2026 Volume: 01 Issue: 04 ISSN 1671-2838 Page No: 01-13

Review Article

Pharmaceutical Care and Research

https://pcrjournal.com

Volume: 01
Issue: 04

COSMO-RS IN PHARMACEUTICAL PREFORMULATION: THERMODYNAMIC FOUNDATIONS, PREDICTIVE APPLICATIONS, AND INTEGRATION INTO MODEL-A REVIEW

Preetha Puthezhath[1]*, Dr C. K. Dhanapal[2], Dr K. Krishna Kumar[1], Dr B. Dinesh Kumar[1] Anjana John[1], Sminu Jos[1] Arathy Jolly[1]

[1]Annamalai Ph D scholar & St James College of Pharmaceutical Sciences, Kerala, India.

[2]Annamalai University, Tamil Nadu, India.

Article Received: 03 June 2026
Article Revised: 23June 2026
Published on: 13 July 2026

***Corresponding Author: Preetha Puthezhath**
Annamalai Ph D scholar & St James College of Pharmaceutical Sciences, Kerala, India.

DOI: https://doi-doi.org/101555/ijpcr.9185

ABSTRACT

Preformulation studies play a critical role in identifying physicochemical properties that influence the development of pharmaceutical dosage forms. Traditional experimental screening methods for solubility, drug–excipient compatibility, and solvent selection are often labor-intensive and time-consuming. The Conductor-like Screening Model for Real Solvents (COSMO-RS) is a quantum chemistry-based statistical thermodynamic approach capable of predicting molecular interactions and thermodynamic properties without extensive experimental data. This review summarizes the theoretical principles of COSMO-RS and its computational implementation through COSMOtherm and related platforms. Applications in pharmaceutical preformulation are discussed, including solubility prediction, solvent screening, Hansen solubility parameter estimation, drug–excipient compatibility assessment, cocrystal screening, pKa prediction, activity coefficient modeling, and permeability estimation. A computational–experimental workflow for integrating COSMO-RS predictions into model-informed formulation development is proposed. The integration of molecular descriptors with thermodynamic modeling provides a promising digital strategy for accelerating pharmaceutical formulation design while reducing experimental workload.

KEYWORDS: Preformulation, COSMO RS, solubility prediction, quantum descriptors, molecular thermodynamics, Drug-excipient compatibility

1. INTRODUCTION

Preformulation studies constitute the scientific cornerstone of rational dosage form development, providing critical insights into the physicochemical and biopharmaceutical properties that govern

formulation performance. Fundamental attributes—including aqueous solubility, partition coefficient (logP), dissociation constant (pKa), polymorphic behavior, hygroscopicity, and drug–excipient compatibility—directly influence formulation selection, processability, stability, and in vivo bioavailability [1–3]. For contemporary drug candidates, particularly those emerging from high-throughput screening pipelines, poor aqueous solubility and complex solid-state behavior remain major barriers to successful translation into clinically viable dosage forms.

Conventional preformulation strategies rely predominantly on empirical and experimentally intensive methodologies such as solvent screening, shake-flask solubility determination, compatibility testing under accelerated conditions, differential scanning calorimetry, thermogravimetric analysis, and phase equilibrium studies. Although scientifically rigorous, these approaches are laborious, time-consuming, and resource-intensive. The challenge is especially pronounced for Biopharmaceutics Classification System (BCS) Class II and IV drugs, where extensive screening across multiple solvents, surfactants, and excipients is often required to identify optimal formulation components [4]. Such iterative experimental workflows not only prolong development timelines but also increase material consumption and development costs.

In response to these limitations, model-informed drug development (MIDD) frameworks increasingly incorporate computational tools to rationalize and streamline preformulation decision-making. Physics-based predictive models provide mechanistic insight into molecular interactions, enabling in silico estimation of solubility, miscibility, and intermolecular compatibility prior to experimental validation. Among these computational approaches, the Conductor-like Screening Model for Real Solvents (COSMO-RS) has emerged as a robust and scientifically validated platform that integrates quantum chemical calculations with statistical thermodynamics to predict thermodynamic properties of fluids and mixtures [5–7]. By generating sigma profiles and chemical potential distributions, COSMO-RS enables quantitative assessment of solvation behavior, activity coefficients, excess enthalpies, and drug–excipient interaction tendencies.

The application of COSMO-RS in pharmaceutical research has demonstrated promising capabilities in predicting solubility in organic solvents, screening lipid excipients, and guiding selection of surfactants for advanced drug delivery systems [5–7]. Its ability to evaluate molecular-level interactions without extensive experimental input positions it as a powerful tool for rational formulation design. Furthermore, integration of COSMO-RS predictions with experimental validation can significantly reduce the scope of solvent screening and minimize trial-and-error experimentation.

However, despite its demonstrated utility, a critical research gap remains. Current applications of COSMO-RS in pharmaceutical sciences are largely limited to solvent solubility prediction and isolated excipient screening. There is insufficient systematic integration of COSMO-RS–based molecular interaction analysis into comprehensive formulation design workflows, particularly for complex colloidal systems such as nanoemulsions and intranasal delivery platforms. Moreover, limited studies correlate COSMO-RS–predicted thermodynamic parameters with experimentally

observed formulation performance indicators such as droplet size distribution, physical stability, drug loading capacity, and mucosal permeation efficiency. The absence of standardized validation frameworks linking in silico predictions to in vitro and in vivo outcomes restricts broader translational adoption of this approach.

Therefore, there is a compelling need to establish a structured, predictive preformulation strategy that integrates COSMO-RS modeling with experimental optimization to rationalize solvent and excipient selection, reduce experimental burden, and enhance formulation robustness. Bridging this gap could accelerate development timelines and provide a scientifically grounded framework for designing advanced drug delivery systems.

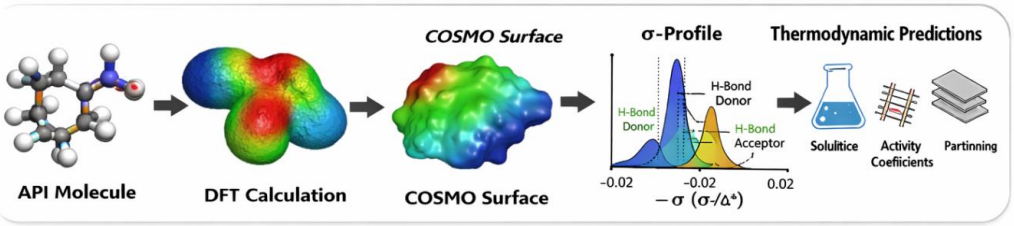
2. Theoretical Foundation of COSMO-RS

2.1 Origin and Conceptual Basis

The Conductor-like Screening Model for Real Solvents (COSMO-RS) was developed by Andreas Klamt in the 1990s as an extension of the Conductor-like Screening Model (COSMO) [5,6]. COSMO-RS integrates quantum chemical surface polarization calculations with statistical thermodynamics to predict thermodynamic equilibrium properties of fluids and mixtures without requiring extensive experimental parameterization.

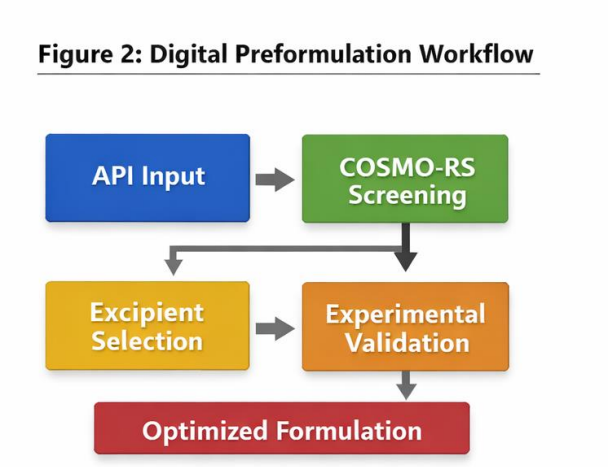
In the initial COSMO step, an isolated molecule is embedded within a virtual conductor environment. Under this condition, polarization charges are induced on the molecular surface to screen the internal electric field. These screening charge densities (σ) are computed using density functional theory (DFT), typically at well-established quantum chemical levels of theory. The resulting surface charge density distribution—commonly referred to as the COSMO surface—serves as a molecular descriptor encoding polarity, hydrogen-bonding capability, and electrostatic character.

Figure 1: COSMO-RS Modeling Principle



The COSMO surface contains spatially resolved positively and negatively charged regions that reflect donor and acceptor functionalities of the molecule. These surface segments constitute the fundamental basis for subsequent thermodynamic interaction calculations within the COSMO-RS framework.

Figure 2: Digital Preformulation Workflow



To contextualize the positioning of COSMO-RS within pharmaceutical preformulation workflows, a comparative analysis of commonly applied predictive and empirical approaches is presented in Table 1. This comparison highlights the mechanistic advantages and practical limitations of each modeling framework in early-stage formulation development.

Table 1:

Approach	Theoretical basis	Parameters predicted	Strengths	Limitation
COSMO-RS	Quantum chemistry + statistical thermodynamics	Solubility, pKa,permeability, drug excipient compatibility (activity coefficient and excess enthalpy)	Mechanistic, no system-specific fitting required, applicable to novel APIs	Does not capture kinetics, requires quantum calculations, limited solid-state correct
UNIFAC	Group contribution thermodynamic model	Predicts solubility in organic solvents and limited drug excipient compatibility	Fast, low computational cost	Requires parameterized groups, limited for novel chemistries
PC-SAFT	Equation-of-state thermodynamic model	Predicts solubility,pKa, drug excipient compatibility and limited permeability	High thermodynamic rigor	Complex parameterization, computationally intensive

2.2. Integrated Digital Preformulation for Model Formulation Design

Stage 1: API Input & Molecular Characterization

Chemical structure (SMILES / 3D structure)

Conformer generation

DFT optimization

Generation of COSMO file

Sigma profile (σ-profile)

Stage 2: Thermodynamic Property Prediction (COSMO-RS Core)

Solubility prediction (pure + mixed solvents)

Activity coefficients ($\gamma\infty$)

Excess enthalpy of mixing (Hex)

Partition coefficients (logP/logD)

pKa estimation

Hansen solubility parameters

Stage 3: Computational Screening Cascade

A. Oil Phase Ranking

→ Activity coefficient ranking

→ Solubility ranking

B. Surfactant Screening

→ Sigma profile complementarity

→ Hydrogen bonding potential

C. Co-solvent Optimization

→ Sigma potential similarity analysis

→ Solvent equivalency mapping

Stage 4: Design Space Reduction

Narrow excipient pool (from ~20 candidates to 3–5)

Eliminate incompatible systems

Define thermodynamically favorable combinations

Stage 5: Targeted Experimental Validation

Shake-flask solubility

Compatibility testing (DSC / FTIR)

Nanoemulsion optimization (DoE)

Droplet size & PDI analysis

Stability studies

Stage 6: Feedback Loop

Compare predicted vs experimental data

Refine computational assumptions

Update digital model

Stage 7: Model-Informed Formulation Strategy

Optimized formulation selection

Reduced development timeline

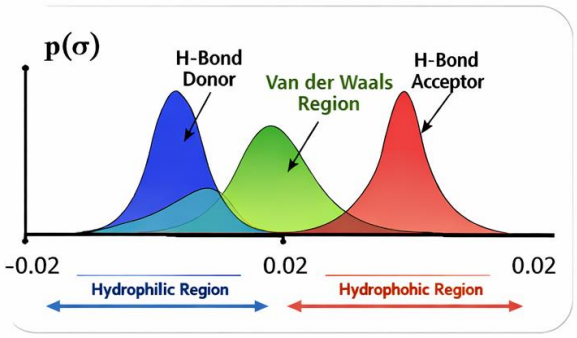
Reduced material consumption

Regulatory documentation support

2.3 Sigma Profile (σ-Profile)

The σ-profile represents the probability distribution of surface charge densities across the molecular surface and serves as the central descriptor in COSMO-RS statistical thermodynamics. It quantitatively describes how much of the molecular surface exhibits specific charge density values and thereby captures polarity distribution and hydrogen-bonding potential.

Figure 3: Sigma Profile Interaction



Molecular interactions in solution are calculated through pairwise interactions between surface segments of contacting molecules. These interactions are decomposed into three principal energetic contributions:

1. Electrostatic misfit interactions (resulting from charge density differences between contacting surface segments)
2. Hydrogen-bonding interactions (arising from donor–acceptor complementarity)
3. Van der Waals interactions (representing dispersion forces)

By statistically averaging these pairwise surface interactions under thermodynamic equilibrium conditions, COSMO-RS predicts macroscopic properties from molecular-level descriptors [6–8].

2.4 Statistical Thermodynamics Framework

COSMO-RS applies a statistical thermodynamic treatment based on several simplifying but mechanistically justified assumptions:

- Close contact between molecular surfaces in the liquid phase
- Pairwise surface segment interactions
- Neglect of explicit three-dimensional structural orientation

- Thermodynamic equilibrium conditions

Within this framework, the chemical potential (μ) of a compound in solution is computed. From the calculated chemical potential, a range of macroscopic thermodynamic properties can be derived, including Activity coefficients (γ), Solubility, Partition coefficients, Phase equilibria and Excess enthalpy. This multiscale translation from quantum chemical surface descriptors to bulk thermodynamic properties enables COSMO-RS to bridge molecular structure and formulation-relevant behavior without empirical fitting to each new system.

3. Computational Workflow

The COSMO-RS workflow consists of a structured multistep computational cascade:

1. Quantum chemical calculation (e.g., DFT using TURBOMOLE) to generate the COSMO file and surface charge distribution
2. Generation of COSMO surface files representing σ-surfaces
3. Thermodynamic evaluation using COSMOtherm software
4. Prediction of target properties, such as solubility, logP, pKa, activity coefficients, and phase behavior
5. For high-throughput screening applications, COSMOquick employs fragment-based σ-profile approximations, significantly reducing computational cost while maintaining reasonable predictive accuracy [9]. This enables rapid preliminary solvent and excipient screening during early-stage preformulation.

4. Applications in Pharmaceutical Preformulation

4.1 Solubility Prediction

One of the most established pharmaceutical applications of COSMO-RS is the prediction of active pharmaceutical ingredient (API) solubility in Pure solvents, Binary solvent mixtures, Co-solvent systems and Deep eutectic solvents

Multiple studies report strong agreement between predicted and experimentally measured solubility trends, particularly for ranking solvent suitability and identifying optimal solubilizing systems [10–12]. While absolute solubility values may vary depending on solid-state considerations, COSMO-RS demonstrates robust capability in predicting relative solubility behavior and solvent compatibility.

By narrowing the candidate solvent pool prior to laboratory validation, COSMO-RS substantially reduces experimental solvent screening workload, material consumption, and development timelines.

4.2 Solvent Screening and Formulation Equivalency

COSMO-based similarity analysis enables comparison of sigma-potentials (σ-potentials) between solvent systems. Solvent mixtures exhibiting similar sigma-potentials possess comparable intermolecular interaction characteristics and thermodynamic behavior.

This approach facilitates identification of alternative solvent systems that are thermodynamically equivalent to target formulations, supporting rational excipient substitution, reformulation under regulatory constraints, cost optimization, green solvent selection. Such predictive equivalency analysis reduces dependency on empirical reformulation cycles and supports data-driven formulation decisions [13].

4.3 Drug–Excipient Compatibility

COSMO-RS provides mechanistic insight into drug–excipient miscibility through calculation of activity coefficients and excess enthalpy of mixing. These parameters enable prediction of compatibility between APIs and lipids, polymers, surfactants and plasticizers.

Highly negative excess enthalpy values indicate favorable intermolecular interactions and strong miscibility potential, whereas positive values may suggest phase separation or limited compatibility [14]. This predictive capability is particularly relevant for lipid-based systems, nanoemulsions, and polymeric drug delivery platforms where molecular-level interactions determine formulation stability.

4.4 Hansen Solubility Parameter (HSP) Prediction

COSMO-RS can also be employed to predict Hansen Solubility Parameters (HSPs), which support rational polymer and excipient selection in amorphous solid dispersions, transdermal systems and semi-solid dosage forms.

Reported predictive accuracies range between 90–95% when compared with experimentally derived HSP values [15]. Integration of COSMO-RS-derived HSP predictions into formulation workflows enhances rational excipient screening and reduces trial-and-error polymer selection.

4.5 Cocrystal Screening

Cocrystal formation is evaluated via mixing enthalpy calculations. Strong negative excess enthalpy indicates favorable heteromolecular interactions [16]. This approach supports early coformer screening prior to crystallization trials.

4.6 pKa Prediction

Accurate pKa prediction requires COSMO files of both neutral and ionized species. Reported deviations from experimental values are often within ±0.2 pKa units [17].This supports salt selection, solubility modeling and absorption prediction

4.7 Activity Coefficients

Activity coefficients quantify deviation from ideal behavior. Infinite dilution activity coefficients support drug loading estimation, release modeling and solvent mixture optimization.

4.8 Partitioning and Permeability

COSMO-based membrane models (e.g., COSMOperm) simulate solute distribution across lipid bilayers [18]. Applications include logP/logD prediction, passive permeability estimation and early ADME profiling

5. Critical Limitations of COSMO-RS in Pharmaceutical Applications

Despite its strong predictive capability and thermodynamic rigor, COSMO-RS presents several limitations that must be critically considered when applying the model to pharmaceutical formulation design.

First, COSMO-RS primarily predicts equilibrium thermodynamic properties and does not inherently account for kinetic phenomena such as crystallization rate, nucleation behavior, supersaturation maintenance, or metastable phase formation. These factors can significantly influence experimentally observed solubility and formulation stability, particularly for poorly water-soluble APIs [5–7].

Second, the model neglects explicit three-dimensional structural orientation and long-range molecular organization. While the pairwise surface segment interaction approach is statistically robust, it may not fully capture complex supramolecular assemblies, micellization phenomena, or interfacial curvature effects relevant to colloidal systems such as nanoemulsions.

Third, solid-state effects including polymorphism, hydrate formation, and amorphous transitions are not intrinsically incorporated into standard COSMO-RS solubility predictions. Since pharmaceutical solubility is governed by both solvation thermodynamics and crystal lattice energy, discrepancies between predicted and experimental absolute solubility values may arise if solid-state corrections are not included.

Fourth, predictions involving highly structured biological environments (e.g., mucus layers, epithelial membranes) remain indirect. COSMO-RS evaluates thermodynamic compatibility and intermolecular interactions but does not explicitly simulate biological transport mechanisms or dynamic interfacial processes.

Therefore, while COSMO-RS is a powerful screening and ranking tool, it should be integrated within a hybrid workflow combining computational prediction with targeted experimental validation rather than used as a standalone decision-making system.

6. Integration of COSMO-RS with Nanoemulsion Formulation Strategy

Nanoemulsions represent thermodynamically and kinetically complex colloidal systems composed of oil, surfactant, co-surfactant, and aqueous phases. The stability, droplet size distribution, drug loading efficiency, and mucosal interaction behavior of nanoemulsions are fundamentally governed by molecular-level interactions among these components.

COSMO-RS provides a rational framework to support nanoemulsion design at multiple levels:

6.1 Oil Phase Selection

Solubility prediction and activity coefficient calculations enable ranking of candidate lipid excipients (e.g., medium-chain triglycerides, oleic acid, ethyl oleate) based on predicted API affinity [10–12,14]. Oils exhibiting lower activity coefficients and favorable excess enthalpy values are expected to enhance drug loading capacity and reduce risk of precipitation upon dilution.

6.2 Surfactant and Co-Surfactant Screening

Sigma-profile analysis allows evaluation of complementary polarity between API and surfactants. Surfactants with compatible hydrogen-bond donor–acceptor balance and electrostatic profiles may improve solubilization and interfacial stabilization. COSMO-based similarity analysis can also identify thermodynamically equivalent surfactant systems, supporting rational substitution or optimization [13].

6.3 Predicting Miscibility and Stability

Excess enthalpy calculations provide insight into the likelihood of phase separation or instability. Strongly negative mixing enthalpies suggest favorable molecular interactions that may contribute to long-term physical stability in nanoemulsion systems [14].

6.4 Reducing Experimental Design Space

By pre-screening oils, surfactants, and co-solvents computationally, COSMO-RS can significantly narrow the experimental design space prior to factorial optimization or response surface methodology studies [32–34]. This hybrid strategy aligns with modern model-informed drug development paradigms and enhances formulation efficiency.

Importantly, while COSMO-RS does not directly predict droplet size or interfacial curvature, it provides a thermodynamic compatibility framework that informs rational component selection prior to high-energy or low-energy emulsification processes.

7. Research Gap and Positioning within Computational Pharmaceutics

Although COSMO-RS has demonstrated strong capability in solvent solubility prediction and excipient compatibility assessment [10–15], its integration into structured nanoemulsion formulation workflows remains limited. Current literature predominantly applies COSMO-RS as an isolated solvent screening tool rather than embedding it within systematic colloidal formulation design strategies.

Specifically, several critical gaps remain:

1. Lack of correlation between COSMO-RS–predicted thermodynamic parameters (activity coefficient, excess enthalpy, sigma similarity) and experimentally measured nanoemulsion performance indicators such as droplet size, polydispersity index, drug loading efficiency, and physical stability.
2. Limited application of COSMO-RS in intranasal nanoemulsion systems, where molecular interactions influence not only solubilization but also mucosal permeation and retention.
3. Absence of standardized hybrid computational–experimental frameworks integrating COSMO-RS screening with design of experiments (DoE) optimization for advanced drug delivery systems.
4. Minimal translational linkage between in silico thermodynamic predictions and in vitro/in vivo performance outcomes.

Therefore, there is a compelling need to establish an integrated computational pharmaceutics strategy that combines COSMO-RS–guided excipient selection with experimental nanoemulsion optimization and biological evaluation. Such an approach would reduce empirical trial-and-error experimentation, enhance mechanistic understanding of formulation behavior, improve resource efficiency and accelerate development timelines for poorly soluble APIs

The present study is positioned within this emerging field of computational pharmaceutics, aiming to bridge molecular-level thermodynamic modeling with practical nanoemulsion formulation development. By correlating COSMO-RS–derived interaction parameters with experimental formulation performance, this work seeks to advance rational, model-informed design of intranasal nanoemulsion systems.

6. Advantages Over Traditional Screening

- No experimental input required
- Mechanistic thermodynamic basis
- Applicable to novel chemical entities
- Reduced material consumption
- Accelerated decision-making

7. Limitations

- Dependent on conformer quality
- Requires accurate quantum chemical descriptors
- Complex hydrogen bonding systems may need validation
- Regulatory acceptance evolving
- Integration with experimental validation remains recommended.

8. Future Perspectives

Emerging directions include:

- AI-assisted conformer selection
- Integration into digital twins
- Coupling with machine learning QSPR frameworks
- Model-informed regulatory submissions
- Digital preformulation platforms integrating COSMO-RS may substantially reduce early-stage attrition.

9. CONCLUSION

COSMO-RS represents a powerful thermodynamic modeling framework for pharmaceutical preformulation. By combining quantum chemical surface descriptors with statistical thermodynamics,

it enables predictive solubility, compatibility, permeability, and ionization modeling. The workflow cascade approach balances speed and accuracy. As computational and regulatory frameworks mature, COSMO-RS-based platforms are positioned to become central tools in model-informed pharmaceutical development.

REFERENCES

1. Sinko PJ. *Martin's physical pharmacy and pharmaceutical sciences*. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2012.
2. Dressman JB, Reppas C. In vitro–in vivo correlations for lipophilic, poorly water-soluble drugs. Eur J Pharm Sci. 2017;101:234–245.
3. Savjani KT, Gajjar AK, Savjani JK. Drug solubility: importance and enhancement techniques. ISRN Pharm. 2012;2012:195727.
4. Kawabata Y, Wada K, Nakatani M, Yamada S, Onoue S. Formulation design for poorly water-soluble drugs based on biopharmaceutics classification system: basic approaches and practical applications. Int J Pharm. 2011;420:1–10.
5. Klamt A. Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. J Phys Chem. 1995;99:2224–2235.
6. Klamt A. *COSMO-RS: from quantum chemistry to fluid phase thermodynamics and drug design*. Weinheim: Wiley-VCH; 2005.
7. Eckert F, Klamt A. Fast solvent screening via quantum chemistry: COSMO-RS approach. AIChE J. 2002;48:369–385.
8. Hsieh CM, Sandler SI, Lin ST. Predictions of drug solubility using COSMO-RS. Fluid Phase Equilib. 2014;367:1–6.
9. Lamoureux G, et al. A fast fragment-based approach for COSMO-RS predictions. J Chem Inf Model. 2006;46:1467–1473.
10. Granberg RA, Rasmuson ÅC. Solubility of pharmaceutical compounds in organic solvents. J Chem Eng Data. 1999;44:1391–1395.
11. Perlovich GL, Kurkov SV, Hansen LK, Bauer-Brandl A. Thermodynamic study of drug solubility in organic solvents. Eur J Pharm Sci. 2012;45:557–564.
12. Abraham MH, Le J, Acree WE Jr, Carr PW. Solubility and partitioning of pharmaceutical compounds. J Pharm Sci. 2010;99:1500–1515.
13. Klamt A, Eckert F. COSMO-RS: a novel and efficient method for the prediction of thermophysical data of liquids. Fluid Phase Equilib. 2000;172:43–72.
14. Jouyban A. Review of the cosolvency models for predicting solubility. J Pharm Pharm Sci. 2008;11:32–58.

15. Hansen CM. *Hansen solubility parameters: a user's handbook*. 2nd ed. Boca Raton: CRC Press; 2007.
16. Childs SL, Stahly GP, Park A. The salt–cocrystal continuum: the influence of crystal structure on solubility. Mol Pharm. 2007;4:323–338.
17. Llinàs A, Glen RC, Goodman JM. Solubility challenge: can you predict solubility? J Med Chem. 2008;51:211–217.
18. Bittermann K, Goss KU, Sabljić A. Prediction of water solubility using COSMO-RS. J Phys Chem B. 2014;118:14833–14842.