eq.2 is published  
47 in (Kipka and Di Toro, 2009) and describes the partitioning to the ‘target lipid’ held  
48 responsible for toxic effects. Both equations are very similar because the membrane is the  
49 target lipid of narcotic toxicity.

### 50 **4 ‘Baseline toxicity-QSAR’ based on $c_{mem}^{tox}$**

51 Despite the variance shown in the summary, section 1.5, Fig. 11, we can use the determined  
52 geometric mean toxic membrane concentration of 105 mmol/kg(lipid) and the  $K_{mem/w}$  values  
53 predicted with pp-LFER in order to test the predictive capability of the baseline toxicity  
54 concept for the different organisms. For that reason we use our first summary of LC50 values  
55 differentiating between chemicals and organisms as described in the summary, section 1.5  
56 (i.e., 1072 organism- and chemical-specific LC50 values). Calculating the LC50 values using  
57 SI-3, Eq. 1 yields a satisfyingly predictive model with  $R^2=0.79$ ,  $RMSE=0.62$  (SI-3, Fig. 2). It  
58 has to be kept in mind, though, that the same data (summarized based only on the different  
59 chemicals) has also been used to calibrate this rather simple ‘baseline toxicity-QSAR’.



60

61 **SI-3, Figure 2: Prediction of 1072 organism- and chemical-specific LC50 values with geometric mean**  
62 **toxic membrane concentration of 105 mmol/kg(lipid) and summary, section 1.5, Eq. 12. The regression**  
63 **analysis was made with Origin 2015.**

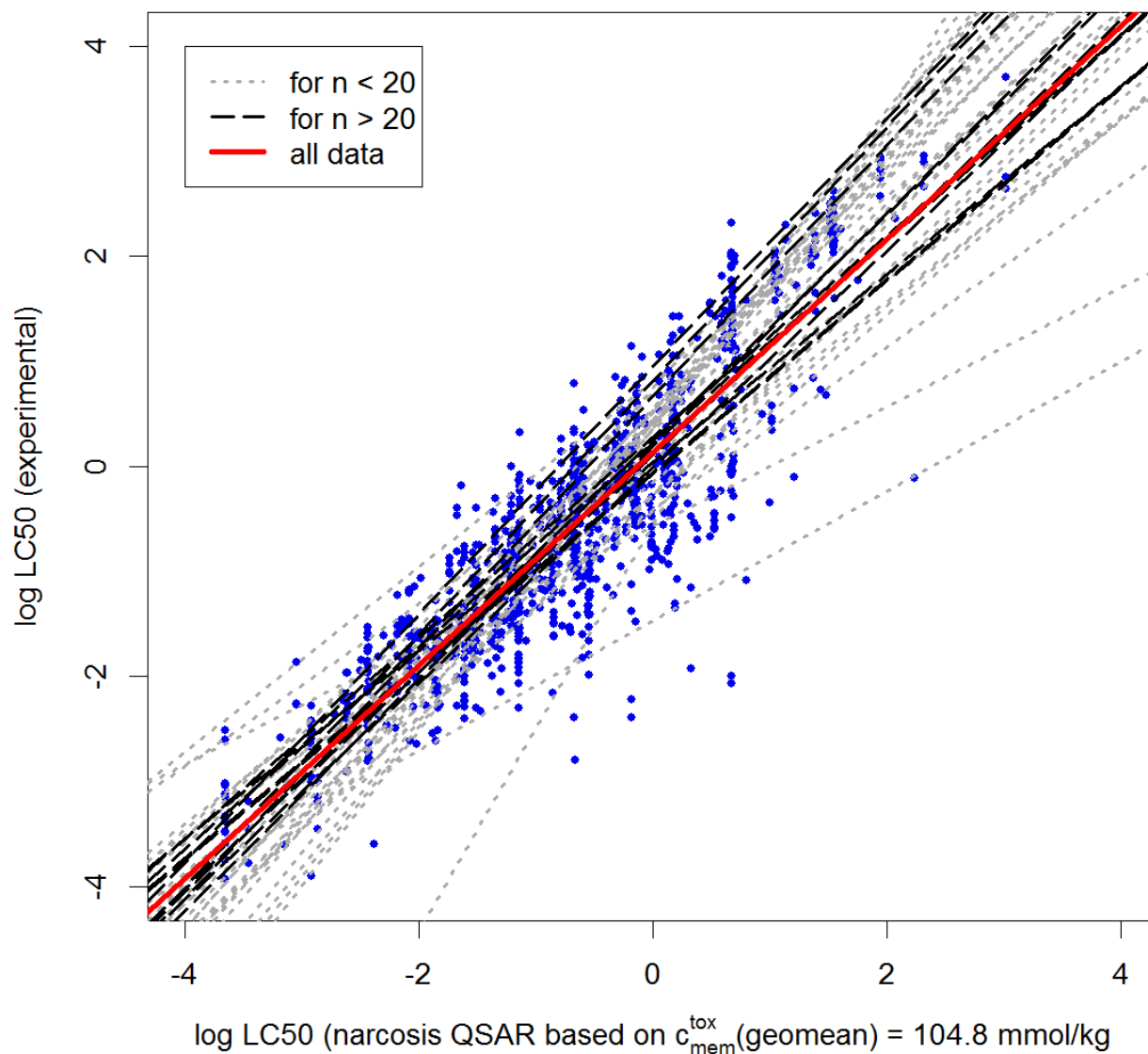
64 The statistics of the 'baseline toxicity-QSAR' shown in SI-3, Fig. 2 varies for different  
65 organisms.  $R^2$  goes from 0.48 (*Chlamydomonas reinhardtii*; n = 10) to 0.99 (*Alburnus*  
66 *alburnus*; n = 7 and *Leptochirus plumulosus*; n = 4), while the RMSE varies from 1.09  
67 (*Chlamydomonas reinhardtii*; n = 10) to 0.09 (*Leptochirus plumulosus*; n = 4) (see SI-3,  
68 Table 4 for the detailed organism-specific analysis).

69 **SI-3, Table 4: Organism-specific analysis of the ‘narcosis-QSAR’: R<sup>2</sup>, RMSE, number of different**  
 70 **chemicals, slope and intercept for the 42 different organisms.**

organism	R <sup>2</sup>	RMSE	number of chemicals	slope	intercept
<i>Leptochirus plumulosus</i>	0.99	0.095	4	0.90	-0.27
<i>Neanthes arenaceodentata</i>	0.81	0.698	4	0.57	-0.56
<i>Portunus pelagicus</i>	0.83	0.746	4	0.62	-1.47
<i>Aedes aegypti</i>	0.84	0.728	5	1.15	0.53
<i>Ambystoma mexicanum</i>	0.90	0.710	5	1.30	0.46
<i>Culex pipiens</i>	0.88	0.677	5	1.26	0.41
<i>Daphnia cucullata</i>	0.86	0.684	5	1.12	0.52
<i>Hydra oligactis</i>	0.83	0.721	5	1.25	0.41
<i>Rana catesbeiana</i>	0.95	0.822	5	1.48	0.44
<i>Jordanella floridae</i>	0.78	0.464	6	0.93	-0.43
<i>Nitocra spinipes</i>	0.97	0.745	6	0.88	0.81
<i>Orconectes immunis</i>	0.93	0.823	6	1.40	0.52
<i>Xenopus laevis</i>	0.96	0.621	6	1.39	0.36
<i>Alburnus alburnus</i>	0.99	0.732	7	1.24	0.55
<i>Ictalurus punctatus</i>	0.95	0.623	7	1.15	0.42
<i>Menidia beryllina</i>	0.93	0.645	7	0.77	0.23
<i>Rhepoxyinus abronius</i>	0.94	0.311	7	1.29	0.75
<i>Gambusia affinis</i>	0.94	0.775	8	1.42	0.37
<i>Oithona davisae</i>	0.96	0.456	8	0.86	0.05
<i>Palaemonetes pugio</i>	0.97	0.296	8	0.77	-0.41
<i>Ankistrodesmus falcatus</i>	0.94	0.491	9	0.97	0.41
<i>Tanytarsus dissimilis</i>	0.95	0.531	9	1.27	0.43
<i>Chlamydomonas reinhardtii</i>	0.48	1.085	10	1.93	-0.55
<i>Tetrahymena ellioti</i>	0.77	0.693	10	1.06	0.69
<i>Danio rerio</i>	0.68	0.439	16	1.12	0.19
<i>Daphnia pulex</i>	0.96	0.406	16	1.08	0.34
<i>Lymnaea stagnalis</i>	0.86	0.424	16	1.13	0.04
<i>Selenastrum capricornutum</i>	0.67	0.471	16	1.05	-0.29
<i>Cyprinodon variegatus</i>	0.82	0.321	18	0.86	-0.14
<i>Mysidopsis bahia</i>	0.89	0.517	19	1.06	-0.26

---

<i>Scenedemus subspicatus</i>	0.71	0.710	23	1.20	0.67
<i>Chlamydomonas angulosa</i>	0.94	0.610	28	1.21	0.82
<i>Oncorhynchus mykiss</i>	0.91	0.564	28	1.11	0.19
<i>Artemia salina</i>	0.84	0.400	33	0.97	0.22
<i>Chlorella vulgaris</i>	0.93	0.805	33	1.18	0.96
<i>Leucisus idus melanotus</i>	0.55	0.703	58	0.96	0.28
<i>Oryzias latipes</i>	0.70	0.662	58	0.89	0.04
<i>Daphnia magna</i>	0.89	0.456	59	1.07	0.25
<i>Carassius auratus</i>	0.53	0.821	62	0.92	-0.07
<i>Lepomis macrochirus</i>	0.84	0.569	69	1.03	0.11
<i>Poecilia reticulata</i>	0.77	0.576	148	1.03	0.00
<i>Pimephales promelas</i>	0.75	0.611	216	0.92	-0.06
<b>total</b>	<b>0.79</b>	<b>0.62</b>	<b>1072</b>	<b>1.01</b>	<b>0.14</b>



71

72 SI-3, Figure 3: Organism-specific analysis of the ‘narcosis-QSAR’ based on the narcotic membrane  
73 concentration of 104.8 mmol/kg. A regression line is drawn for each organism according to the number  $n$   
74 of different chemicals per organism (dotted grey for  $n < 20$ , dashed black for  $n > 20$  and thick red for all  
75 values combined).

76

77 **5 Toxicity of ILs**

78 **SI-3, Table 5: Statistics for toxic membrane concentrations for ILs for different organisms calculated**  
 79 **with Eq. 13 of the summary, section 1.5.**

organism	median toxic membrane concentration	geometric mean toxic membrane concentration	standard deviation of the log-normal distribution of toxic membrane concentrations
<i>A. fischeri</i>	8.6	22.8	1-563
<i>E. coli</i>	1100.3	1431.6	334-6143
<i>P. subcapitata</i>	8.2	1.45	0.02-89
<i>S. vacuolatus</i>	2.8	1.19	0.01-99
IPC-81	48.6	109.2	2-4984
HeLa	187.0	185.6	23-1503
MCF7	57.6	21.6	0.3-1598
<i>L. Minor</i>	5.8	10.7	0.2-647
<i>D. magna</i>	0.2	0.6	0.1-7

80

81 **SI-3, Table 6: Predicted log  $K_{lipw}$  (COSMOmic 1601). For IL-acronyms see (Stolte et al., 2007).**

IL chemical	log $K_{lipw}$ (COSMOmic 1601)	smiles
IM12 cation	-0.73	<chem>[N+]1(C)=CN(C=C1)CC</chem>
IM13 cation	-0.40	<chem>[N+]1(C)=CN(C=C1)CCC</chem>
IM14 cation	0.12	<chem>[N+]1(C)=CN(C=C1)CCCC</chem>
IM15 cation	0.62	<chem>[N+]1(C)=CN(C=C1)CCCCC</chem>
IM16 cation	1.23	<chem>[N+]1(C)=CN(C=C1)CCCCCC</chem>
IM17 cation	1.76	<chem>[N+]1(C)=CN(C=C1)CCCCCCC</chem>
IM18 cation	2.31	<chem>[N+]1(C)=CN(C=C1)CCCCCCCC</chem>
IM19 cation	2.90	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCC</chem>
IM1-10 cation	3.45	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCC</chem>
IM1-14 cation	5.78	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCCCCCC</chem>
IM1-16 cation	6.85	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCCCCCCCC</chem>
IM1-18 cation	8.08	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCCCCCCCCCC</chem>
IM1-19 cation	8.69	<chem>[N+]1(C)=CN(C=C1)CCCCCCCCCCCCCCCCC</chem>
IM24 cation	0.44	<chem>[N+]1(CC)=CN(C=C1)CCCC</chem>

IM25 cation	0.91	[N+]1(CC)=CN(C=C1)CCCC
IM26 cation	1.36	[N+]1(CC)=CN(C=C1)CCCCC
IM2-10 cation	3.69	[N+]1(CC)=CN(C=C1)CCCCCCCCC
Py4 cation	-0.05	[n+]1(ccccc1)CCCC
Py8 cation	2.17	[n+]1(ccccc1)CCCCCCCC
Py4-2Me cation	-0.01	[n+]1(c(C)cccc1)CCCC
Py4-3Me cation	0.15	[n+]1(cc(C)ccc1)CCCC
Py6-3Me cation	1.17	[n+]1(cc(C)ccc1)CCCCC
Py6-4Me cation	1.15	[n+]1(ccc(C)cc1)CCCCC
Py8-3Me cation	2.22	[n+]1(cc(C)ccc1)CCCCCCCC
Py8-4Me cation	2.15	[n+]1(ccc(C)cc1)CCCCCCCC
Pyr14 cation	-0.23	[N+]1(C)(CCCC1)CCCC
Pyr16 cation	0.85	[N+]1(C)(CCCC1)CCCCC
Mor14 cation	-0.38	[N+]1(C)(CCOCC1)CCCC
Pip14 cation	0.00	[N+]1(C)(CCCCC1)CCCC
Quin4 cation	0.56	[n+]1(cccc2cccc12)CCCC
Quin6 cation	1.55	[n+]1(cccc2cccc12)CCCCC
Quin8 cation	2.75	[n+]1(cccc2cccc12)CCCCCCCC
N1114 cation	-0.26	[N+](C)(C)(C)CCCC
N1124 cation	-0.10	[N+](C)(C)(CC)CCCC
N2222 cation	-0.80	[N+](CC)(CC)(CC)CC
N2226 cation	1.22	[N+](CC)(CC)(CC)CCCCC
N4444 cation	3.09	[N+](CCCC)(CCCC)(CCCC)CCCC
P4444 cation	3.42	[P+](CCCC)(CCCC)(CCCC)CCCC
Cl anion	0.16	[Cl-]
BF4 anion	1.42	[B-](F)(F)(F)F
PF6 anion	2.79	[P-](F)(F)(F)(F)(F)F
(CF3SO2)2N anion	3.02	[N-](S(=O)(=O)C(F)(F)F)S(=O)(=O)C(F)(F)F
Br anion	0.30	[Br-]
(CN)2N anion	1.23	[N-](C#N)C#N

## 82 SI-3, Table 7: Duration of test/exposure

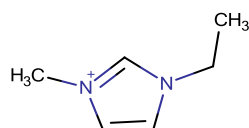
test system	duration of test/exposure	ref
<i>Aliivibrio fischeri</i>	30 min (DIN)	(Matzke et al., 2007)

	38412-L34.58)	
<i>E. coli</i>	8 h	(Lee et al., 2005)
<i>Pseudokirchneriella subcapitata</i>	72 h (OECD 201)	(Wells and Coombe, 2006)
<i>Scenedesmus vacuolatus</i>	24 h	(Matzke et al., 2007)
IPC-81	4 h	(Matzke et al., 2007)
HeLa	24 to 48 h	(Matzke et al., 2007)
MCF7	24 h	(Kumar et al., 2009)
<i>Lemna Minor</i>	7 days	(Jastorff et al., 2005)
<i>Daphnia magna</i>	48 h (OECD 202)	(Wells and Coombe, 2006)

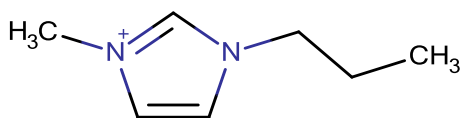
83

84 **SI-3, Table 8: Structures of the IL chemicals**

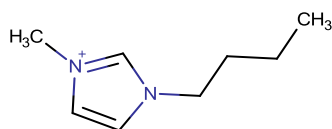
IM12 cation



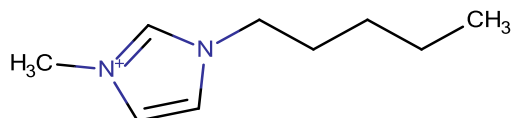
IM13 cation



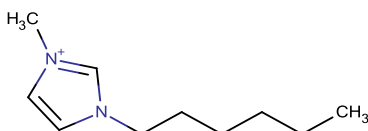
IM14 cation



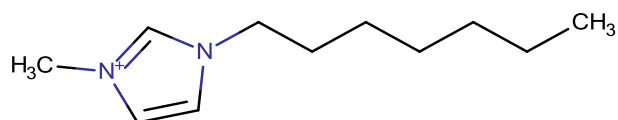
IM15 cation



IM16 cation

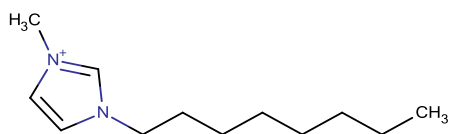


IM17 cation

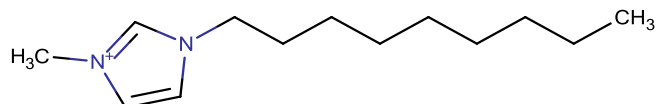




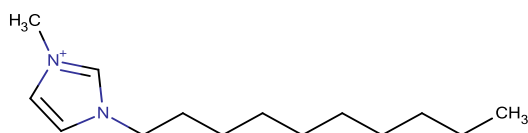
IM18 cation



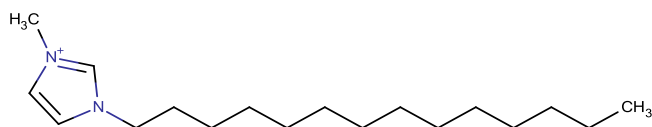
IM19 cation



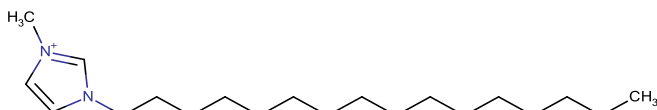
IM1-10 cation



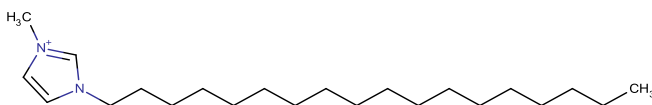
IM1-14 cation



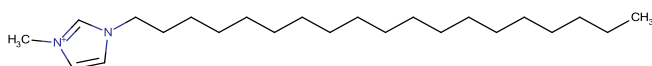
IM1-16 cation



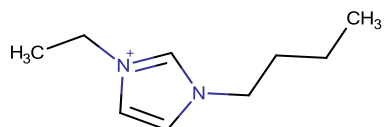
IM1-18 cation



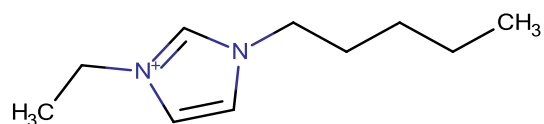
IM1-19 cation



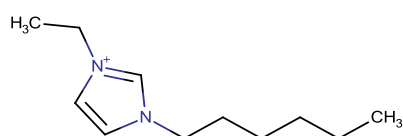
IM24 cation



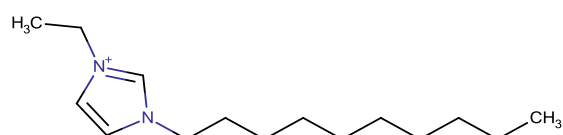
IM25 cation



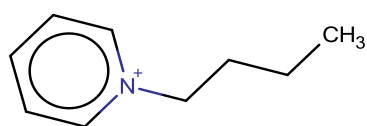
IM26 cation



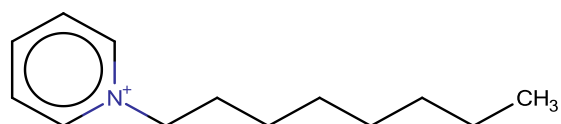
IM2-10 cation



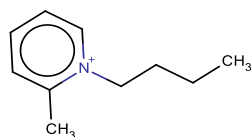
Py4 cation



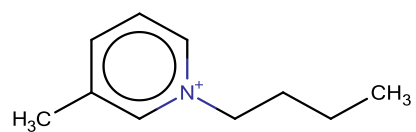
Py8 cation



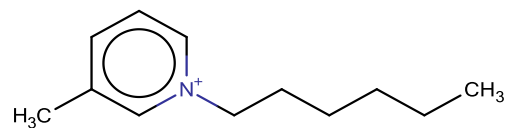
Py4-2Me cation



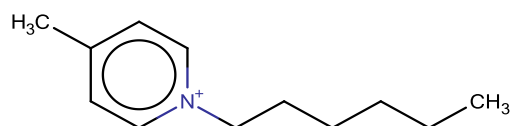
Py4-3Me cation



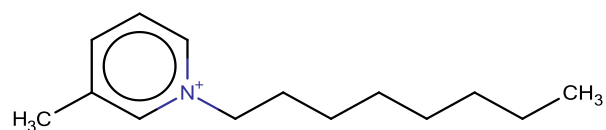
Py6-3Me cation



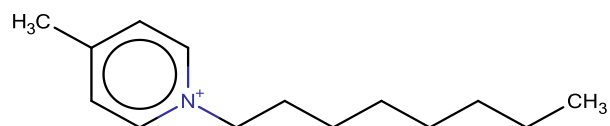
Py6-4Me cation



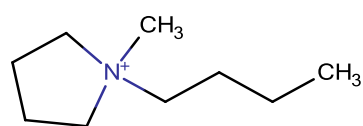
Py8-3Me cation



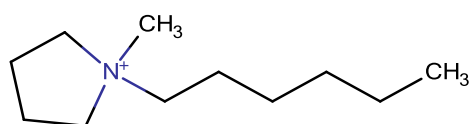
Py8-4Me cation



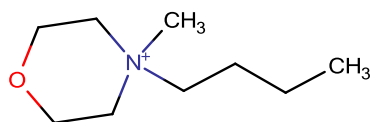
Pyr14 cation



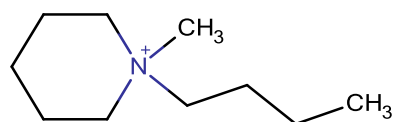
Pyr16 cation



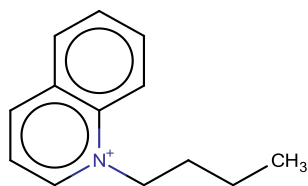
Mor14 cation



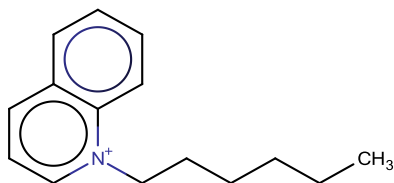
Pip14 cation



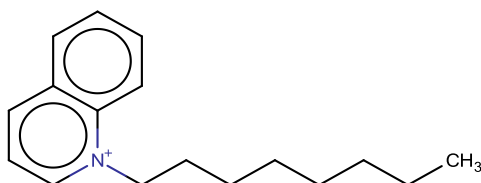
Quin4 cation



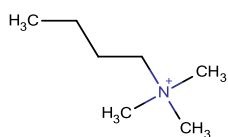
Quin6 cation



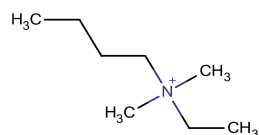
Quin8 cation



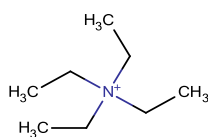
N1114 cation



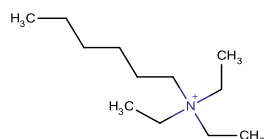
N1124 cation



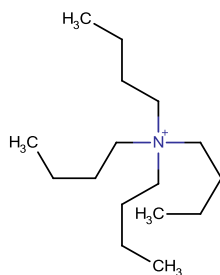
N2222 cation



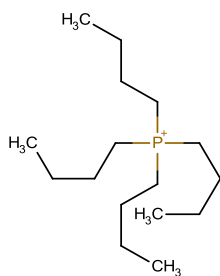
N2226 cation



N4444 cation



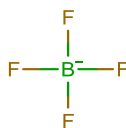
P4444 cation



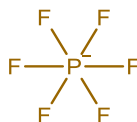
Cl anion



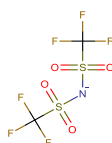
BF4 anion



PF6 anion



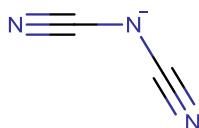
(CF3SO2)2N anion



Br anion



(CN)2N anion



85 SI-3, Table 9: Analysis of the toxic ratios (TRs) according to Eq. 14 of the summary, section 1.5.

A.	E.	P.	S.	IPC-81	HeLa	MCF7	L.	D.		
fischeri	coli	subcapitata	vacuolatus				Minor	magna	cation	anion
2.3			105.9						IM12	Cl
	0.022			1.4	0.40					BF4
				0.021						PF6
					0.055					(CF3SO2)2N
0.45				1.3					IM13	BF4
16.2		172.4	207.3	10.6			57.1	443.2	IM14	Cl
11.8		10.9		11.7	11.4			688.6		Br
2.1	0.10		29.5	2.9	0.8		12.3	80.3		BF4
0.14	0.012	1.1		0.14	0.012			2.4		PF6
0.041	0.28	1.6	1.6	0.21	0.085		0.36			(CF3SO2)2N
2.7				4.0						(CN)2N
2.5									IM15	BF4
23.3		476005.0	4760.1	8.1					IM16	Cl
426.3		15.0						667.9		Br
1.6										BF4
1.1				0.17						PF6
				0.12						(CF3SO2)2N
				5.2					IM17	Cl
4.5				3.3						BF4

			0.78				PF6
40.2	14606.0	236881.4	4.9			IM18	Cl
225.6	11.3			1.7	1254.2		Br
17.6		90231.4	11.6	1.5			BF4
14.3	0.29		1.4				PF6
			1.9	0.44			(CF3SO2)2N
24.5						IM19	BF4
27.0		136917.7	1.7			IM1-10	Cl
55.3			6.2				BF4
			1.0				PF6
0.25		52.8	0.46			IM1-14	Cl
0.01			0.023			IM1-16	Cl
0.000031			0.00085			IM1-18	Cl
			0.0000086			IM1-19	Cl
			0.0000049				BF4
			0.0000031				PF6
5.7			2.0	0.16		IM24	BF4
2.2						IM25	BF4
			41.0			IM26	Br
15.1			11.7				BF4
			6.3			IM2-10	Br
37.4	120.2	114.8			213.7	Py4	Cl

31.1		4.6	11.4			Br
		2.5				BF4
6.4						(CN)2N
		38.2				Py8 Cl
		2.2				Py4- 2Me BF4
145.5						Py4- 3Me Br
		1.9				BF4
26.8						N(CN)2
235.3						Py6- 3Me Cl
106.4					1610.9	Br
		17.6				Py6- 4Me BF4
189.7						Py8- 3Me Br
				62.8	250.0	Py8- 4Me Cl
		20.3				BF4
		21.8			353.5	Pyr14 Cl
	8.6	6.8	1.4			Br



		4.9					BF4
	0.30	0.10	0.073	0.11			(CF3SO2)2N
		0.35					N(CN)2
12.6		15.1				Pyr16	Cl
				33.6		Mor14	Br
	1.0	0.037		0.071			(CF3SO2)2N
1.9	18.7	3.3	2.5	11801.5		Pip14	Br
	0.83	0.039	0.12	0.14			(CF3SO2)2N
		89.3				Quin4	Br
		24.3					BF4
		145.1				Quin6	BF4
		201.3				Quin8	Br
		121.7					BF4
		0.025				N1114	(CF3SO2)2N
			6874.6			N1124	Cl
	1.7	0.037					(CF3SO2)2N
			2.7			N2222	Br
19.6						N2226	Br
0.046		0.48				N4444	Br
0.078		0.88			4.5	P4444	Br

## Appendix Bibliography

- Abraham, M.H., Acree, Jr, W.E., 2010a. Equations for the Transfer of Neutral Molecules and Ionic Species from Water to Organic phases. *J. Org. Chem.* 75, 1006–1015. doi:10.1021/jo902388n
- Abraham, M.H., Acree, Jr, W.E., 2010b. Solute Descriptors for Phenoxide Anions and Their Use To Establish Correlations of Rates of Reaction of Anions with Iodomethane. *J. Org. Chem.* 75, 3021–3026. doi:10.1021/jo100292j
- Abraham, M.H., Acree, Jr, W.E., 2010c. The transfer of neutral molecules, ions and ionic species from water to ethylene glycol and to propylene carbonate; descriptors for pyridinium cations. *New J. Chem.* 34, 2298. doi:10.1039/c0nj00222d
- Austin, R.P., Andrew, D.M., Carol, M.N., 1995. Partitioning of ionizing molecules between aqueous buffers and phospholipid vesicles. *J. Pharm. Sci.* 84, 1180–1183.
- Avdeef, A., Box, K.J., Comer, J.E.A., Hibbert, C., Tam, K., 1998. pH-Metric logP 10. Determination of liposomal membrane-water partition coefficients of ionizable drugs. *Pharm. Res.* 15, 209–215. doi:10.1023/A:1011954332221
- Barzanti, C., Evans, R., Fouquet, J., Gouzin, L., Howarth, N.M., Kean, G., Levet, E., Wang, D., Wayemberg, E., Yeboah, A. a., Kraft, A., 2007. Potentiometric determination of octanol–water and liposome–water partition coefficients (logP) of ionizable organic compounds. *Tetrahedron Lett.* 48, 3337–3341. doi:10.1016/j.tetlet.2007.03.085
- Becke, A.D., 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* 38, 3098–3100. doi:10.1103/PhysRevA.38.3098
- Betageri, G.V., Rogers, J.A., 1987. Thermodynamics of partitioning of  $\beta$ -blockers in the n-octanol- buffer and liposome systems. *Int. J. Pharm.* 36, 165–173. doi:10.1016/0378-5173(87)90152-9
- Bittermann, K., Spycher, S., Endo, S., Pohler, L., Huniar, U., Goss, K.-U., Klamt, A., 2014. Prediction of Phospholipid–Water Partition Coefficients of Ionic Organic Chemicals Using the Mechanistic Model COSMOmic. *J. Phys. Chem. B* 118, 14833–42. doi:10.1021/jp509348a
- Brooks, B.W., Foran, C.M., Richards, S.M., Weston, J., Turner, P.K., Stanley, J.K., Solomon, K.R., Slattery, M., La Point, T.W., 2003. Aquatic ecotoxicology of fluoxetine. *Toxicol. Lett.* 142, 169–183. doi:10.1016/S0378-4274(03)00066-3
- Cohen, F.S., Eisenberg, M., McLaughlin, S., 1977. The kinetic mechanism of action of an

- uncoupler of oxidative phosphorylation. *J. Membr. Biol.* 37, 361–96.
- Demura, M., Kamo, N., Kobatake, Y., 1987. Binding of lipophilic cations to the liposomal membrane: thermodynamic analysis. *Biochim. Biophys. Acta - Biomembr.* 903, 303–308. doi:10.1016/0005-2736(87)90220-3
- Dilger, J., McLaughlin, S., 1979. Proton transport through membranes induced by weak acids: A study of two substituted benzimidazoles. *J. Membr. Biol.* 46, 359–384. doi:10.1007/BF01868755
- Eckert, F., Klamt, A., 2002. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* 48, 369–385. doi:10.1002/aic.690480220
- Eichkorn, K., Treutler, O., Öhm, H., Häser, M., Ahlrichs, R., 1995. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* 240, 283–290. doi:10.1016/0009-2614(95)00621-A
- Endo, S., Escher, B.I., Goss, K.-U., 2011. Capacities of Membrane Lipids to Accumulate Neutral Organic Chemicals. *Environ. Sci. Technol.* 45, 5912–5921. doi:10.1021/es200855w
- Escher, B.I., Berger, C., Bramaz, N., Kwon, J.-H., Richter, M., Tsinman, O., Avdeef, A., 2008. Membrane-water partitioning, membrane permeability, and baseline toxicity of the parasiticides ivermectin, albendazole, and morantel. *Environ. Toxicol. Chem.* 27, 909–918. doi:10.1897/07-427.1
- Escher, B.I., Bramaz, N., Richter, M., Lienert, J., 2006. Comparative ecotoxicological hazard assessment of beta-blockers and their human metabolites using a mode-of-action-based test battery and a QSAR approach. *Environ. Sci. Technol.* 40, 7402–8.
- Escher, B.I., Hunziker, R.W., Schwarzenbach, R.P., 2001. Interaction of phenolic uncouplers in binary mixtures: concentration-additive and synergistic effects. *Environ. Sci. Technol.* 35, 3905–3914.
- Escher, B.I., Schwarzenbach, R.P., 1996. Partitioning of Substituted Phenols in Liposome–Water, Biomembrane–Water, and Octanol–Water Systems. *Environ. Sci. Technol.* 30, 260–270. doi:10.1021/es9503084
- Escher, B.I., Schwarzenbach, R.P., Westall, J.C., 2000. Evaluation of Liposome–Water Partitioning of Organic Acids and Bases. 1. Development of a Sorption Model. *Environ. Sci. Technol.* 34, 3954–3961. doi:10.1021/es0010709
- Escher, B.I., Sigg, L., 2004. Chemical Speciation of Organics and of Metals at Biological Interphases, in: v. Leeuwen, H.P., Köster, W. (Eds.), *Physicochemical Kinetics and*

- Transport at Biointerfaces. John Wiley & Sons, Ltd, Chichester, UK, pp. 205–269. doi:10.1002/0470094044.ch5
- Flewelling, R.F., Hubbell, W.L., 1986. Hydrophobic ion interactions with membranes. Thermodynamic analysis of tetraphenylphosphonium binding to vesicles. *Biophys. J.* 49, 531–540. doi:10.1016/S0006-3495(86)83663-3
- Fruttero, R., Caron, G., Fornatto, E., Boschi, D., Ermondi, G., Gasco, A., Carrupt, P.A., Testa, B., 1998. Mechanisms of liposomes/water partitioning of (p- methylbenzyl)alkylamines. *Pharm. Res.* 15, 1407–1413. doi:10.1023/A:1011953622052
- Inoue, T., Iwanaga, T., Fukushima, K., Shimozawa, R., 1988. Effect of sodium octanoate and sodium perfluorooctanoate on gel-to-liquid-crystalline phase transition of dipalmitoylphosphatidylcholine vesicle membrane. *Chem. Phys. Lipids* 46, 25–30. doi:10.1016/0009-3084(88)90109-0
- Inoue, T., Miyakawa, K., Shimozawa, R., 1986. Interaction of surfactants with vesicle membrane of dipalmitoylphosphatidylcholine. Effect on gel-to-liquid-crystalline phase transition of lipid bilayer. *Chem. Phys. Lipids* 42, 261–270. doi:10.1016/0009-3084(86)90085-X
- Jakobtorweihen, S., Ingram, T., Smirnova, I., 2013. Combination of COSMOmic and molecular dynamics simulations for the calculation of membrane-water partition coefficients. *J. Comput. Chem.* 34, 1332–1340. doi:10.1002/jcc.23262
- Jastorff, B., Mölter, K., Behrend, P., Bottin-Weber, U., Filser, J., Heimers, A., Ondruschka, B., Ranke, J., Schaefer, M., Schröder, H., Stark, A., Stepnowski, P., Stock, F., Störmann, R., Stolte, S., Welz-Biermann, U., Ziegert, S., Thöming, J., 2005. Progress in evaluation of risk potential of ionic liquids—basis for an eco-design of sustainable products. *Green Chem.* 7, 362. doi:10.1039/b418518h
- Kaiser, S., Escher, B.I., 2006. The evaluation of liposome-water partitioning of 8-hydroxyquinolines and their copper complexes. *Environ. Sci. Technol.* 40, 1784–1791.
- Kasianowicz, J., Benz, R., McLaughlin, S., 1987. How do Protons Cross the Membrane-Solution Interface? Kinetic Studies on Bilayer Membranes Exposed to the Protonophore S-13 (5-chloro-3-tert-butyl-2'-chloro-4'-nitrosalicylanilide). *J. Membr. Biol.* 95, 73–89.
- Kipka, U., Di Toro, D.M., 2009. Technical basis for polar and nonpolar narcotic chemicals and polycyclic aromatic hydrocarbon criteria. III. A polyparameter model for target lipid partitioning. 28, 1429–1438. doi:10.1897/08-364.1
- Klamt, A., 2005. COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics

- and Drug Design, First Edit. ed. Elsevier Science, Amsterdam.
- Klamt, A., 1995. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* 99, 2224–2235. doi:10.1021/j100007a062
- Klamt, A., Eckert, F., Arlt, W., 2010. COSMO-RS: an alternative to simulation for calculating thermodynamic properties of liquid mixtures. *Annu. Rev. Chem. Biomol. Eng.* 1, 101–22. doi:10.1146/annurev-chembioeng-073009-100903
- Klamt, A., Huniar, U., Spycher, S., Keldenich, J., 2008. COSMOmic: A Mechanistic Approach to the Calculation of Membrane–Water Partition Coefficients and Internal Distributions within Membranes and Micelles. *J. Phys. Chem. B* 112, 12148–12157. doi:10.1021/jp801736k
- Kumar, R.A., Papaiconomou, N., Lee, J.-M., Salminen, J., Clark, D.S., Prausnitz, J.M., 2009. In vitro cytotoxicities of ionic liquids: Effect of cation rings, functional groups, and anions. *Environ. Toxicol.* 24, 388–395. doi:10.1002/tox.20443
- Lee, S.-M., Chang, W.-J., Choi, A.-R., Koo, Y.-M., 2005. Influence of ionic liquids on the growth of *Escherichia coli*. *Korean J. Chem. Eng.* 22, 687–690. doi:10.1007/BF02705783
- Lehmler, H., Xie, W., Bothun, G., Bummer, P.M., Knutson, B.L., 2006. Mixing of perfluorooctanesulfonic acid (PFOS) potassium salt with dipalmitoyl phosphatidylcholine (DPPC). *Colloids Surfaces B Biointerfaces* 51, 25–29. doi:10.1016/j.colsurfb.2006.05.013
- Lombardi, D., Cuenoud, B., Krämer, S.D., 2009. Lipid membrane interactions of indacaterol and salmeterol: do they influence their pharmacological properties? *Eur. J. Pharm. Sci.* 38, 533–47. doi:10.1016/j.ejps.2009.10.001
- Matzke, M., Stolte, S., Thiele, K., Juffernholz, T., Arning, J., Ranke, J., Welz-Biermann, U., Jastorff, B., 2007. The influence of anion species on the toxicity of 1-alkyl-3-methylimidazolium ionic liquids observed in an (eco)toxicological test battery. *Green Chem.* 9, 1198. doi:10.1039/b705795d
- Miyoshi, H., Nishioka, T., Fujita, T., 1987. Quantitative relationship between protonophoric and uncoupling activities of substituted phenols. *Biochim. Biophys. Acta* 891, 194–204.
- Nakamura, Y., Yamamoto, H., Sekizawa, J., Kondo, T., Hirai, N., Tatarazako, N., 2008. The effects of pH on fluoxetine in Japanese medaka (*Oryzias latipes*): acute toxicity in fish larvae and bioaccumulation in juvenile fish. *Chemosphere* 70, 865–73.

- doi:10.1016/j.chemosphere.2007.06.089
- Neuwoehner, J., Fenner, K., Escher, B.I., 2009. Physiological Modes of Action of Fluoxetine and its Human Metabolites in Algae. *Environ. Sci. Technol.* 43, 6830–6837. doi:10.1021/es9005493
- Ottiger, C., Wunderli-Allenspach, H., 1997. Partition behaviour of acids and bases in a phosphatidylcholine liposome–buffer equilibrium dialysis system. *Eur. J. Pharm. Sci.* 5, 223–231. doi:10.1016/S0928-0987(97)00278-9
- Pallicer, J.M., Krämer, S.D., 2012. Evaluation of fluorescence anisotropy to assess drug–lipid membrane partitioning. *J. Pharm. Biomed. Anal.* 71, 219–227. doi:10.1016/j.jpba.2012.08.009
- Perdew, J., 1986. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* 33, 8822–8824. doi:10.1103/PhysRevB.33.8822
- Plempers van Balen, G., Caron, G., Ermondi, G., Pagliara, A., Grandi, T., Bouchard, G., Fruttero, R., Carrupt, P.-A., Testa, B., 2001. Lipophilicity Behaviour of the Zwitterionic Antihistamine Cetirizine in Phosphatidylcholine Liposomes/Water Systems. *Pharm. Res.* 18, 694–701. doi:10.1023/A:1011049830615
- Sadowski, J., Gasteiger, J., Klebe, G., 1994. Comparison of Automatic Three-Dimensional Model Builders Using 639 X-ray Structures. *J. Chem. Inf. Model.* 34, 1000–1008. doi:10.1021/ci00020a039
- Schäfer, A., Huber, C., Ahlrichs, R., 1994. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* 100, 5829. doi:10.1063/1.467146
- Schellenberg, K., Leuenberger, C., Schwarzenbach, R.P., 1984. Sorption of chlorinated phenols by natural sediments and aquifer materials. *Environ. Sci. Technol.* 18, 652–657. doi:10.1021/es00127a005
- Schwarzenbach, R.P., Gschwend, P., Imboden, D., 2003. *Environmental organic chemistry*, 2nd editio. ed, Journal of Chemical Education. John Wiley & Sons, Hoboken, New Jersey.
- Schwarzenbach, R.P., Stierli, R., Folsom, B.R., Zeyer, J., 1988. Compound properties relevant for assessing the environmental partitioning of nitrophenols. *Environ. Sci. Technol.* 22, 83–92. doi:10.1021/es00166a009
- Schweigert, N., Hunziker, R.W., Escher, B.I., Eggen, R.I., 2001. Acute toxicity of (chloro-

- )catechols and (chloro-)catechol-copper combinations in *Escherichia coli* corresponds to their membrane toxicity in vitro. *Environ. Toxicol. Chem.* 20, 239–247.
- Smejtek, P., Blochel, A., Wang, S., 1996. Hydrophobicity and sorption of chlorophenolates to lipid membranes. *Chemosphere* 33, 177–201. doi:10.1016/0045-6535(96)00158-0
- Spycher, S., Smejtek, P., Netzeva, T.I., Escher, B.I., 2008. Toward a Class-Independent Quantitative Structure–Activity Relationship Model for Uncouplers of Oxidative Phosphorylation. *Chem. Res. Toxicol.* 21, 911–927. doi:10.1021/tx700391f
- Stolte, S., Arning, J., Bottin-Weber, U., Müller, A., Pitner, W.-R., Welz-Biermann, U., Jastorff, B., Ranke, J., 2007. Effects of different head groups and functionalised side chains on the cytotoxicity of ionic liquids. *Green Chem.* 9, 760–767. doi:10.1039/B615326G
- Thomae, A. V., Koch, T., Panse, C., Wunderli-Allenspach, H., Krämer, S.D., 2007. Comparing the lipid membrane affinity and permeation of drug-like acids: the intriguing effects of cholesterol and charged lipids. *Pharm. Res.* 24, 1457–1472. doi:10.1007/s11095-007-9263-y
- Thomae, A. V., Wunderli-Allenspach, H., Krämer, S.D., 2005. Permeation of aromatic carboxylic acids across lipid bilayers: the pH-partition hypothesis revisited. *Biophys. J.* 89, 1802–11. doi:10.1529/biophysj.105.060871
- Vaes, W.H.J., Ramos, E.U., Verhaar, H.J.M., Hermens, J.L.M., 1998. Acute toxicity of nonpolar versus polar narcosis: Is there a difference? *Environ. Toxicol. Chem.* 17, 1380–1384. doi:10.1002/etc.5620170723
- Vainio, M.J., Johnson, M.S., 2007. Generating Conformer Ensembles Using a Multiobjective Genetic Algorithm. *J. Chem. Inf. Model.* 47, 2462–2474. doi:10.1021/ci6005646
- Wells, A.S., Coombe, V.T., 2006. On the freshwater ecotoxicity and biodegradation properties of some common ionic liquids. *Org. Process Res. Dev.* 10, 794–798. doi:10.1021/op060048i
- Yamamoto, H., Hayashi, A., Nakamura, Y., Sekizawa, J., 2005. Fate and partitioning of selected pharmaceuticals in aquatic environment. *Environ. Sci.* 12, 347–58.