



Predictions of flavonoid solubility in ionic liquids by COSMO-RS: experimental verification, structural elucidation, and solvation characterization

Guo, Zheng; Lue, Bena-Marie; Thomsen, Kaj; Meyer, Anne S.; Xu, Xuebing

Published in:
Green Chemistry

Link to article, DOI:
[10.1039/b709786g](https://doi.org/10.1039/b709786g)

Publication date:
2007

[Link back to DTU Orbit](#)

Citation (APA):

Guo, Z., Lue, B.-M., Thomsen, K., Meyer, A. S., & Xu, X. (2007). Predictions of flavonoid solubility in ionic liquids by COSMO-RS: experimental verification, structural elucidation, and solvation characterization. *Green Chemistry*, 9(12), 1362-1373. <https://doi.org/10.1039/b709786g>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Predictions of flavonoid solubility in ionic liquids by COSMO- RS: experimental verification, structural elucidation, and solvation characterization

Zheng Guo¹, Bena-Marie Lue¹, Kaj Thomasen², Anne S. Meyer² and Xuebing Xu^{1,*}

¹ BioCentrum-DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark; ² Department
of Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark.

Running title: Prediction and verification of the solubilities of flavonoids in ionic liquids

* Correspondence:

Xuebing Xu, BioCentrum-DTU, Technical University of Denmark, Building 222, DK-2800 Lyngby,
Denmark

Tel. +45-45252773

Fax. +45-45884922

E-mail: xx@biocentrum.dtu.dk

Abstract

2 Predictions of the solubility of flavonoids in a large variety of ionic liquids (ILs) with over 1800
available structures were examined based on COSMO-RS computation. The results show that the
4 solubilities of flavonoids are strongly anion-dependent. Experimental measurement of the
solubilities of esculin and rutin in 12 ILs with varying anions and cations show that predicted and
6 experimental results generally have a good agreement. Based on the sound physical basis of
COSMO-RS, the solubility change of flavonoids were quantitatively associated with solvation
8 interactions and structural characteristics of ILs. COSMO-RS derived parameters, i. e. misfit, H-
bonding and van der Waals interaction energy, are shown to be capable of characterizing the
10 complicated multiple interactions in IL system effectively. H-bonding interaction is the most
dominant interaction for ILs (followed by misfit and van der Waals interactions) to determine the
12 solubility of flavonoids, and the anionic part has greater effect on the overall H-bonding capability
of the IL. Based on basicity of anions, ILs were categorized into 3 groups, corresponding to the
14 classification of the solubility of flavonoid. While COSMO sigma-moment descriptors, which
roughly denote the characteristic properties of the ILs, might be of general value to have a fast
16 estimation for the solubilities of flavonoids as well as those compounds with massive moieties as H-
bonding donor. **The results obtained in this work may be important for achieving an improved**
18 **understanding of IL solvations and** the tailoring of the desired structures of ILs used as the media
for efficient enzymatic esterification of flavonoids.

20 **Key words:** Ionic liquids (ILs), COSMO-RS (Conductor-like Screening Model for Real Solvent),
flavonoid, esculin, prediction, solvation and enzymatic esterification.

Introduction

2 The effectiveness and absorption of many drugs are largely controlled by their low solubility,
4 therefore, modification of drugs and producing so-called prodrug is a useful method to obtain
improved properties ¹⁻³. Flavonoids are such a group of prominent molecules with multiple
6 physiological activities. However, their functions are limited by their low aqua- and lipo- solubility
and resultant low bioavailability ⁴⁻⁶. Lipophilisation of flavonoids into fatty acid esters has been
8 proven to be an efficient way to expand functionalities and applications in human nutrition ^{5,7}. The
presence of solvents is essential in all steps of pharmaceutical processes (reaction, separation and
10 formulation) ⁸. The particular importance of a good choice of solvents for a reaction results from the
fact that the medium often affects the overall reaction rate, selectivity or yield ^{9,10}. However, the
12 attempts to establish an efficient enzymatic esterification of flavonoids with fatty acids using
conventional solvents have been seriously upset by their low solubility ^{7,11}. As neoteric “green”
14 solvents, the unique properties and tunable physical and chemical characteristics of ionic liquids
(ILs) guide one to resort to the novel media for a better solution^{12,13}.

16 The characteristics and the state of art of ILs in technology development and applications have
been described in a few excellent reviews ^{14,15}. Among those characteristics of ILs distinguishing
18 them from conventional solvents, the tailorability of properties by judicious selection of cation,
anion and substituents and an extended family of available structures for selection constitutes the
20 most interesting and attractive features of ionic liquids ^{13,16}. Actually, from engineering point of
view, it is the above distinct features that qualify ILs as “a designer solvent” and offer a huge
22 potential for practical applications ^{17,13}. However, screening a desired structure with required
properties for a particular task from a large pool of available ILs represents a big challenge facing
24 chemists and engineers ^{13,18,19}. To establish an efficient IL-based enzymatic reaction system to
product lipophilic derivatives of flavonoids, there exist similar obstacles for the setup of such a

reaction system to fulfill the requirement of high productivity and simultaneously being benign for enzyme activity²⁰. Solubility estimation approaches could be quite valuable in reducing time and resources required to identify a good solvent^{8,9,21}. Particularly for numerous possible ILs, a priori screening is needed, because it is impractical to use trial and error methods to find a suitable IL from enormous number of ILs structurally possible for a given function. As reviewed elsewhere^{8,9}, Hansen solubility parameters, group contribution method (UNIFAC) and quantitative structure-property relationship (QSPR) method etc. have been employed for solvent selection for solid solutes. These methods generally need property data experimentally available to create reliable training sets, which are currently scarce for ILs, as a group of relatively new compounds^{9,22}. As a physically well-founded computation approach and independent of experimental data, COSMO-RS (Conductor-like Screening Model for Real Solvent) provides a different and feasible alternative for solubility estimation in ILs^{23,24}.

As a physically founded model, COSMO-RS integrates dominant interactions (electrostatic (polarity), H-bonding and van der Waals (dispersion)) among IL systems, which adequately summarize multiple solvation interactions of ILs^{23,24}. Therefore, this method is able, at least qualitatively, to describe structural variations correctly. Previous publications also have demonstrated the general applicability of COSMO-RS concerning the prediction of solubilities of solids and liquids in ILs^{25,8}. Importantly, COSMO-RS introduces the vivid concept of σ -profiles for a qualitative and quantitative comparison of the polarity distribution on molecular surfaces, which integrate the description of electrostatic, hydrogen-bonding and hydrophobicity of a structure. The derived interaction items from COSMO-RS calculation can be conversely used for molecular force field analysis to correlate interaction parameters with structural characteristics qualitatively and quantitatively, which is of general instruction value in the design and development of ILs for specific tasks^{20,26,27}. Thus, besides presenting the predicted results of flavonoids in ILs and

experimental validation of the COSMO-RS prediction, this work gives more attention on the analysis and assortment of interactions resulted from the specific environment of the ILs applied. These efforts are expected to contribute to an improved understanding of structure-functionality relationship of ILs and serve later structural optimization of the ionic liquids possibly with high solubilities of flavonoids and benign to enzyme activity as well.

6 Theoretical basis of COSMO-RS for solubility calculation

Solubility denotes the solute concentration in a solution that is in thermodynamic equilibrium with the solute in the solid state. Therefore, the solubility depends on the difference of the chemical potentials of the solute in the solvent and in pure solute. For a solid solute, the energy change of a compound from the “supercooled” liquid state to the ordered solid state has to be taken into account. The solubility (x_i) at temperature T is thus expressed as a function of pure component properties of the solute (the melting temperature T_m and the free energy difference ($\mu_i^i - \mu_i^j$) (solid state related to its liquid state) calculated from heat of fusion (ΔH_m^{fus}) and the heat capacity difference of solute in melt state and “hypothetical” “supercooled” melt state) ($C_p^m - C_p^s$) and of the interactions between the solute and solvent in the solution (the activity coefficients, $\ln \gamma_i$)²⁸:

$$\ln x_i = \frac{1}{RT} \cdot (\mu_i^i - \mu_i^j) - \ln \gamma_i = \frac{1}{R} \cdot \left[\Delta H_m^{fus} \left(\frac{1}{T_m} - \frac{1}{T} \right) + \int_{T_m}^T \frac{(C_p^m - C_p^s) dT}{T^2} \right] - \ln \gamma_i \quad (1)$$

where R is gas constant.

The theory of COSMO-RS has been described by Klamt and coworkers^{23, 24}. Briefly, COSMO-RS is a statistical thermodynamics approach based on the results of quantum chemical-COSMO calculations. Starting from the surface polarization charge densities σ from density functional theory (DFT) COSMO calculations, COSMO-RS considers all interactions, especially electrostatic interaction and H-bonding, in a liquid system as contact interactions of the molecular

surfaces, which are written as pair interactions of the respective polarization change densities σ and σ' of contacting surface. Then, the chemical potential of a surface segment with SCD (screening charge density) σ in an ensemble can be described by the normalized distribution function $p_S(\sigma')$ given by Eqn.(2).

$$\mu_S(\sigma) = -\frac{RT}{a_{eff}} \ln \left[\int p_S(\sigma') \exp \left\{ \frac{a_{eff}}{RT} [\mu_S(\sigma') - E_{misfit}(\sigma, \sigma') - E_{HB}(\sigma, \sigma')] \right\} d\sigma' \right] \quad (2)$$

where $\mu_S(\sigma)$ is a measure for the affinity of the system S to a surface polarity σ ; E_{misfit} represents the electrostatic contact interaction energy; E_{HB} represents the energy contribution from H-bonding interaction; and a_{eff} is the effective contact area between two surface segments. Not being a function of individual surface contacts, E_{vdw} is not included in Eqn. (2) but added to the reference energy in solution *a posteriori*. The chemical potential of compound X_i in system S is then available from integration of the σ -potential over the surface of X_i . The capability of COSMO-RS to predict the chemical potential μ_S^X of any solute X in any pure or mixed solvent S at variable temperature T enables the calculation of any thermodynamic liquid-liquid equilibrium, which also constitutes the basis of COSMO-RS for solubility estimation. In the current COSMOtherm program, the free energy of fusion of a solid solute is estimated via a QSPR approach and approximated from following COSMOtherm descriptors²⁵:

$$-\Delta G_{fus}^X = c_1 \mu_i^{(H_2O)} + c_2 N_i^{Ring} + c_3 V_i + c_4 \quad (3)$$

where c_1 to c_4 is the QSPR parameters. $\mu_i^{(H_2O)}$ is the chemical potential of solute i in water, N_i^{Ring} is the number of ring atoms in compound i and V_i is the molecular volume of the solute. Crystal structure prediction for drugs has to be considered as an unsolved problem. In Eqn. (3), COSMO-RS treats the quantity size, rigidity, polarity and number of H-bonds as plausible driving forces of crystallization. Molecular size is described by V_i , which is available in the framework of COSMO-RS. The molecular rigidity of drugs is largely resulted from ring structures and therefore the number

of ring atoms N_i^{Ring} is used as descriptor of rigidity. $\mu_i^{(H_2O)}$ is a combined measure of polarity and H-bonding, which can be estimated by COSMO $therm$. Therefore, the chemical potential of a compound in pure water is of special importance in the computation of solubility in COSMO-RS. The parameterizations of the QSPR parameters for fully relaxed Turbomole DFT / COSMO calculations with the larger TZVP basis set (namely BP_TZVP_C21_0106.ctd) used in this calculation includes solubility parameters that derived from a set of solubility data of 150 aqueous solubilities²⁵.

8 Computational details and calculation sets of flavonoids and ionic liquids

The molecular structures of flavonoids (rutin, esculin, quercetin, isoquercetrin and naringin, etc.) (Scheme 1) were sketched as two-dimensional structures and subsequently converted to three-dimensional geometries using Chemdraw Ultra 8.0. To obtain the lowest energy conformations for each flavonoid, special care has been taken into consideration in choosing cis-trans and conformational isomerism of the sugar ring structures that are able to build internal hydrogen bonds. The energy of molecular conformations was minimized by MOPAC (Molecular Orbital PACKage) 2000²⁶. Using the geometries thus optimized, the computation of the COSMO polarization charge densities σ of the molecular surfaces were performed with the TURBOMOLE 5.7 program package on the density functional theory level, utilizing the BP (B88-VWN-P86) functional with a triple- ζ valence polarized basis set (TZVP)²⁵. Most of the COSMO files of the cations and anions involved in this work adopt the provision from a latest database of BP-COSMO-IL (COSMO/logic GmbH & Co KG, Leverkusen, Germany). The COSMO files of the cations and anions of ILs excluded in this database but used in this work were generated following a procedure similar to the one we used for flavonoid molecules. All solubility calculations of flavonoids are performed using

1 COSMOthermX_2.1 program (COSMOlogic GmbH & Co KG, Leverkusen, Germany). For
2 flavonoid inputs only the conformation lowest in energy was used, and for cations and anions of ILs
multiple conformers were input with activated conformer treatment in the automatic computation.

4 In all calculations of the solubilities of flavonoids in ILs, the cation and anion of an IL are
treated as separate molecules with equal molar fractions ($n_{cation} = n_{anion} = n_{IL}$). Thus, the
6 COSMOtherm calculation is based on a ternary mixture (cation, anion and solute), differing from an
experimental determination treated as a binary system consisting of IL and solute. Therefore, a
8 transition calculation from the solubility of a ternary system ($x_s^{ternary} = n_i / (n_i + 2 \cdot n_{ion})$) to the
solubility of a binary system ($x_s^{binary} = n_i / (n_i + n_{IL})$) was done to acquire the comparative datum
10 with the experimental value. The intensive interests of this work are not placed on the discussion of
methodology itself but devoted to the revelation of the dominant interactions among IL systems to
12 govern the solubility of flavonoids. For a better comparison of solubilities of flavonoids in different
types of ILs under the same conditions, all calculations (regardless of small or large solubility) were
14 performed with a non-iterative mode, which means the solubility computed is a zeroth order
approximation.

16 To achieve a comprehensive evaluation on the properties of ILs, the calculation set of cations
involved in this work covered most important types of possible cations, such as imidazolium,
18 pyridinium, pyrrolidinium, ammonium, phosphonium, Sulfonium, guanidinium, isoquinolinium,
and isouronium, etc (Scheme 2). To investigate the effects of the incorporated substituents on the
20 hydrophobic-hydrophilic property of ILs and the subsequent influence on the solubility of
flavonoids, varied chain length of alkyl groups were appended to the parent structures at different
22 positions, with total 59 structures. While the anions examined include the often used types (e.g. PF_6^- ,
 BF_4^- and tf_2N^-) and those uncommon groups^{29,30}, that have the properties varying from “non-polar”
24 to moderate and strong polar anions. The number of the anion types involved amounts to 32.

Therefore, the calculation set of ILs in this work is 1888 combinations, which almost covered all hitherto known important commercially available ILs (not included are those functionalized ILs developed for special task). **It should be mentioned that not all combinations lead to ionic liquids, at least some of them do not exist as liquid at room temperature. However, they are perceived as ionic liquids herein for the convenience of model processing.**

The common structural nature of flavonoids is poly-phenyl ring with saccharide substituents. Therefore, rutin with 2-phenylchromone parent structure (three phenolic rings) and disaccharide substituents and esculin with chromone parent structure (two phenolic rings) and monosaccharide substituents were selected as the representative structures for an extensive investigation of solubility (Scheme 1). Another reason for choosing these two flavonoids is that their lipophilic derivatives have previously been shown to have improved anti-oxidation properties⁷.

The average absolute error (AAE) was determined by Eqn (4)

$$AAE = \sum \left| \log S_{predict} - \log S_{exp} \right| / n \quad (4)$$

and root mean square deviations (RMSD) from Eqn (5)

$$RMSD = \left[\frac{1}{n} \sum \left(\left| \log S_{predict} - \log S_{exp} \right| - AAE \right)^2 \right]^{1/2} \quad (5)$$

where $\log S$ are the logarithms of the molar percentage of the predicted and experimental solubility (mol%) and n is the number of the compounds used.

Results and Discussion

COSMO-RS solubilities of esculin in ILs

The primary objective of this study is to acquire fundamental knowledge regarding what are the dominant interaction parameters that govern the solubilities of flavonoids in ILs and what are structural characteristics behind these interactions. Therefore, a large pool of ILs (59 cations \times 32 anions) was employed for solubility evaluation by COSMO-RS calculation. Fig. 1 shows the predicted solubilities of esculin in the ILs of different cations paired with differing anions (A) and in the ILs of different anions paired with varying cations (B). It is difficult to find any regularity in

the plotting of the solubilities of esculin versus cation alteration in Fig. 1A. For most of the ILs with same cation, the solubilities of esculin (logarithm of molar fraction) varied over a wide range (from 0 to 10^{-6}) with the change of the anion paired. However, if the same data are plotted as the function of anion variation, a different and interesting observation can be visualized (Fig. 1B). That is, in the same anion interval, apart from some exceptions (Examinations of the structures of ILs for the exception cases reveal that these ILs exist as solid at room temperature, which are perceived as liquids in the prediction), the group of ILs with same anion gave a narrow variation range of the solubilities of esculin. For instance, regardless of the structural variations of cations paired, acetate, deaconate and chlorion etc. based ILs generally have very high solubility of esculin (some are even theoretically mutual miscible according to COSMO-RS estimation). For PF_6^- based ILs, the solubilities of esculin varied within 10^{-3} - 10^{-5} (log x) and for tris(nonafluorobutyl)trifluorophosphate based ILs the solubilities are generally lower than 10^{-4} . The results suggested that the solubilities of esculin in ILs are determined, to more extent, by the anion rather than the cation part of the solvent IL. In another word, the solubilities of esculin in the solvent ILs with different anion and cation combinations are largely anion-dependent. To examine whether other flavonoids have a similar tendency, a similar computation of the solubilities of rutin in total 1888 combinations of ILs was carried out, and the solubilities of quercetin, isoquercetrin and naringin were calculated in part of the databases (data not shown). The results demonstrated that the anion-dependency of the solubilities of flavonoids in the solvent ILs with different cation-anion pairings is universal for rutin and other test flavonoids, which indicate that there are some intrinsic interactions between flavonoid molecules and solvent ILs deserving to be explored.

Based on the anion-dependent solubilities of esculin, the ILs examined could be classified into 3 groups (Fig. 1B). The first group dissolves flavonoids at very high concentration (log x, 0 ~ -1), in which the ILs with the anions of Cl^- , Br^- , decanoate, dimethylphosphate, dihydricphosphate, acetate,

trifluoroacetate, bis(2,4,4-trimethylpentyl)phosphinate, and toluene-4-sulfonate etc. The second group of the ILs have moderate solubilities ($\log x$, -1 ~ - 2.5) and with the anions of dicyanamide, bisbiphenyldiolatoborate, bis-pentafluoroethyl-phosphinate, bissalicylatoborate, ethoxyethylsulfate, methoxyethylsulfate, alkylsulfate (methyl-, ethyl-, butyl-, and octyl-), trifluoromethane-sulfonate, bis(trifluoromethyl)imide, bis(trifluoromethylsulfonyl)methane, etc. The third group of the ILs with the anions of BF_4^- , PF_6^- , bismalonatoborate, bisoxalatoborate, ClO_4^- , tetracyanoborate, tf_2N (bis(trifluoromethylsulfonyl)imide), tris(nonafluorobutyl)trifluorophosphate, tris(pentafluoroethyl)trifluorophosphate, etc., in which the solubilities of esculin ($\log x$) are generally less than -2 regardless of whatever cations paired (Fig. 1B). Interestingly, the above categorization of ILs for esculin is also seen to be validated for rutin with few exceptions according to our calculation (data not shown), indicating, in some way, a similar solvation behaviour between ILs with esculin and rutin. This categorization of ILs is apparently useful to quantify solute-solvent interactions and thus deserves to be further examined and investigated.

14 **Validation of COSMO-RS predictions**

To examine the accuracy of COSMO-RS predictions of the solubilities of flavonoids, 12 different types of ILs, representing different anion types and dissolution abilities as categorized above, were selected as solvents and esculin and rutin were chosen as model flavonoid molecules for evaluation (Fig. 2 and 3). Most of the ILs tested has higher viscosity, and 1-ethyl-3-methylimidazolium toluene-4-sulfonate (EMIM.OTs) is even solid at room temperature. Therefore, the measurements were conducted at 40 °C and 60 °C to accelerate dissolution. This test temperature is far from the boiling points of ILs and the evaporation of ILs can be neglected, thereby the measurements stay in a safe temperature range. The scatter plot of Fig. 2A shows a rather homogeneous error distribution of esculin solubilities at 40 °C, which means the solubilities of esculin in the ILs were not systematically overestimated or underestimated. The predicted and experimental values gave

average absolute error (AAE) of 0.29 log-unit and root mean square deviations (RMSD) 0.25 log-unit. These data suggested a better quality of esculin solubility prediction in ILs than a previous report concerning prediction of aqueous solubility of drugs and pesticides with COSMO-RS²⁵. The predictions of esculin solubility at the temperature of 60°C achieve the accuracy with the AAE of 0.39 and RMSD 0.22. The accuracy is comparable with that at 40 °C. No systemic deviation is observed, suggesting that 60 °C (the temperature range often used for lipase catalysis) is within the safety interval for COSMO-RS prediction.

Predicted and experimental data for the solubilities of rutin in 12 types of ILs at 40 and 60 °C are depicted in Fig. 3A and 3B, respectively. It is clear that for solvent BMIM.PF₆, BMIM.BF₄ and tOMA.tf₂N at either 40 or 60 °C, the experimental values are significantly higher than the predicted data. These significant differences lead to a bigger error for the total test set of 12 ILs. The average absolute error for the solubility of rutin at 40 °C is 1.41 log-units and the root mean square deviations is 1.51; while the AAE for 60 °C is 1.16 and RMSD amounts to 1.10 log-units. This result is still acceptable compared to the prediction of solid solute by other method²⁸. Inspection of the data in Fig. 3 reveals that the major errors come from the greater deviations of the predicted solubilities of rutin in BMIM.PF₆, BMIM.BF₄ and tOMA.tf₂N from the corresponding experimental data, which cover over 50% of the total absolute error. Interestingly, even for the solubility of esculin in tOMA.tf₂N, the predicted value also seriously deviates from the experimental datum in the same direction. This result seems to suggest that COSMO-RS did not give a sufficient and accurate description of the interaction between tOMA.tf₂N and flavonoid molecules.

It should be pointed out that an accurate measurement of flavonoids in those ILs with high viscosity is methodologically difficult. For instance, HMIM.Cl (7985 mPa s at 25 °C), EMIM.OTs (solid at 25°C) and dMIM.dMP (391.1 mPa s at 25 °C) have high viscosity, in which both rutin and

esculin have higher solubility according to COSMO-RS estimation. However, both flavonoids have
2 bigger molecular weight and are structurally cohesive compounds. With more solute added to the
solvents, the viscosity of the system becomes much greater (like semi-solid). In this case, agitation
4 to promote dissolution is impossible and prolonging equilibrating time (over 2 month) doesn't help
so much for diffusion. This fact demonstrated that the presented experimental data for HMIM.Cl,
6 EMIM.OTs and dMIM.dMP in Fig 2 and 3 are far less than real solubilities in these solvents due to
methodological impossibility, which means that the model estimation for the ILs with high
8 solubility is actually more accurate than the shown data. This probably may explain a better
agreement of experimental solubility of rutin with the predicted value at 60 °C, because the
10 dissolution of rutin at 60 °C is observed faster than at 40 °C in high-viscosity ILs and enable the
measured value closer to its real solubility. However, this doesn't represent the reasons accounting
12 for greater deviation of the predictions for the solubilities of rutin in BMIM.PF₆, BMIM.BF₄ and
tOMA.tf₂N, because of their relatively lower viscosity and lower solubility of flavonoids. Most
14 likely, the solubilities of rutin for the cases in these three types of ILs could be systematically
underestimated by COSMO-RS. More studies are needed to look into the issue more thoroughly.

16 Overall, COSMO-RS essentially gave a good prediction of the solubilities of representative
flavonoid molecules in the ILs with varied cation structures and different anions, as experimentally
18 validated. The predicted results at 60 °C give almost the same accuracy as at 40 °C, which indicates
that the temperature range tested in this work is within the safety and effective prediction range of
20 COSMO-RS prediction for ILs with high upper limit of liquidus. The prediction quality for the case
of esculin is better than that for rutin (Fig. 2 and 3). The model shows a better estimation for strong
22 solvation ionic liquids than those ILs with lower solubility (paired anions of BF₄⁻, PF₆⁻ and tf₂N)
(Fig. 2 and 3). The results demonstrated that the model is capable of producing a reasonable
24 prediction of the solubilities in almost arbitrary cation-anion combinations existing as liquid at the

test temperature range^{22, 25, 31}. This is perhaps of particular interest to serve as a first guide in the selection of solvents from a large pool. Most importantly, being a sound physical founded model, COSMO-RS could give reasonable force field analysis in a complicated system and therefore is able to quantify the function property of the structural moiety. This function of COSMO-RS approach is particularly useful for IL structural design for a specific task, as demonstrated in a recent work²⁶, by provision of a molecular level understanding of structure-function relationship of ILs, which constitutes intensive contents in the following section.

8 **Analysis of the multiple solvation interactions between esculin and ILs**

It is known that ionic liquids are among the most complex solvents³². Clearly, single parameter like “polarity”, normally used for the characterization of conventional solvents, is not sufficient to describe the structure and diversity of functionality of ILs. Several approaches have been proposed that allow one to examine and categorize the different solvent-solute interactions^{32, 33}. These solvatochromic or chromatographic approaches employ probe molecules to characterize the most dominant interactions of ILs, namely, polarity, hydrogen bond basicity, and dispersion, etc^{32, 33}. These efforts are capable of categorizing the types and strength of interactions of an extensive number of ILs that effectively delineate their similarities and differences. However, those descriptions could not or at least have not been associated with the quantification of the thermodynamic properties of ILs, which is just the need for a practical application²⁶. As a physically well-founded computation approach, COSMO-RS integrates dominant interactions among IL systems (electrostatic (polarity), H-bonding and van der Waals (dispersion)), which adequately summarize multiple solvation interactions of ILs³². Importantly, this methodology provides a direct quantitative scaling of the thermodynamic properties of ILs at a specific solvation environment²⁶. Therefore, it is theoretically possible to associate the specific interactions

determining the solubility of flavonoids with the cationic and/or anionic part of the ILs through force field analysis of the measures derived from COSMO-RS computation.

Fig. 4 shows the predicted solubilities of esculin in BMIM-based ILs with varying anions and the corresponding solvation interaction energies. According to the solubility of esculin, the ILs (in terms of the anions paired) could be classified into 3 groups: $\log_{10}(\text{solu}_S)$ of 0-1, 1-2 and >2 (Fig. 4). Among 3 descriptor parameters (Fig. 4), van der Waals interaction is strongly negative for all ILs and varies within a narrow range of value. Misfit interaction could be regarded as a disguised form of polarity, and in COSMO-RS it is defined as electrostatic interaction between the two contacting ensembles. Actually this term integrates the molecular shape, size and charge density and distribution³⁴, which reflects, to some extent, the dissimilarity and mismatching property of the two interacting molecules. Its positive value indicates the ionic nature of ILs is thermodynamically unfavourable for the dissolution of a neutral molecule (herein esculin). The declining values with the anions paired denote a decreasing polarity of the anions and corresponding ILs, also indicates that a less polar IL favour the dissolution of esculin if judged only by this parameter. However, the decrease of misfit interaction energy is apparently not enough to compensate for the continuous reduction of the H-bonding power of ILs to stabilize the dissolved esculin (Fig. 4). Clearly, the decrease of solubility has shown a pronounced dependency of the increase of H-bonding interaction energy (Fig. 4). In the first group of ILs the dissolved esculin is stabilized by the strong H-bonding between anion and solute, yielding a very high solubility; while for the anion in the third group like PF_6^- , the H-bonding capability attenuates seriously (H-bonding interaction energy close to 0), leading to a lower solubility of esculin. The results suggested that, for a solute with the presence of the structure (herein saccharide rings) being good H-bonding donor, the anion part of ILs or the H-bonding capability of anion plays a decisive role in the determination of solute solubility.

To have a close look into the effects of the structural variation of cations on the solubility of
2 esculin, the prediction was carried out in the ionic liquids of the 4 basic structures of imidazolium,
pyridinium, ammonium and phosphonium with variable substituents (Fig. 5). Acetate (A),
4 methylsulfate (B) and PF_6^- , representing three different groups (Fig. 4), were selected as the paired
anions to examine how structural change of the cationic part influence the solvation behaviours of
6 the ILs with different anionic property. In agreement with the results in Fig. 1B, $\log_{10}(\text{solu}_S)$ is
zero no matter what substituents incorporated into the 4 basic cation structures of the ILs when the
8 anion is acetate (Fig. 5A); in the ILs with the anion of PF_6^- the esculin solubilities varies between -3
- -4 (Fig. 5C); and in the ILs with the anion of methylsulfate the solubility has a wider fluctuating
10 range (-0.5 - -2) (Fig. 5B). As depicted in Fig. 5, for the same solute molecule of esculin, the
structural variation in the cation part results in little change of the van der Waals interaction and
12 different anions also have comparable values (around -15 kcal/mol). The strongly negative values
for acetate (Fig. 5A) and less negative value for PF_6^- (5C) of H-bonding interaction could explain
14 the high and low solubility of esculin in respective ILs. The case for the ILs with anion of
methylsulfate is in the between of acetate and PF_6^- , of which the absolute values of the H-bonding
16 and misfit interaction energies are very close. This probably may explain a slightly wide variation
range of esculin solubility. Briefly, the results presented in Fig. 5 further demonstrate that the
18 solubility of esculin in an ionic liquid is largely governed by the H-bonding capability of its anionic
part (great difference at the orders of magnitude), and the change of the cationic part generally
20 results in small variation of solubility (within 1 order of magnitude).

Analysis of the multiple solvation interactions of rutin and ILs

22 Compared to esculin, rutin has 1 more phenolic ring and a disaccharide substituent. To examine the
similarities and differences of the solvation behaviours of ILs between rutin and esculin, a similar
24 computation has been conducted for solute rutin as done for esculin (Fig. 6). Fig. 6 displays that

rutin has a wider varying range of solubility, and the minimum solubility (molar fraction) is around
2 the magnitude of 10^{-13} . The classification of solubility by the orders of magnitude of the values
appears to be very clear (Fig. 6). In the high solubility zone the logarithmic solubility of rutin is
4 zero (Group 1'); in the low solubility zone the solubility is less than 10^{-4} (Group 3') and the
solubility varies within the range of $10^{-1} - 10^{-4}$ in the second group. Comparison of Fig. 4 and 6
6 reveals that the solubility of rutin decrease against anion type in a generally similar but not totally
the same order like esculin, indicating a similar but different solvation behaviour between rutin and
8 esculin. Similar to esculin, the van der Waals interactions for rutin are nearly constant, but the
absolute values (about 25 kcal/mol) are markedly higher than those for esculin (around 15 kcal/mol).
10 This strong interaction results from bigger molecular size and mass of rutin. Different from esculin,
the misfit interaction for rutin with anion alteration shows a smaller variability, indicating that in the
12 same solvent environment different solutes behave different solvation or induce solvents to exhibit
different polarities, as reported elsewhere^{26, 32}. The data in Fig. 6 show that there are strong H-
14 bonding interactions between rutin and anions of ILs in the high soluble zone to stabilize the
dissolved molecules, and this interaction is generally greater than van der Waals interaction for the
16 ILs in this zone except for the cases of Br^- and toluene-4-sulfonate. This result suggests that the
additional saccharide ring of rutin adds much to the H-bonding interaction with ILs. However, this
18 structural characteristic does not always generate a desirable effect for the dissolution of rutin;
because more saccharide rings will result in a bigger molecular misfit or increase the molecular
20 dissimilarity with hydrophobic ILs, thus, lead to a significant lower solubility of rutin in those ILs
with weak H-bonding capability (Group 3' in Fig. 6).

22 As did for esculin, we also calculated the rutin solubility in the ILs having cations with
varying substituents. Similar to the observations for esculin, the rutin solubility in the ILs with
24 acetate anion is zero and in the ILs with PF_6^- anion is less than 10^{-6} (logarithm of molar fraction)

regardless of substituent variation in the cation (details not shown). A greater variation range of the rutin solubility in the ILs containing methylsulfate as anion but with the alteration of substituents in the cation can be seen in Fig. 7. This change plausibly corresponds to the change of misfit interaction, indicating that rutin appears to be more sensitive to the incorporation of hydrophobic substituents due to the massive occurrence of polar saccharide rings. The increasing hydrophobicity of the cation may enhance the repulsion between IL and rutin, and counteract the H-bonding interaction that facilitates dissolution of rutin. This interaction may thus exert a more significant effect for those ILs with moderate H-bonding capability (e.g. methylsulfate), and leads to an evident decrease of the solubility of rutin (Fig. 7).

10 **Concluding remark and outlook**

The primary interest of this study is not only to validate COSMO-RS predictions, but also to achieve a better understanding of the functionality of structural moiety of ILs through the powerful force field analysis function of COSMO-RS methodology based on its sound physical basis. It turns out that COSMO-RS predictions generally have a good agreement with the experimental data, compared to the predictions of the solubilities of solid solute by other approaches^{25,28}. Considering that the predictions of this model cover vast structural diversity of ILs and the calculations are performed without any specific parameter adjustment, the results and accuracy are encouraging and acceptable. The physically founded basis of COSMO-RS and validation of this work proved that the predicted results are reliable, and thus the necessary experimental efforts for quantitative determination of the solubilities of flavonoids in ILs could be reduced. The results of this work also revealed a reasonable anion-dependency of the solubility of flavonoids in ILs, and accordingly presented a first systemic categorization of anions based on flavonoid solubility. With this interesting finding, we anticipate that the grouping for anions of ILs might be generally applicable to those solutes with massive moieties as H-bonding donor. We note the reports so far concerning

the ILs that are capable of dissolving cellulose³⁵, carbohydrate³⁶, and protein,³⁷ etc, the anions of which (Cl^- , Br^- , H_2PO_4^- and $(\text{CN})_2\text{N}^-$, etc) are exclusively included in the first group of the categorization in this study ($(\text{CN})_2\text{N}^-$ on the border) (Fig. 4 and 6). To confirm this interesting finding, we calculated the solubility of two repetitive units of cellulose, carbohydrate and protein (used in place of macromolecular structures) in BMIM-based ILs with the anion spectrum as in Fig. 6. The predictions give surprisingly good agreement with the categorization of anions for rutin (data not shown). The results strengthened the experimental basis of COSMO-RS, and further identified the general applicability of this physically founded model, as well as the reasonability of a logical extension of some conclusions in this work.

Importantly, the descriptors derived from COSMO-RS are also shown to give a different, but surprisingly equivalent description of the physics of molecules and a similar quantitative scaling in the corresponding terms of another experimental based solvation model – Abraham equation^{32, 38}. The overlap of chemical content of the molecular descriptors between COSMO sigma-moments of CSOMO-RS and experimental descriptors in the Abraham equation has been demonstrated elsewhere³⁸. **COSMO sigma-moments (total 5 parameters) are molecular descriptors derived from COSMO-RS calculation, among which the second and third sigma moment (Sig2 and Sig3) and the hydrogen bond moments (HB_acc3 and HB_don3), are chemically corresponding to the measures of polarity/polarizability, H-bonding basicity and H-bonding acidity, respectively³⁸. The zeroth sigma-moment is identical with the molecular surface³⁸.** Table 1 lists the HB_acc3 of the anions evaluated in this work. Anions are good H-bonding acceptors but negligible donors (HB_don3 is generally zero, data not shown). This parameter could be used to classify ILs into 3 groups based on basicity (Table 1). Clearly, the categorization of anions in Table 1 is similar but not the same as in Fig. 4 and 6. For example, ethoxyethylsulfate and methoxyethylsulfate (Group I) belong to Group 2 in Fig. 4 or Group 2' in Fig. 6. The reason is that HB_acc3 is an intrinsic property of the solvent

and does not change with specific solute. However, the specific interactions between IL and
2 different solutes can be different due to varying molecular size, surface charge density and
distribution, symmetry, etc of the solute. Therefore, misfit, H-bonding and van der Waals
4 interaction energies employed in this work could theoretically give a more correct and accurate
description for the solvation behaviour of flavonoids in ILs. Our results show a comparable
6 visualization of the solvation behaviours of ILs in the characterization of ionic liquids as described
by Anderson et al ³². Namely, the anion has a greater influence on the overall H-bond basicity of
8 ILs; hydrogen bond basicity varies significantly with anions but vary little for cation alternation;
and the dispersion forces show only a slight variability for the ILs evaluated. However, as a first
10 rough selection of ILs for flavonoid dissolution, HB_acc3 is very useful to evaluate the H-bonding
capability of anions (Table 1 and Fig. 4 and 6). Based on the changes of measures from COSMO-
12 RS versus cationic variation depicted in Table 2, it is also easy to reason why misfit interaction
energy decreases, H-bonding little changes and the corresponding decrease of solubility of
14 flavonoids at small scale follows the changes of cationic substituents (Fig. 5). In brief, the results in
this study demonstrated that, as an experimentally independent approach, COSMO-RS is capable of
16 producing a comparable high-quality characterization of multiple solvation interactions of ILs
comparable to that obtained from other experimental models. The knowledge and understanding of
18 the relationship of property with interactions and characteristic moiety of ILs can thereby serve
molecular design and structural optimization for constructing a desirable structure with high
20 solubility of flavonoids. However, the establishment of an efficient enzymatic reaction system
involves some unpredictable parameters, such as enzyme activity. Fortunately, COSMO-RS can
22 also aid our effort in some way, because our results show that some parameters, such as water
activity and activity coefficients in ILs relevant to enzyme activity, could be accurately estimated
24 by COSMO-RS. The molecular design of ILs assisted by COSMO-RS is in progress in our group.

Experimental

2 Esculin and rutin hydrate (with the purity >99%) were purchased from Sigma-Aldrich Co. (St.
Louis, USA). Dimethylsulfonate (DMSO), methanol, acetic acid and triethylamine are from Sigma-
4 Aldrich Co. (St. Louis, USA) and of HPLC grade. 1-Hexyl-3-methylimidazolium (HMIM.Cl), 1-
butyl-3-methylpyridinium dicyanamide (BMPy_i.N(CN)₂), 1-ethyl-3-methylimidazolium toluene-4-
6 sulfonate (EMIM.OTs), 1,3-dimethylimidazolium dimethylphosphate (dMIM.dMP), 1-ethyl-3-
methylimidazolium n-octylsulfate (EMIM.OctSO₄), 1-ethyl-3-methylimidazolium 2(2-
8 methoxyethoxy)ethylsulfate (EMIM.MDEGSO₄), 1-ethyl-3-methylimidazolium ethylsulfate
(EMIM.ES), methyltrioctylammonium bis(trifluoromethylsulfonyl)imide (tOMA.tf₂N), 1-butyl-3-
10 methylimidazolium tetrafluoroborate (BMIM.BF₄) and 1-Butyl-3-methylimidazolium
hexafluorophosphate (BMIM.PF₆) were procured from Solvent Innovation GmbH (Köln, Germany)
12 and of minimum 98% purity. Methyltrioctylammonium trifluoroacetate (tOMA.TFA) is from
Merck KGaA (Darmstadt, Germany) and with a purity > 99.7%. 1-butyl-1-methyl pyrrolidinium
14 dicyanamide (BMPy_o.N(CN)₂) is purchased from IoLiTec Ionic Liquids Technologies GmbH & Co
KG and of >98% purity (Denzlingen, Germany). The dissolution and equilibration of esculin or
16 rutin in the solvent ILs were performed in a thermostat oven at 40 or 60 °C with ±0.1°C accuracy.
Typically, 2 mL of ionic liquid were accurately added in a 10 mL capped bottle fixed on a
18 Variomag Telesystem with multiple magnetic stirrers (H+P Labortechnik AG, Oberschleissheim,
Germany). The batchwise added esculin or rutin was dissolved with continuous magnetic stirring
20 and the dissolution lasted over 2 months to allow sufficient equilibration. For the case of
EMIM.OTs, ultrasonication was employed to assist the dissolution of rutin. The undissolved solid
22 was removed by pressure filtration with Syringe Filter (with 0.45 µm PTFE membrane) (Pall Life
Science, Ann Arbor, MI, USA) at the same temperature. The resulting IL solution was immediately
24 dissolved in DMSO for HPLC analysis.

The standard curves of esculin, rutin and different ILs were established, respectively. A series of sample concentrations of 0.05, 0.1, 0.2, 0.5, 1, 2, 3, 5 and 10 mg/mL in DMSO was used and the means of triplicate determinations were adopted. Based on the property of ILs, two elution systems were used for the HPLC analysis of esculin (rutin) dissolved in ILs. The solubility of flavonoids in BMIM.PF₆, BMIM.BF₄, tOMA.tf₂N and tOMA.TFA were eluted with methanol/water (containing 0.1% acetic acid); while other types of ILs use methanol/acetate-triethylamine (TEA) buffer (20 mM; pH, 4.0) elution system. The HPLC analysis was performed on Hitachi-Merck HPLC Series 7000 (Hitachi-Merck, Japan), conjugated with a PL-ELS 2100 evaporative light scattering detector (ELSD) (Polymer Laboratories, Shropshire, UK). The reverse phase column employed was Supelcosil LC-18 (250 mm × 4.6 mm) (Supelcosil Inc., Bellefonte, PA). The ELSD was operated at an evaporating temperature of 100 °C and a nebulizing temperature of 50 °C with air as the nebulizing gas at 1.2 SLM. For either methanol/water or methanol/buffer, the elution gradient follows the same program: starts with 30% methanol phase and increase to 100% methanol phase in 10 min; and holds 6 min and then reduces to 30% methanol phase in 3 min, and keeps at this phase ratio for another 10 min. The mobile phase flow rate was 1.0 mL/ min. Area percentage was used as mass for solubility calculation. The measured values of the IL and flavonoid were calibrated using standard curves, respectively. All HPLC analyses were determined in triplicate and the means were used for evaluation.

Acknowledgements

The authors thank A. Klamt and M. Diedenhofen for their assistance in model processing. Financial support from Danish Research Council for Technology and Production (FTP) (274-05-0286) and Center for Advanced Food Studies (LMC) is gratefully acknowledged.

References

1 A. Wong, I. Toth, *Curr. Med. Chem.* 2001, **8** (9), 1123-1136.

- 2 M. Gulati, M. Grover, S. Singh, M. Singh, *Inter. J. Pharm.* 1998, **165** (2), 129-168.
- 2 3 W. N. Charman, C. J. H. Porter, *Adv. Drug Delivery Rev.* 1996, **19** (2), 149-169.
- 4 C. Rice-Evans, *Curr. Med. Chem.* 2001, **8** (7), 797-807.
- 4 5 S. Lesser, R., Cermak, S. Wolffram, *J. Nutri.* 2004, **134** (6), 1508-1511.
- 6 R. Hirano, W. Sasamoto, A. Matsumoto, H. Itakura, O. Igarashi, K. Kondo, *J. Nutr. Sci. Vitamin.*
6 2001, **47** (5), 357-362.
- 7 S. Riva, G. Carrea, G. Ottolina, F. Secundo, B. Danieli, P. De Bellis, *Annn. N. Y. Acad. Sci.* 1996,
8 712-715.
- 8 P. Kolář, J.-W. Shen, A. Tsuboi, T. Ishikawa, *Fluid Phase Equilibria* 2002, **194-197**, 771-782.
- 10 9 T. C. Frank, J. R. Downey, S. K. Gupta, *Chem. Eng. Prog.* 1999, **95** (12), 41-61.
- 10 C. Reichardt, *Solvents and solvent effects in organic chemistry*, 2nd ed., 1996, VCH, New York.
- 12 11 M. H. Katsoura, A. C. Polydera, L. Tsironis, A. D. Tselepis, H. Stamatis, *J. Biotechnol.* 2006,
123 (4), 491-503.
- 14 12 R. A. Sheldon, R. M. Lau, M. J. Sorgerdrager, F. van Rantwijknd and K. R. Seddon, *Green
Chem.* 2002, **4**, 147-151.
- 16 13 F. Brennecke and E. J. Maginn, *AIChE Journal* 2001, **47**, 2384-2389.
- 14 J. Dupont, R. F. de Souza and P. A. Z. Suarez, *Chem. Rev.*, 2002, **102**, 3667-3692.
- 18 15 F. van Rantwijk, R. A. Sheldon, *Chem. Rev.* 2007, **107**, 2757-2785.
- 16 Z. Guo, B.-M. Lue, X. Xu, *Inform*, 2007, **18** (2), 78-82
- 20 17 Z. Guo, X. Xu, *Org. Biomol. Chem.* 2005, **3**, 2615-2619.
- 18 Z. Guo, X. Xu, *Green Chem.* 2006, **8**, 54-62.
- 22 19 C. Jork, C. Kristen, D. Pieraccini, A. Stark, C. Chiappe, Y. A. Beste, W. Arlt, *J. CHEM.
Thermodyn.* 2005, **37** (6), 537-558.
- 24 20 Z. Guo, B. Chen, R. L. Murillo, T. Tan, X. Xu, *Org. Biomol. Chem.* 2006, **4**, 2772-2776.

- 21 K. Wichmann, M. Diedenhofen, A. Klamt, *J. Chem. Inf. Model.* 2007, **47**, 228-233.
- 2 22 S. Oleszek-Kudlak, M. Grabda, E. Shibata, F. Eckert, T. Nakamura, *Environ. Toxic. Chem.* 2005, **24** (6), 1368-1375.
- 4 23 A. Klamt, *J. Phys. Chem.* 1995, **99**, 2224-2235.
- 24 F. Eckert, A. Klamt, *AIChE Journal.* 2002, **48**, 369-385.
- 6 25 A. Klamt, F. Eckert, M. Hornig, M. E. Beck, T. Bürger, *J. Comput. Chem.* 2002, **23**, 275-281.
- 26 B. Chen, Z. Guo, T. Tan, X. Xu, *Biotechnol. Bioeng.* 2007, **97**, in press.
- 8 27 M. Fermeglia, P. Braiuca, L. Gardossi, S. Pricl, P. J. Halling, *Biotechnol. Prog.* 2006, **22**, 1146-1152
- 10 28 S. Gracin , T. Brinck , A. C. Rasmuson, *Ind. Eng. Chem. Res.* 2002, **41** (20), 5114-5124.
- 29 J. L. Kaar, A. M. Jesionowski, J. A. Berberich, R. Moulton, A. J. Russell. *J. Am. Chem. Soc.*
- 12 2003, **125**, 4125–4131.
- 30 S. Stolte, J. Arning, U. Bottin-Weber, M. Matzke, F. Stock, K. Thiele, M. Uerdingen, U. Welz-
- 14 Biermann, B. Jastorff, J. Ranke, *Green Chem.* 2006, **8** (7), 621-629.
- 31 A. Klamt, *Fluid Phase Equilibria* 2003, **206**, 223-235.
- 16 32 L. Anderson, J. Ding, T. Welton, D. W. Armstrong, *J. Am. Chem. Soc.* 2002, **124**, 14247-14254.
- 33 C. Reichardt, *Green Chem.* 2005, **7**, 339-351.
- 18 34 A. Klamt, F. Eckert, *Fluid Phase Equilib.* 2000, **172**, 43-72.
- 35 R. P. Swatloski, S. K. Spear, J. D. Holbrey, R. D. Rogers, *J. Am. Chem. Soc.* 2002, **124** (18),
- 20 4974-4975.
- 36 Q. B. Liu, M. H. A. Janssen, F. van Rantwijk, R. A. Sheldon, *Green Chem.* 2005, **7** (1), 39-42.
- 22 37 K. Fujita, D. R. MacFarlane, M. Forsyth, *Chem. Commun.* 2005, **38**, 4804-4806.
- 38 A. M. Zissimos, M. H. Abraham, A. Klamt, F. Eckert, J. Wood, *J. Chem. Inf. Comput. Sci.* 2002,
- 24 **42**, 1320-1331.

Tables and Figures

- 2 **Table 1** COSMO-RS descriptor of HB_acc3 (hydrogen bonding acceptor moment indicates hydrogen bond basicity) for 32 anions of ionic liquids and classification of solvation interactions.
- 4 For structural formulae, see supplementary information.

| Group I | | Group II | | Group III | |
|--------------------------------------------|---------|-------------------------------------|---------|----------------------------------------------|---------|
| Anion types | HB_acc3 | Anion types | HB_acc3 | Anion types | HB_acc3 |
| acetate | 39,0533 | ethylsulfate | 19,2952 | bis(trifluoromethyl)imide | 4,0642 |
| decanoate | 38,2772 | octylsulfate | 19,1774 | bis(trifluoromethyl sulfonyl)methane | 3,8915 |
| bis(2,4,4-trimethyl pentyl) phosphinate | 37,8887 | butylsulfate | 19,1238 | clo4 | 3,7574 |
| cl | 36,6278 | methylsulfate | 18,5006 | tetracyanoborate | 3,4554 |
| phosphate | 35,9622 | dicyanamide | 15,5327 | bisoxalatoborate | 3,0140 |
| dimethylphosphate | 35,1611 | Bisbiphenyl diolatoborate | 12,7544 | bf4 | 2,4740 |
| br | 29,6444 | bis-pentafluoroethyl phosphinate | 12,6559 | tf2n | 2,3089 |
| toluene-4-sulfonate | 25,9431 | bissalicylatoborate | 12,6200 | tris(nonafluorobutyl) trifluorophosphate | 0,0747 |
| ethoxyethylsulfate | 20,6579 | bismalonatoborate | 10,8007 | tris(pentafluoroethyl) trifluorophosphate | 0,0108 |
| methoxyethylsulfate | 20,5119 | trifluoromethane- sulfonate | 10,6147 | pf6 | 0.0000 |
| trifluoroacetate | 20,3944 | | | | |
| nitrate | 19,4858 | | | | |

6

8

Table 2 COSMO-RS descriptors of molecular surface area, the second sigma moment 2 (Sig 2 indicates polarity/polarizability) and HB_don3 (hydrogen bonding donor moment indicates hydrogen bond acidity) for the cations of 1-alkyl-3-methylimidazolium (alkylMIM).

| Cations | Area (Å ²) | Sig 2 | HB_don3 |
|---------------|------------------------|---------|---------|
| methylMIM | 143.6001 | 88.0259 | 2.0735 |
| ethylMIM | 161.7652 | 85.3777 | 2.0727 |
| butylMIM | 201.9685 | 84.7111 | 2.0647 |
| pentylMIM | 221.9982 | 85.2520 | 2.0537 |
| hexylMIM | 241.7418 | 86.0083 | 2.0371 |
| heptylMIM | 261.8546 | 86.6088 | 1.9864 |
| octylMIM | 281.5146 | 87.3381 | 2.0014 |
| decylMIM | 321.5566 | 88.7644 | 1.9834 |
| dodecylMIM | 361.3977 | 90.2467 | 1.9988 |
| tetradecylMIM | 401.0784 | 91.7587 | 1.9814 |
| hexadecylMIM | 441.0983 | 93.3169 | 1.9932 |
| octadecylMIM | 480.4567 | 94.8666 | 1.9960 |

Scheme 1 Structures of rutin and esculin.

2

4

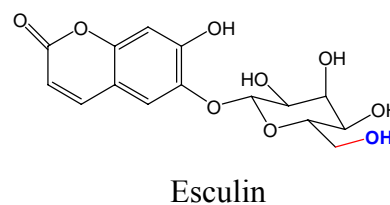
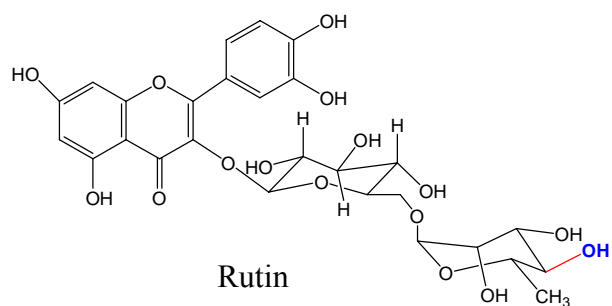
6

8

10

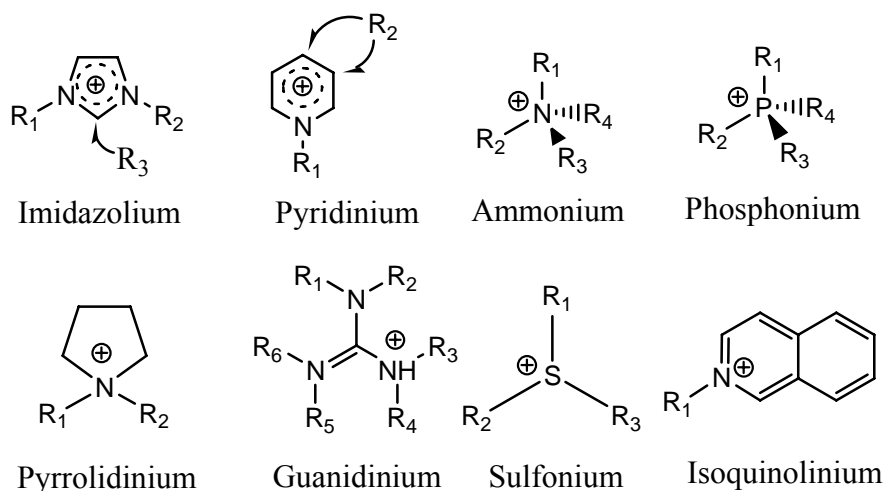
12

14



Scheme 2 Basic structures of the cations of ionic liquids evaluated in this study

2



R1-R6 represent substituents, typically alkyl-, H-, etc.

4

6

8

10

12

14

16

18

20

Captions to Figures

2 **Fig. 1** COSMO-RS predictions of the solubility of esculin in 1888 types of ionic liquids
(theoretically possible) combined by 59 types of cations and 32 types of anions at 298.15 K. The
4 data shown are categorized by cations (A) and by anions (B), respectively. 59 Cations have basic
structures as shown in Scheme 2 but different substituents, which are not shown in this figure. The
6 data in Fig. 1A and 1B are the same data but plotted versus cation and anion variation, respectively.
See the supplementary information for the chemical structures of the anions evaluated.

8 **Fig. 2** COSMO-RS predictions plotted versus experimental values of the solubility of esculin in 12
types of ILs at 313.15 K (A) and 333.15 K (B). Abbreviations for ILs see Experimental section.

10 **Fig. 3** COSMO-RS predictions plotted versus experimental values of the solubility of rutin in 12
types of ILs at 313.15 K (A) and 333.15 K (B). Abbreviations for ILs see Fig. 2.

12 **Fig. 4** COSMO-RS derived descriptors used to characterize the solvation interactions of esculin and
BMIM-based ionic liquids with different anions. (○) predicted solubility. (◇) misfit interaction
14 energy, (□) H-bonding interaction energy and (△) van der Waals interaction energy in respective
ILs at 298.15 K. *All interactions are calculated at indefinite dilution. See the supplementary
16 information for the chemical structures of the anions evaluated.*

Fig. 5 Solvation interactions of esculin and acetate- (A), methylsulfate- (B) and
18 hexafluorophosphate- (PF₆) based ionic liquids with different cations characterized by COSMO-RS
derived descriptive parameters. (○) predicted solubility. (◇) misfit interaction energy, (□) H-
20 bonding interaction energy and (△) van der Waals interaction energy in respective ILs at 298.15 K.
All interactions are calculated at indefinite dilution. Abbrev: C1MIM to C18MIM represents 1-
22 methyl- to octadecyl- 3-methylimidazolium, respectively. PhMIM is 1-benzyl-3-methyl-
imidazolium. C2Py to C8Py stands for 1-ethyl- to octyl- pyridinium, respectively, and C4MPy,

C6MPy and C8MPy corresponds to 1-butyl-, hexyl- and octyl-3-methylpyridinium. Tetra-
2 methylammonium (tetra-C1A), tetra-ethylammonium (tetra-C2A), tetra-n-butylammonium (tetra-
C4A) and methyl-trioctyl-ammonium (t-C8C1A or tOMA). tetrabutyl-phosphonium (tetra-C4P),
4 triisobutyl-methyl-phosphonium (*t-i*-C4C1P), trihexyl-tetradecyl-phosphonium (t-C6C14P) and
benzyl-triphenyl-phosphonium (t-PhCH₂PhP).

6 **Fig. 6** COSMO-RS derived descriptors used to characterize the solvation interactions of rutin and
BMIM-based ionic liquids with different anions. (○) predicted solubility. (◇) misfit interaction
8 energy, (□) H-bonding interaction energy and (△) van der Waals interaction energy in respective
ILs at 298.15 K. **All interactions are calculated at indefinite dilution. See the supplementary**
10 **information for the chemical structures of the anions evaluated.**

Fig. 7 Solvation interactions of rutin and methylsulfate based ionic liquids with different cations
12 characterized by COSMO-RS derived descriptive parameters. (○) predicted solubility. (◇) misfit
interaction energy, (□) H-bonding interaction energy and (△) van der Waals interaction energy in
14 respective ILs at 298.15 K. **All interactions are calculated at indefinite dilution.** The abbreviations
are the same as in Fig. 5.

16

Fig. 1

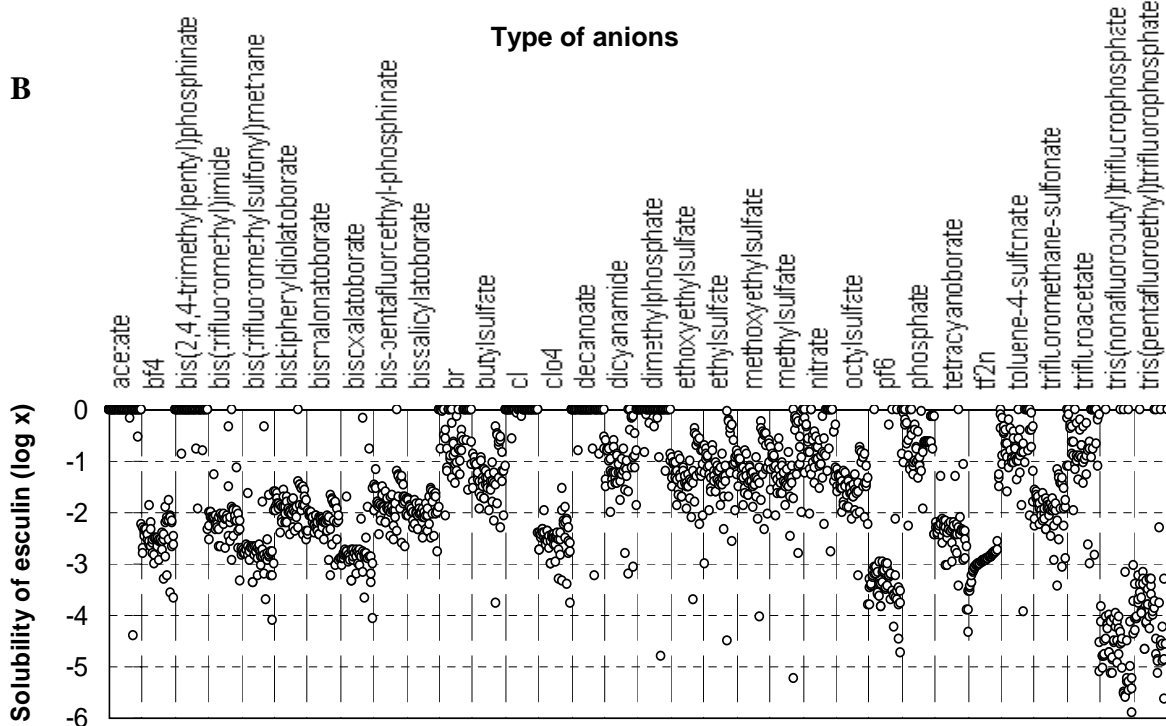
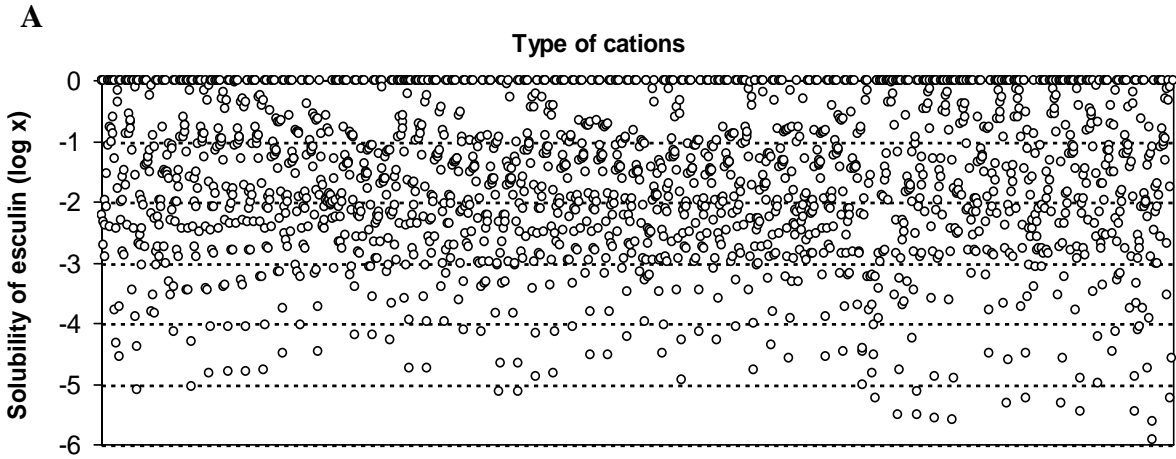


Fig. 2

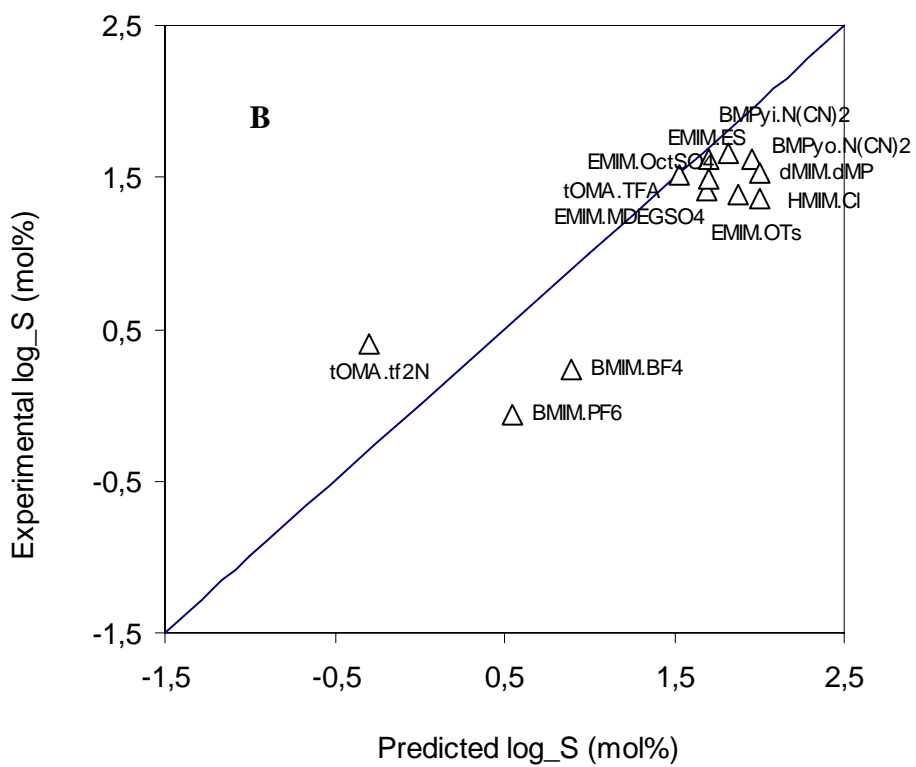
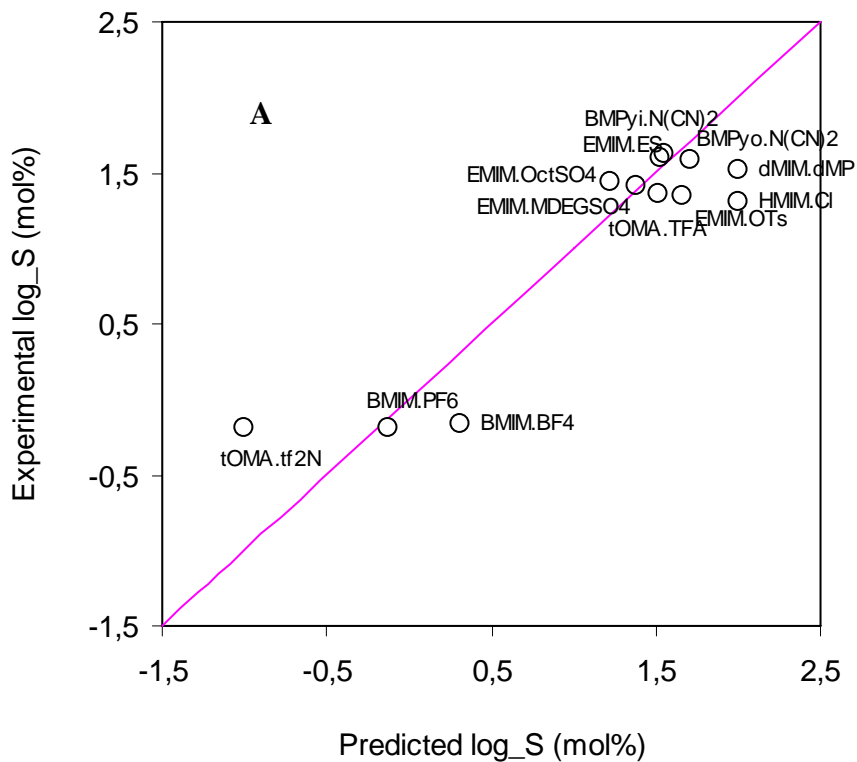


Fig. 3

2
4
6
8

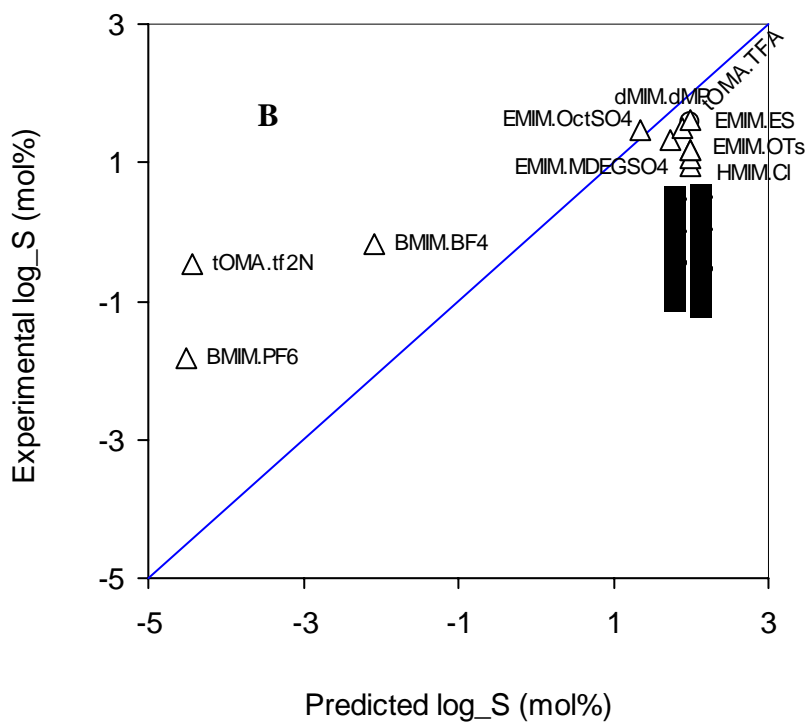
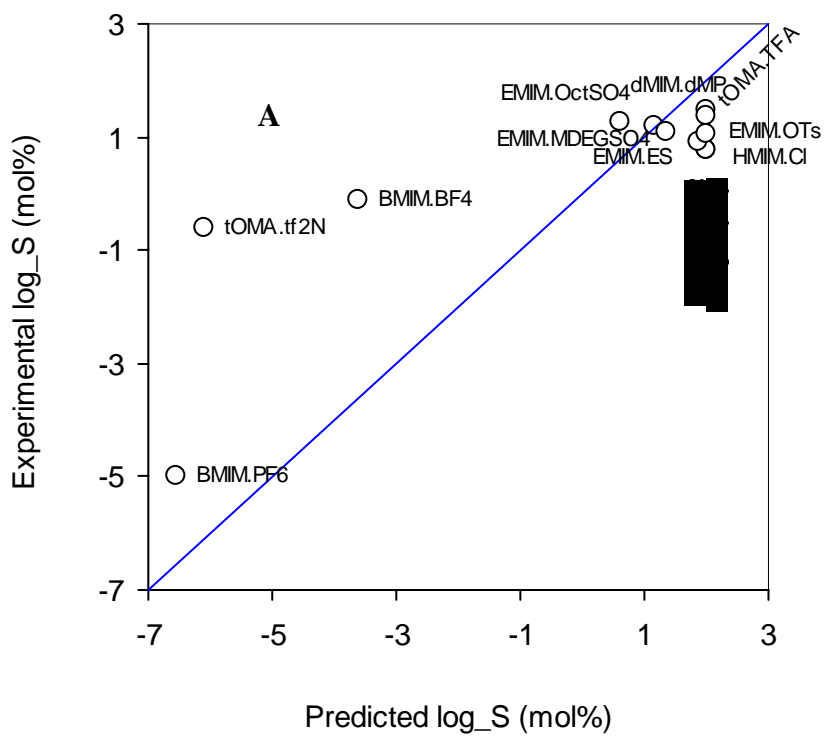


Fig. 4

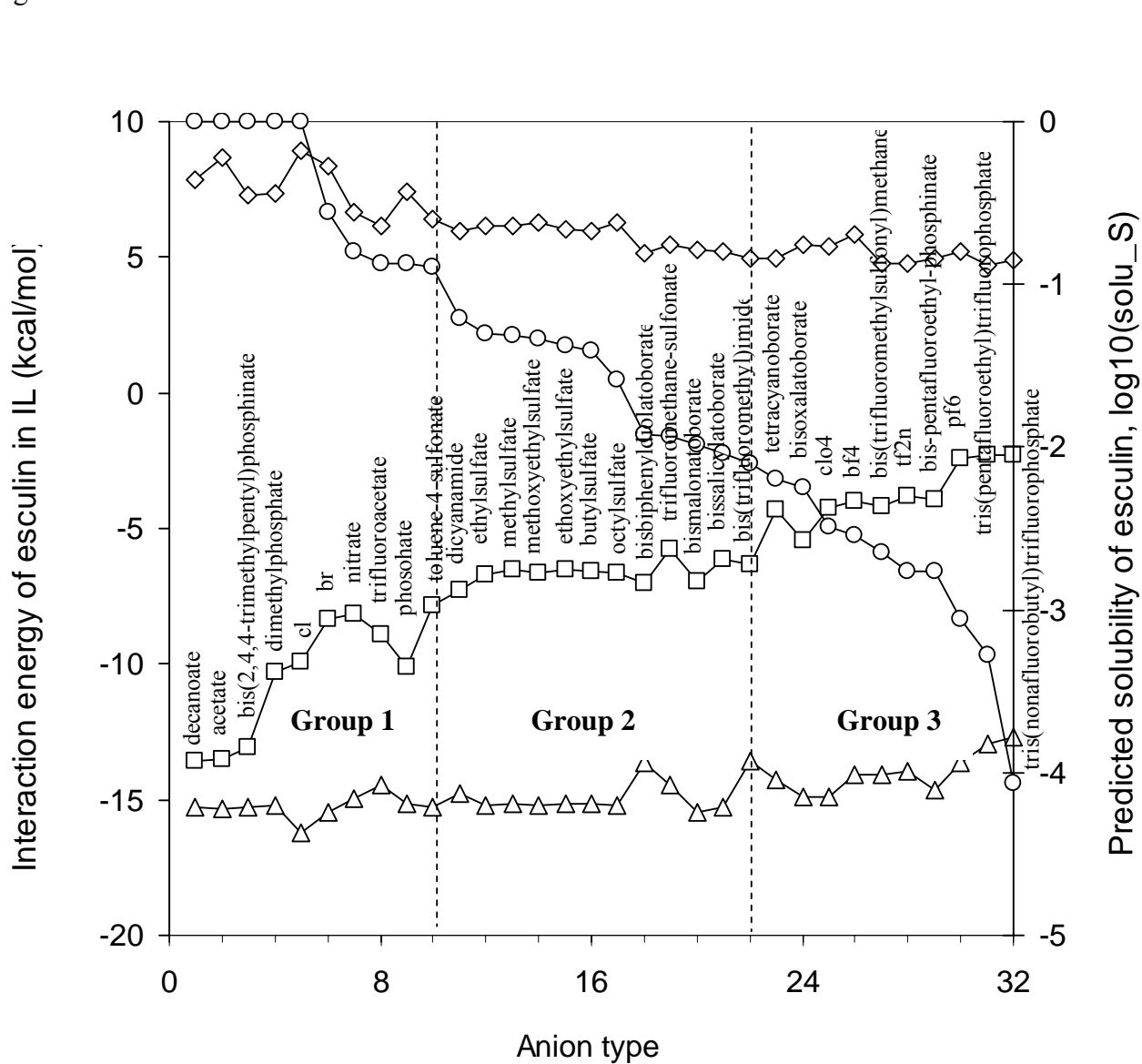


Fig. 5

2
4

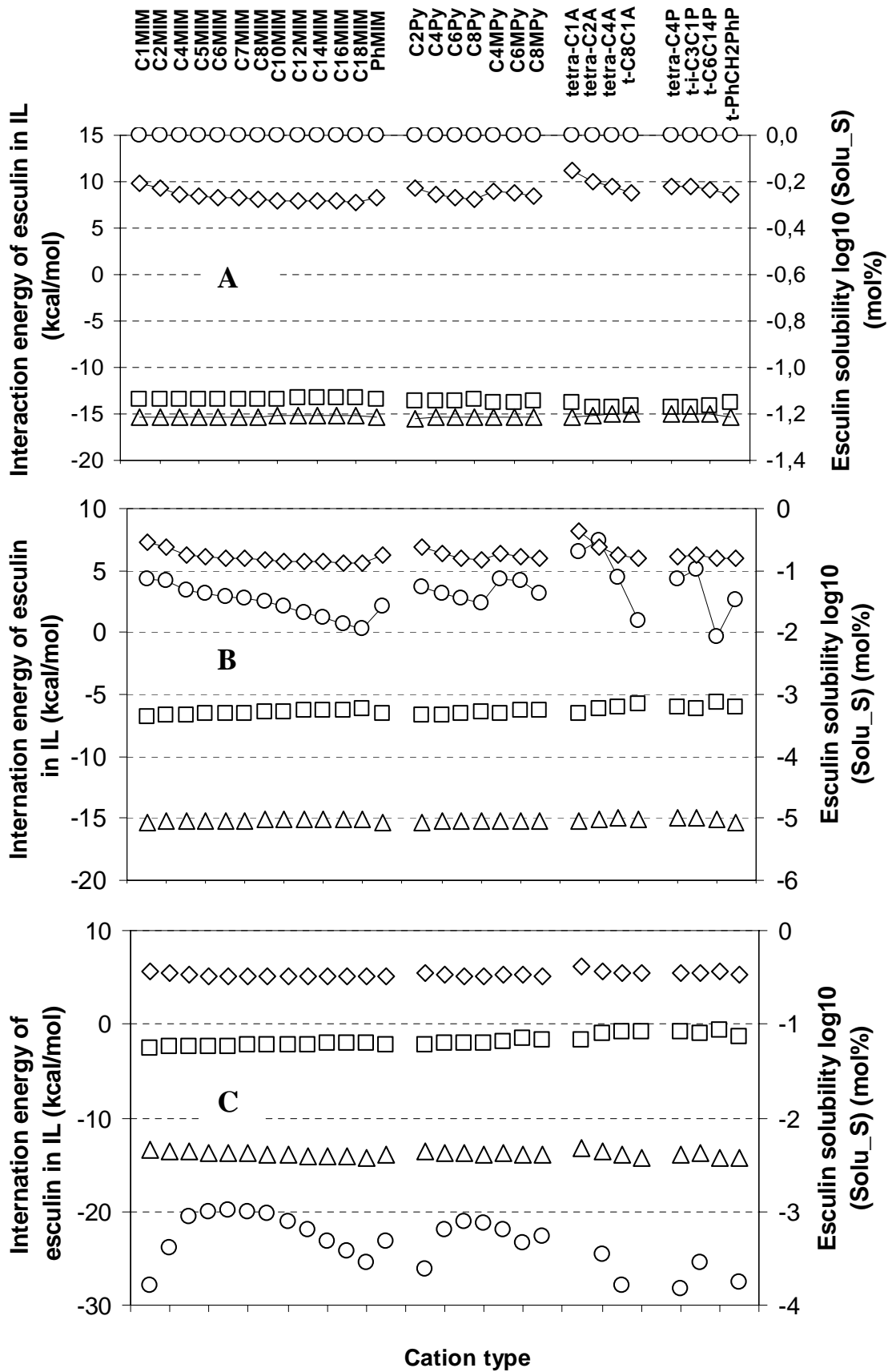


Fig. 6

2
4
6
8
10
12
14

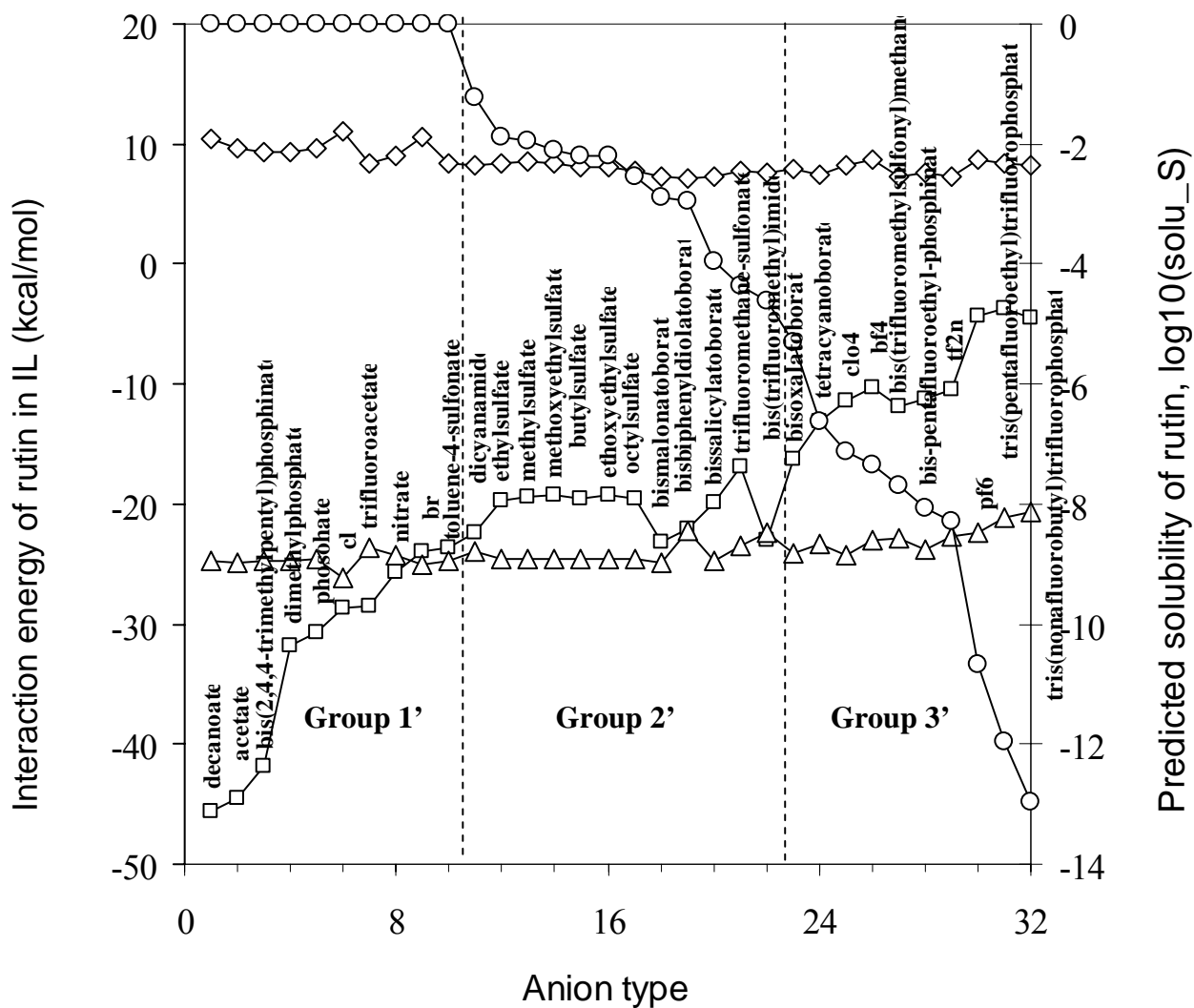
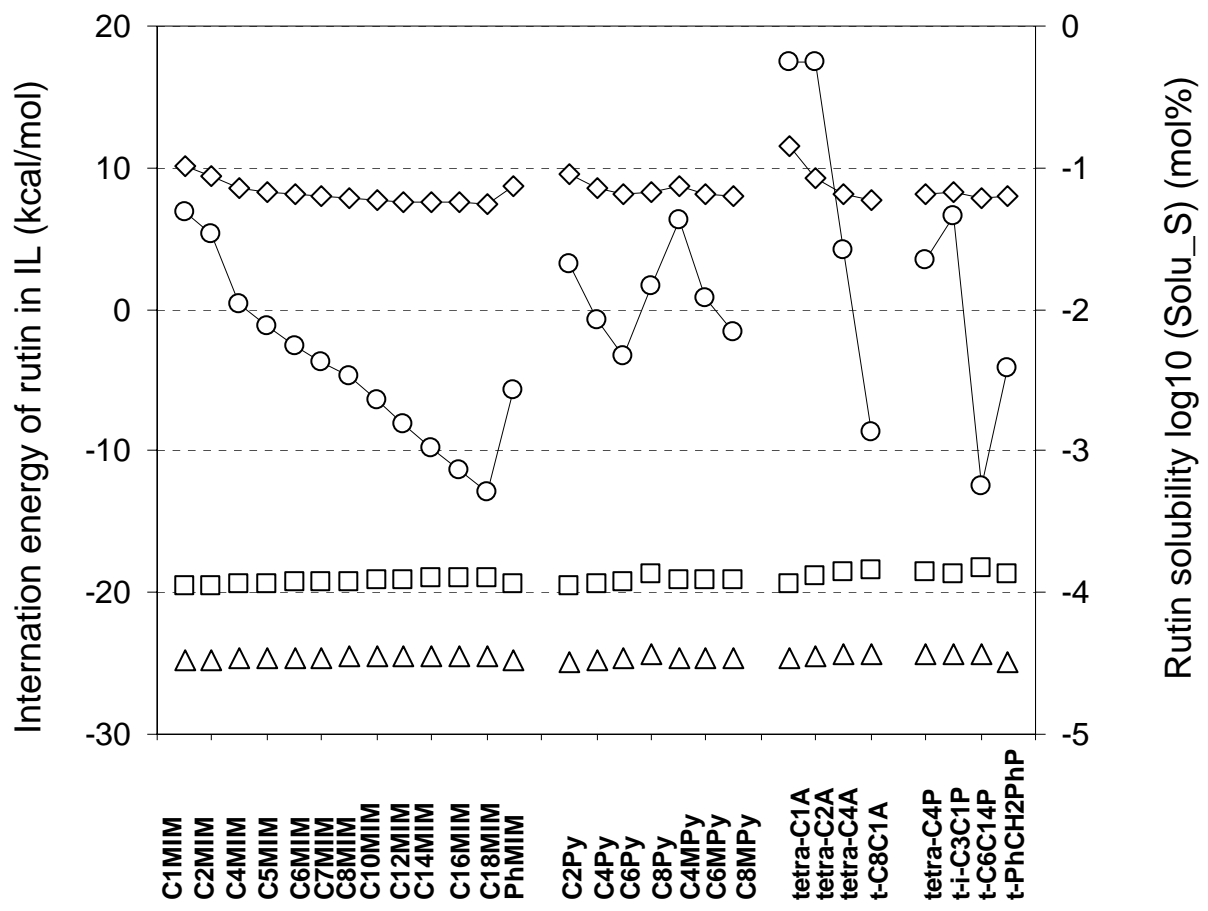


Fig. 7

2



4

6