



Article

Solvent Screening for Solubility Enhancement of Theophylline in Neat, Binary and Ternary NADES Solvents: New Measurements and Ensemble Machine Learning

Piotr Cysewski ^{*} , Tomasz Jeliński, Patryk Cymerman and Maciej Przybyłek

Department of Physical Chemistry, Faculty of Pharmacy, Collegium Medicum of Bydgoszcz, Nicolaus Copernicus University in Toruń, Kurpińskiego 5, 85-950 Bydgoszcz, Poland; tomasz.jelinski@cm.umk.pl (T.J.); patryk.cymerman@onet.pl (P.C.); m.przybylek@cm.umk.pl (M.P.)

* Correspondence: Piotr.Cysewski@cm.umk.pl

Abstract: Theophylline, a typical representative of active pharmaceutical ingredients, was selected to study the characteristics of experimental and theoretical solubility measured at 25 °C in a broad range of solvents, including neat, binary mixtures and ternary natural deep eutectics (NADES) prepared with choline chloride, polyols and water. There was a strong synergistic effect of organic solvents mixed with water, and among the experimentally studied binary systems, the one containing DMSO with water in unimolar proportions was found to be the most effective in theophylline dissolution. Likewise, for NADES, the addition of water (0.2 molar fraction) resulted in increased solubility compared to pure eutectics, with the highest solubilisation potential offered by the composition of choline chloride with glycerol. The ensemble of Statistica Automated Neural Networks (SANNs) developed using intermolecular interactions in pure systems has been found to be a very accurate model for solubility computations. This machine learning protocol was also applied as an extensive screening for potential solvents with higher solubility of theophylline. Such solvents were identified in all three subgroups, including neat solvents, binary mixtures and ternary NADES systems. Some methodological considerations of SANNs applications for future modelling were also provided. Although the developed protocol is focused exclusively on theophylline solubility, it also has general importance and can be used for the development of predictive models adequate for solvent screening of other compounds in a variety of systems. Formulation of such a model offers rational guidance for the selection of proper candidates as solubilisers in the designed solvents screening.

Keywords: theophylline; solubility; machine learning; ensemble neural networks; COSMO-RS; NADES; binary solvents



Citation: Cysewski, P.; Jeliński, T.; Cymerman, P.; Przybyłek, M. Solvent Screening for Solubility Enhancement of Theophylline in Neat, Binary and Ternary NADES Solvents: New Measurements and Ensemble Machine Learning. *Int. J. Mol. Sci.* **2021**, *22*, 7347. <https://doi.org/10.3390/ijms22147347>

Academic Editors: Uko Maran and Csaba Hetényi

Received: 17 May 2021

Accepted: 6 July 2021

Published: 8 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Methylxanthines belong to the purine alkaloids and, as such, they comprise a fused heterocyclic system with pyrimidine and imidazole rings. Specifically, the studied compound, i.e., theophylline (T), is a bi-substituted derivative of xanthine (1,3-dimethylxanthine). Methylxanthines are abundant in nature and can be found in tea and other plant leaves, coffee and cocoa beans, as well as cola seeds [1]. Theophylline is mostly found in cocoa beans, with trace amounts in other sources [2]. It was chemically identified for the first time in 1888 and synthesised seven years later [1]. Methylxanthines play an important role in different biological processes, and their biological activities include the stimulation of the central nervous system, increased blood pressure, kidney diuresis, relaxation of smooth muscles, strengthening of the concentration of skeletal muscles and gastric acid secretion [3,4]. The mode of action of methylxanthines arises from their activities as phosphodiesterase inhibitors [5] and nonselective adenosine receptor antagonists [6]. Theophylline, in particular, is used to treat asthma, chronic obstructive pulmonary disease and neonatal apnoea [7–12]. It is usually administered in the form of a capsule or tablet

and is extensively metabolised in the liver, with only a small part excreted in urine [13]. Theophylline has a low octanol-water partition coefficient (P) and is sparingly soluble in water [1].

The problem of the limited solubility of chemical compounds in water and various other solvents poses a challenge to many areas of modern industry, notably including pharmaceuticals. Many substances are regarded as insoluble in aqueous solutions [14–16], which makes it necessary to develop techniques that would overcome this difficulty. Having the above in mind, several different approaches have been offered to address poor solubility. Among the most often utilized are micronization [17], monocrystal formation [18,19], amorphization [20–22], complexation with cyclodextrins [23,24], solid dispersion formation [25], pH modification [26,27], formation of salts [28,29] and cocrystals [30,31]. The usage of hydrotropes or cosolvation techniques also offers an interesting and relatively simple way of enhancing solubility [32,33]. The cosolvation effect takes place when a particular amount of a cosolvent is added to the primary solvent and causes a solubility increase [34]. Among the organic compounds primarily used for this task, one can include alcohols, DMSO, propylene glycol or glycerin [35,36]. However, the usage of organic solvents may have limited applicability due to their potential toxicity. This generates a need for alternative solvation media that would combine good solubilising potential with environmental and health safety. These simultaneous requirements focus the attention of many researchers on natural deep eutectic solvents (NADES). The key feature of deep eutectic solvents (DES) is the lowering of their melting point in comparison with the corresponding individual components. This enables them to form liquids even at room temperature [37]. Natural deep eutectic solvents are a specific part of this larger group of compounds and can be defined as bio-based DES prepared using such compounds as organic acids, alcohols, amino acids, sugars, choline chloride (ChCl) or other plant-based primary metabolites [38,39]. The desired physicochemical properties of NADES, including low volatility, sustainability and biodegradability, low cost and simplicity of preparation, as well as the potential of fine-tuning for specific applications, make these solvents very favourable in terms of efficiency, safety and economics [40–43]. All of this leads to the widespread usage of natural deep eutectic solvents in many different applications [42,44–49].

During the process of extraction, synthesis, evaluation and optimisation of active pharmaceutical ingredients (APIs), a large volume of solubility data is required, both in neat solvents as well as their mixtures [36]. For example, when considering cosolvation, it would be beneficial to know which potential cosolvents are the most promising before conducting actual experiments. These considerations mean that a large number of experiments are needed to supply these data, which in turn results in a large amount of waste originating from a number of processes, including synthesis, purification and analysis. Therefore, it would be beneficial from an economical, ecological and environmental perspective to utilise such methods that would provide the necessary data but would simultaneously reduce the amount of environmental pollution. The answer to this problem might be the application of computational methods that would direct and partially replace the actual experiments. Unfortunately, to date, there are few reliable models enabling the estimation of solubility from the molecular structure of API and solvent molecules. Although many theoretical approaches were proposed [35,50–52], their usage is seriously prohibited by inadequate accuracy or restricted applicability domains resulting from limited training sets. That is why post-factum models, which use some equations with parameters fitted to experimental data, are so popular [53–57]. Furthermore, several alternative approaches were proposed to utilise physicochemical properties as molecular descriptors; for example, the general solubility equation proposed by Jain and Yalkowsky [58] utilises values of the melting point and the octanol-water partition coefficient. Moreover, the set of models developed by Abraham and coworkers [59] is worth mentioning; this type of approach allows for predicting the relative solubility with respect to a selected reference solvent, which is typically water. The set of equations proved to be very effective in solubility computations not only in organic solvents [60,61] but also in ionic liquids [62] after parametrisation with the aid of linear

free energy relationships. On the other hand, *ab initio* approaches were developed for a priori characteristics of chemical systems without external parametrisation. Among them, one of the outstanding is the COSMO-RS model [63,64], extensively utilised for predicting solubility [65,66], solvent screening [67–69] and multiphase chemical equilibria [70–72] such as SLE, LLE, VLE, SLLE, etc. It combines density functional theory calculations with statistical thermodynamics as a post-treatment [73] and has become a standard approach in many scientific and industrial applications. Furthermore, dissolution in NADES [74,75] was studied using this theoretical framework. Unfortunately, the accuracy of this approach is very often not sufficient, not only for reproducing experiment data but also for screening purposes. That is why the advantages of machine learning [76] in general and the artificial neural networks (ANNs) [77] particularly have become a very powerful tool for solubility predictions proving their efficiency in many applications [78–82]. Although such non-linear modelling sacrifices the analytical solution, it can address serious accuracy problems, as previously documented [83]. That is why the effort was undertaken to effectively combine the machine learning protocol with molecular descriptors characterising properties of the solute in a variety of solvents for a quantitative API solubility computation. Although here the solubility of theophylline is addressed, the proposed procedure is general and can be easily applied to any API in virtually any solvent, including screening of new hydrotropic and green solvents. There are three main features of the protocol applied here. The starting point was the collection of available experimental solubility data of theophylline in the literature, which were seriously extended with new measurements. It is worth emphasising that quite a unique set of data was composed. It encompasses not only the values characterising solubility in neat solvents and five aqueous binary mixtures with organic liquids but also a set of NADES containing choline chloride mixed with several other compounds. In the second stage, the machine learning protocol was used for the development and validation of an ensemble of artificial neural networks. Finally, extensive screening was performed for finding the most promising solvents in all three categories, which led to proposing new solubility enhancers.

2. Results and Discussion

The screening protocol for solvents with enhanced solubility was divided into three steps. First, the data set of theophylline solubility was prepared by the collection of published data and seriously extended with new measurements. Then, using machine learning protocols, the non-linear model was developed by formulation and validation of an ensemble of artificial neural networks. Finally, extensive screening was performed for finding new promising solvents in all three categories as potential solubility enhancers.

2.1. Theophylline Solubility

The starting point was the experimental determination of theophylline solubility in eleven neat organic solvents at room temperature. This enabled the initial solvent screening and preselection for the binary mixture solubility measurements. The resulting solubilities expressed as mole fraction x_T , that were collected are shown in Figure 1. It happened that the highest solubilities were found for DMSO ($x_T = 70.96 \pm 0.28 \times 10^{-4}$) and DMF ($x_T = 59.21 \pm 0.43 \times 10^{-4}$). A slightly lower solubility of theophylline was observed for neat 1,4-dioxane ($x_T = 31.67 \pm 0.58 \times 10^{-4}$), with other most effective solvents being methanol, 1-propanol and 1-butanol ($x_T = 13.64 \pm 0.04 \times 10^{-4}$, $13.39 \pm 0.13 \times 10^{-4}$ and $10.23 \pm 0.11 \times 10^{-4}$, respectively). On the other hand, as expected, the smallest solubility was obtained for pure water ($x_T = 6.10 \pm 0.12 \times 10^{-4}$) at ambient conditions. The five most effective solvents were selected for the next stage of experiments aiming for the determination of solubility in aqueous binary mixtures. It is worth mentioning that some of our measurements intentionally repeated the already available solubility data in four neat solvents. This allowed for direct comparisons of the accuracy of the performed experiments with published data. As demonstrated in Figure 1, our results are consistent with the previously determined solubility of theophylline by Zhong et al. [84]. A

detailed comparison of our measurements with already published data is provided in the Supplementary Materials (see Table S2). It was found that the measurements presented in this work are in good accordance with the literature data, with a mean relative difference as low as 1.9%.

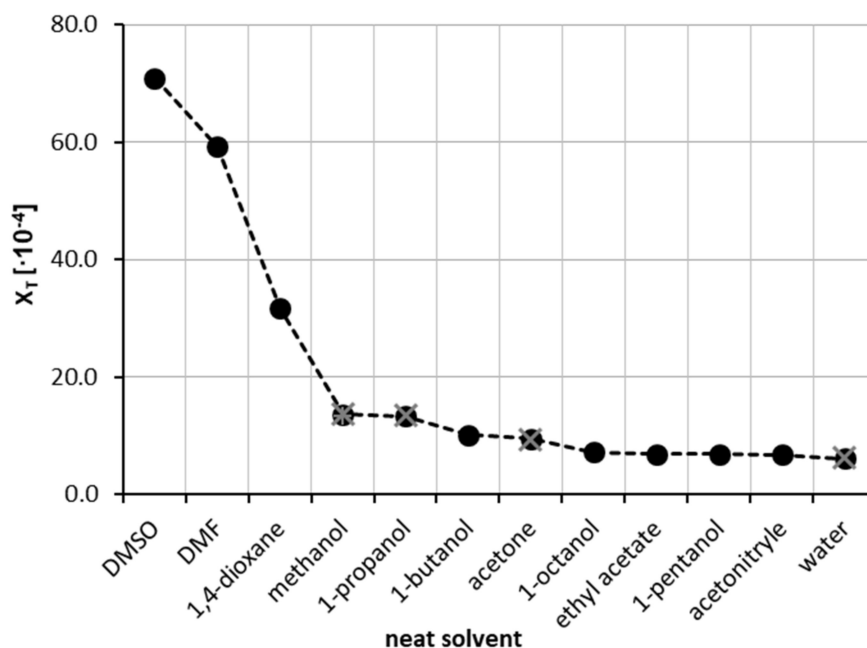


Figure 1. Solubility of theophylline at 25 °C, expressed as its mole fraction, x_T , in eleven neat organic solvents and water. Black circles represent measurements of this work, while grey crosses indicate values published by Zhong et al. [84].

In the next step, the solubility of theophylline was determined in five binary mixtures comprising DMSO, DMF, 1,4-dioxane, 1-propanol and 1-butanol mixed with water in twelve molar ratios covering the whole range of concentrations. The results of these measurements are presented in Figure 2. The most interesting observation is that all of the studied organic solvents exhibit synergistic behaviour since the highest solubility was obtained when the binary mixture comprised water and solvent in unimolar proportions. Likewise, all of the considered solvents can be regarded as good cosolvents since a systematic and significant solubility increase is observed with the rise of the organic solvent content in water solutions. The highest solubility at the point corresponding to $x_2^* = 0.5$ was found to be equal to $x_T = 90.13 \pm 0.56 \times 10^{-4}$ and $x_T = 76.41 \pm 0.58 \times 10^{-4}$ for DMSO and DMF, respectively. In the case of 1,4-dioxane, the synergistic effect was slightly less pronounced, and the largest solubility was found to be equal to $x_T = 62.37 \pm 0.45 \times 10^{-4}$. For the last two solvents, namely 1-butanol and 1-propanol, the largest solubility at unimolar solvent composition was found to be equal to $x_T = 43.47 \pm 0.22 \times 10^{-4}$ and $x_T = 35.03 \pm 0.08 \times 10^{-4}$, respectively. Interestingly, although the solubility of theophylline in neat 1-propanol was larger than in 1-butanol, for unimolar compositions with water, the order is reversed. Detailed results of theophylline solubility in both neat organic solvents and binary mixtures with water were collected in Tables S3 and S4 in Supplementary Materials.

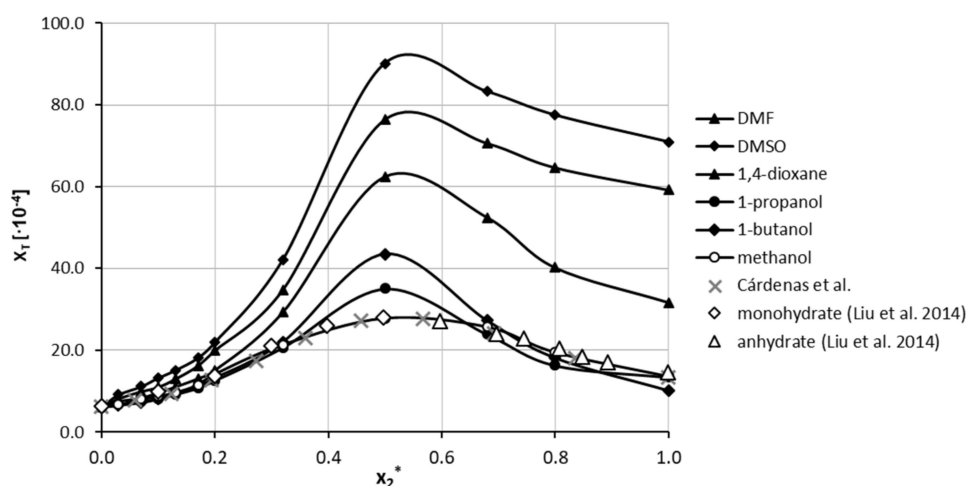


Figure 2. Solubility of theophylline at 25 °C, expressed as its mole fraction, x_T , in binary solvents comprising water and five different organic solvents in varying compositions. On the ordinate, x_2^* represents the mole fractions of organic solvent in solute-free binary solutions. The values of Theophylline solubility in aqueous methanolic solutions published by Cárdenas et al. [85] were marked with grey crosses and are anhydrous. Additionally, monohydrate and anhydrate theophylline solubilities measured by Liu et al. [86] were plotted as open diamonds and triangles.

The third series of measurements was devoted to the extension of the experimental pool of theophylline solubility by including some designed solvents. Hence, seven natural deep eutectic solvents containing choline chloride and one of the selected polyols were prepared by mixing both the cationic and anionic counterparts in unimolar proportions. This led to the identification of the NADES constituents, which are the most suited as potential theophylline solubility enhancers. The obtained results are presented in Figure 3. The highest solubility of theophylline was found in the NADES made of choline chloride and glycerol. The measured mole fraction in saturated solution was equal to $x_T = 128.17 \pm 0.68 \times 10^{-4}$ and enhanced solubility 21 times compared to pure water and was about 80% higher with respect to pure DMSO. This observation provides direct evidence of the effectiveness of NADES as a proficient solubiliser of theophylline. Other polyols such as sorbitol and xylitol were not as efficient, resulting in lower theophylline solubility, and were equal to $x_T = 103.63 \pm 1.20 \times 10^{-4}$ and $x_T = 92.61 \pm 0.94 \times 10^{-4}$, respectively. NADES made of glucose, and fructose resulted in slightly lower solubility of theophylline, i.e., $x_T = 73.45 \pm 0.75 \times 10^{-4}$ and $x_T = 67.79 \pm 0.72 \times 10^{-4}$, respectively. It is worth noting that NADES formulated with the first four constituents gave solubility values of theophylline greater than in DMSO, which was characterised by the highest solubility among studied aqueous organic solvents. Taking this into account, the four most effective NADES were selected as cosolvents in mixtures with water, analogically to the earlier procedure with neat organic solvents. Hence, NADES with kept proportions of cationic and anionic counterparts were successfully diluted with water in the whole range of concentrations. This procedure allowed for determining the effectiveness of NADES as cosolvents for theophylline solubilisation, which was summarised in Figure 4. Additionally, the synergistic effect was observed, corresponding to the mixtures containing $x_2^* = 0.8$ of designed solvents. The order of increased solubility observed for pure NADES solutions was also preserved in all studied mixtures with water. In the case of NADES comprising choline chloride and glycerol, the highest solubility of theophylline was found to be $x_T = 146.57 \pm 0.72 \times 10^{-4}$, which stands for 24 times higher solubility with respect to pure water and 62% higher than the most effective binary mixture. When sorbitol and xylitol were used as the second NADES constituent, the resulting solubility was equal $x_T = 123.83 \pm 0.70 \times 10^{-4}$ and $x_T = 113.94 \pm 0.31 \times 10^{-4}$, respectively. Finally, for the NADES made with glucose, the solubility in the optimal composition was $x_T = 93.76 \pm 0.96 \times 10^{-4}$. When comparing pure NADES and NADES-mixtures with water, the increase in theophylline

solubility in the most effective composition is less pronounced than it was for the binary solvents tested earlier, although the solubility values are still greater than for DMSO. For the NADES with sorbitol, xylitol and glucose as the second constituent, the solubility of theophylline in the NADES-water mixture with $x_2^* = 0.8$ is also larger than in pure NADES. It is thus expected that a modest amount of water is a promoting factor of theophylline solubility in organic solutions. Detailed results of theophylline solubility in both pure NADES and NADES-water mixtures can be found in Tables S5 and S6 in Supplementary Materials.

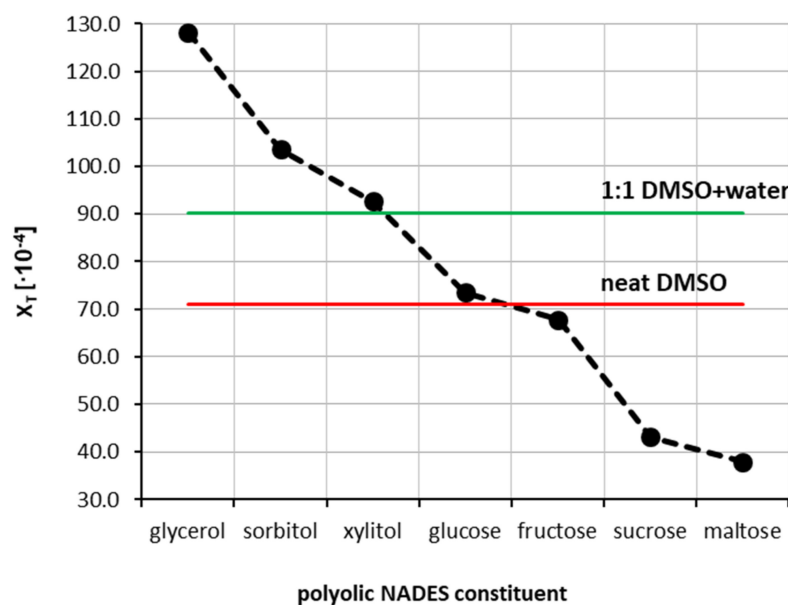


Figure 3. Mole fraction solubility of theophylline determined in water-free natural deep eutectic solvents at 25 °C. All designed solvents were prepared in unimolar proportions.

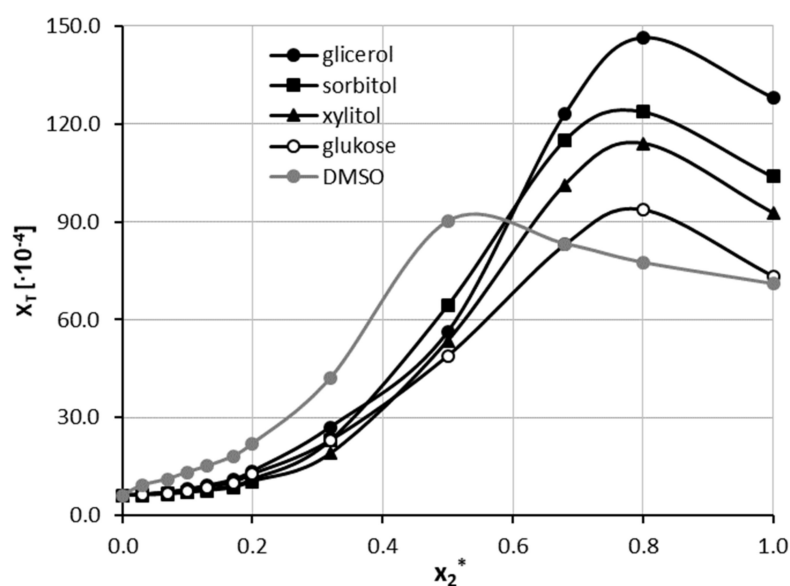


Figure 4. Solubility profiles of theophylline at 25 °C in mixtures comprising NADES successfully diluted with water. On the ordinate, x_{NADES}^* , represents values of the mole fractions of Natural Deep Eutectic Solvent in aqueous solutions. For comparison purposes, solubility in DMSO+water was plotted as a function of x_2^* (grey line).

2.2. Precipitate Characteristics

Theophylline belongs to a class of organic solutes that exhibit rich structural diversity in the solid-state. The commercial theophylline is usually offered in the form of polymorph II since it easily crystallises from many neat organic solvents [87–89], and for a long time, it was considered to be the most stable form thermodynamically at room temperature [87,90,91]. However, it was documented by Seton et al. [89] that form IV is more stable at ambient conditions, although the II \rightarrow IV conversion encounters a kinetic barrier significantly slowing the spontaneous transformation [89]. Form I is stable at higher temperatures [87], and form III is highly metastable at any temperature [92]. Hence, it is expected that the precipitate of theophylline from solutions under the solubility measurement conditions adopts the form II. However, theophylline can also exist as a stable monohydrate at conditions of sufficiently high water activity in the crystallisation medium [93]. For example, a spontaneous transformation of theophylline anhydrate into monohydrate occurs in contact with an aqueous methanolic solution if the water activity exceeds 0.252 [93]. On the contrary, anhydrous theophylline remained unchanged at conditions with a water activity below 0.227. This was also studied in detail [94] by precise determination of the transformation point of solution mediated transformation of monohydrate to anhydrate in form II.

To further explore this aspect, the mechanochemical synthesis was used for theophylline wetted with pure water, methanol and aqueous methanolic mixtures. The resulting thermograms of co-grounded samples provided in Figure 5 confirm the existence of a hydrate in water-rich systems since a characteristic small peak appeared at the low-temperature values. This endothermic process is typical for the theophylline-water complex [95,96]. As one can see, the commercial theophylline and the same compound in the presence of pure methanol adopt an anhydrous form, as clearly evidenced by the lack of peak around 340 K on lines (1) and (3) plotted in Figure 5. However, milling in the presence of pure water and aqueous methanol mixtures with a modest concentration of organic solvent results in the formation of a monohydrate with a characteristic peak on lines (2) and (4). Similar conclusions were also drawn from milling in neat water by González et al. [97]. Here, additional experiments were conducted by re-milling the freshly obtained hydrate in conditions of higher content of methanol. This resulted in the disappearance of the hydrate, as evidenced by line (5) in Figure 5. This is in good accord with the quantitative characteristics of aqueous methanol binary mixtures made by Liu et al. [86]. This observation serves as a general characteristic of theophylline behaviour in saturated solutions. Hence, the conditions of the solubility measurements determine the form of theophylline precipitate since the monohydrate-anhydrate transition is energetically reversible [94].

It is worth emphasising that despite the fact that in the case of theophylline, different precipitates occur with varying binary mixture compositions, the same set of parameters were used for modelling the solubility in the whole range of solvent ratios. For example, CNIBS/R-K and the modified Jouyban–Acree models were applied for interpreting the solubility of theophylline in methanol-water and isopropanol-water conditions [86]. Moreover, rich solid-liquid phase diagrams of theophylline were computed using the same set of parameters in PC-SAFT modelling [90]. Besides, preferential solvation of theophylline in methanol-water mixtures was characterised using the same idea [85]. This can be justified by the fact that the physicochemical properties of dissolved theophylline depend only on the bulk liquid in an equilibrated state, and the form of the solid-state does not interfere with the molecular parameters of the dissolved solute. The same approach was adopted here in formulating machine learning models since all parameters used for training characterise theophylline in the dissolved state. Hence, the solubility predictions can inform the bulk concentration but not about the form of the precipitate. Although this restriction is important for any model of solubility computations, it does not prohibit its applications for back solubility computations or screening purposes.

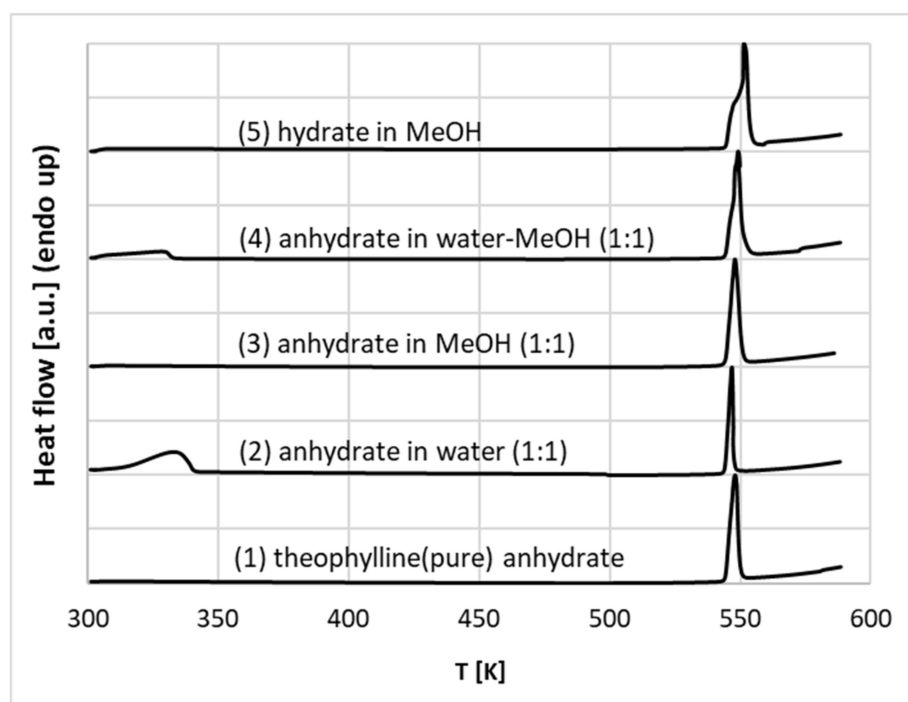


Figure 5. Thermograms of the products of milling of pure commercial anhydrated theophylline (1), anhydrate co-grinded with equimolar amounts of neat water (2), neat methanol (3), aqueous unimolar mixture with methanol (4) and monohydrate co-grinded with methanol (5).

2.3. Predictive Solubility Model

The set of experimentally derived data comprising 160 measurements was used for the theoretical dissemination. In the first stage, the first principle approach utilising COSMO-RS theory was applied for direct solubility predictions. Unfortunately, it is clear that this approach is hardly suited for solubility predictions of theophylline in the studied solvents, as documented in Figure 6. The average percentage error exceeds 103%, and such high inaccuracy prevents the utilisation of these computations as a screening tool and guidance for the rational selection of solvents or their mixtures for enhancing solubility. The error introduced by COSMO-RS is of a not-systematic type, and no linear relationship was found for the improvement of the computed solubility. That is why an alternative approach was adopted, taking advantage of the non-linear machine learning protocol described in the methodology. The predictive power of any machine learning procedure relies on an adequate selection of molecular descriptors. The first choice of such parameters, very often utilised in modelling, comes from well-established [98–100] sets of 2D or 3D descriptors. Such an approach was successfully demonstrated in the case of co-crystal screening [101] or screening for solid curcumin formulations [102]. Many alternative molecular descriptors are of potential interest, as reviewed by Bergström and Larsson [103]. However, the application of such datasets for studies of temperature-dependent properties and/or multinary systems is problematic. Here, an alternative source of molecular descriptors was proposed. According to common chemical intuition, the concentration of saturated solution depends strongly on the intermolecular interactions. Hence, it seems plausible to utilise the energetic contributions as they can vary with the composition and measurements conditions. The first principle computations offer direct insight into the amount of hydrogen bonding energetics along with quantified dispersion and electrostatics. These data were taken from routine COSMO-RS computations, as described in the methodology. Although the solubility computed using COMSOtherm software suffers from serious inaccuracies, the energetic contributions provide sufficiently detailed information about chemical systems. It happened that the set of seven descriptors characterised the whole chemical space accurately enough to allow for the precise prediction of solubility in such a diverse set of solvents,



including neat solvents, their binary mixtures and ternary NADES formulations. Since solubility values computed using COSMO-RS methodology have been found to be highly inaccurate, their values were not used as molecular descriptors, and only intermolecular interactions were included. The accuracy of the obtained model, along with its applicability domain, was presented in Figure 6. The mean RMSD value and the one characterising the best SANN were 36.6×10^{-3} and 29.3×10^{-3} , respectively. The ensemble of neural networks comprises one-layer networks with 6 to 12 neurons. The detailed characteristics of the developed ensemble are provided in Supplementary Materials, Table S7. Since one of the inclusion criteria was the number of outliers not exceeding four, all SANNs are well suited for the theophylline solubility reproduction. The most common outliers in many networks were the cases of cold water solubility ($T = 14.5 \text{ }^\circ\text{C}$) and half diluted solvents [ChCl+Glycerol], i.e., NADES mixed with unimolar proportions with water.

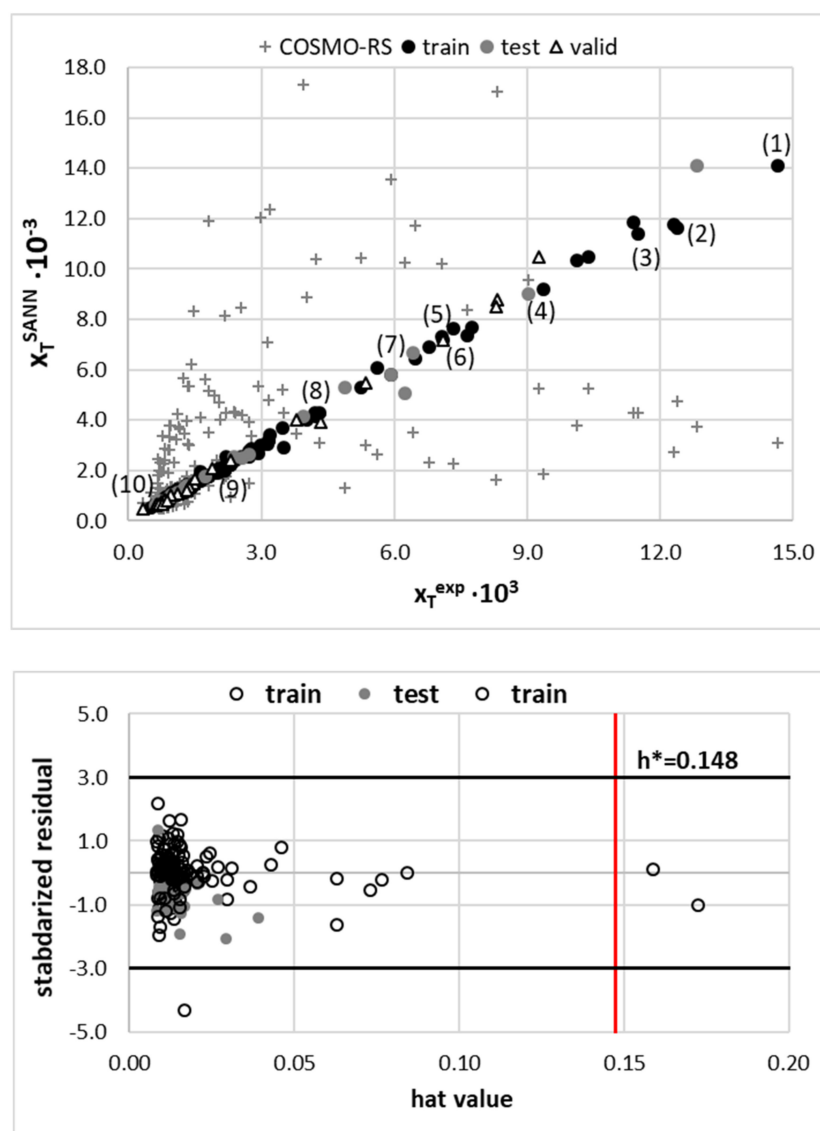


Figure 6. Accuracy of solubility prediction using COSMO-RS and an ensemble of SANNs, (top) relationship between experimental and predicted values of theophylline solubility in studied solvents and (bottom) applicability domain of a SANNs ensemble defined by relationship between standardized residuals and hat values. Some selected solvents were marked as (1) [ChCl+Glycerol+water] (0.8), (2) [ChCl+Sorbitol+water] (0.8), (3) [ChCl+Xylitol+water] (0.8), (4) [DMSO+water] (0.5), (5) [DMF+water] (0.5), (6) [DMSO], (7) [1,4-dioxane+water] 0.5, (8) 1-propanol (9) acetone and (10) water.

2.4. Theophylline Solubility Screening

The purpose of the theoretical part of this paper is not merely the formulation of a theoretical model enabling accurate back computations of theophylline solubility but also the development of a reliable and extendable approach applicable for screening of new candidates for promising solubility enhancers. Bearing in mind that any extrapolation based on a single model cannot guarantee the expected reliability, even if back-computations match the fitted data set very closely, the ensemble of SANNs was developed, and the averaged values of predicted solubility were used for final screening purposes. Hence, the sets of 150 neat solvents, 250 binary mixtures and 1500 NADES were analysed, for which theophylline solubility was computed using the developed model. It is fair to say that not all values obtained during this stage fulfil the formal restrictions, and occasionally (less than 0.4%), the computed solubility values were outside of the meaningful physical range of mole fractions, and as such, they were discarded. That is why for every solvent, theophylline solubility was averaged over all SANNs included in the ensemble. The necessity of this step is documented in Figure 7, where the values of predicted solubility were plotted as a function of the size of the ensemble formed by the systematic inclusion of an increasing number of SANNs sorted according to increasing values of their RMSD. Interestingly, the green-solid line plotted in Figure 7 representing theophylline solubility prediction in DMSO is almost constant. This proves that the solubility values back-computed using different SANNs ensembles are stable and match the experimental ones. Among considered neat solvents, it was found that theophylline has the highest affinity toward solvents of class IV, grouping aprotic highly dipolar solvents. It is worth mentioning that DMSO is also classified as this type of solvent. The highest predicted solubility in neat solvents was found for 1-methyl-2-pyrrolidone (NMP), suggesting slightly higher efficiency compared to DMSO. However, it is worth emphasising that the computed solubility value was quite sensitive to the size of the SANNs ensemble. As it was presented in Figure 7 by a green-dashed line, the predicted value of theophylline solubility in NMP exhibits quite slow convergence, and prediction stabilises after as many as 35 networks used for averaging. Besides, some representatives of class II, i.e., weak electron-pair donor bases, such as morpholine and sulfolane, reached a solubility level comparable to DMSO. Meanwhile, *m*-cresol belonging to class IX of organic acidic compounds was identified as a solvent as effective as DMSO. Furthermore, screening of binary mixtures was also quite efficient, and one binary composition was found with potentially higher solubility compared to the DMSO-water mixture. It is quite interesting to note that the NMP-water mixture was identified as a promising candidate for solubility enhancement, which is consistent with predictions for neat solvents. No other binary mixture offering higher solubility was found. The last screening was devoted to finding NADES components more effective than glycerol. As it is documented in Figure 7, this step was also successful, and several compounds were identified as potentially more efficient solubilisers. Among all considered candidates, two tetrahydrofuran analogues (3-hydroxytetrahydrofuran [CAS: 86087-23-2] and *cis*-tetrahydrofuran-3,4-Diol [CAS: 4358-64-9]) were found as the components of the most efficient NADES. Likewise, some hydroxy-analogues of cyclopropane and cyclopropene were identified as potentially interesting from the perspective of an optimisation of NADES compositions for enhancing theophylline solubility. The more detailed results of the screening are provided in Supplementary Materials, Figures S2–S4.



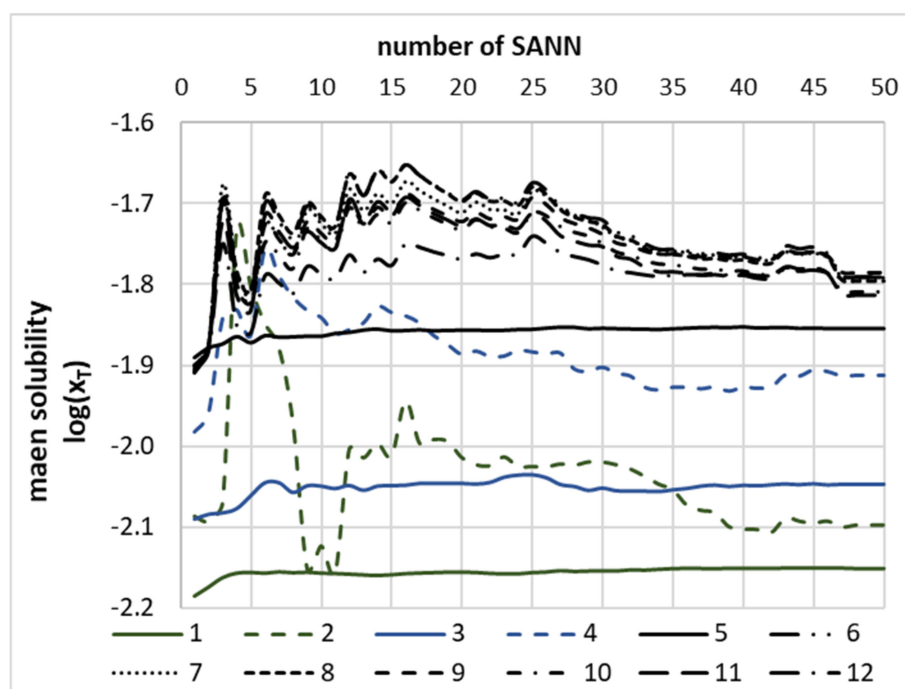


Figure 7. The distribution of mean values of theophylline solubility, expressed as the logarithm of mole fraction in the selected solvents, as a function of the number of neural networks taken for averaging. All accepted SANNs were sorted with decreasing accuracy. Legend caption: neat solvents: 1: neat DMSO, 2: neat 1-methyl-2-pyrrolidone, aqueous DMSO binary mixtures: 3: DMSO, 4: 1-methyl-2-pyrrolidone; NADES (0.8) + water (0.2) formulations with 5: glycerol, 6: (3S)-oxolan-3-ol, 7: (3R,4S)-oxolane-3,4-diol, 8: (1R,2S)-cyclobut-3-ene-1,2-diol, 9: 2,3-dihydroxycyclopropan-1-one, 10: 2,3-dihydroxycyclopentan-1-one, 11: 1-(hydroxymethyl) cycloprop-2-en-1-ol, 12: 2,2-dihydroxycyclopropan-1-one.

3. Materials and Methods

3.1. Materials

The theophylline (CAS: 58-55-9) considered in this study was obtained from Sigma-Aldrich (Poznań, Poland). The following organic solvents were used: dimethyl sulfoxide (CAS: 67-68-5), dimethylformamide (CAS: 68-12-2), 1,4-dioxane (CAS: 123-91-1), acetonitrile (CAS: 75-05-8), acetone (CAS: 67-64-1), methanol (CAS: 67-56-1), 1-propanol (CAS: 71-23-8), 1-butanol (CAS: 71-36-3), 1-pentanol (CAS: 71-41-0), 1-octanol (CAS: 111-87-5) and ethyl acetate (CAS: 141-78-6); all of which were reagent grade and purchased from Avantor Performance Materials, Gliwice, Poland. Natural deep eutectic solvents constituents included: choline chloride (CAS: 67-48-1), glucose (CAS: 50-99-7), fructose (CAS: 57-48-7), sorbitol (CAS: 50-70-4), xylitol (CAS: 87-99-0), maltose (CAS: 69-79-5), saccharose (CAS: 57-50-1) and glycerol (CAS: 56-81-5), all supplied by Sigma-Aldrich.

3.2. Preparation of Calibration Curve

A stock solution of theophylline with a concentration of 1.6 mg/mL in methanol was prepared in a 100 mL volumetric flask. From this solution, successively smaller amounts were taken and diluted with methanol in 10 mL volumetric flasks, which led to solutions with decreasing concentrations of theophylline. Solutions prepared in this manner were measured spectrophotometrically, which enabled for obtaining a relationship between the concentration of theophylline in solution and the obtained absorbance values, measured at 270 nm. Three separate curves were obtained, and their average was used as the final calibration curve for theophylline determination. The determination coefficient R^2 , as well as the limits of detection (LOD) and quantification (LOQ), were also calculated. The



detailed values of concentration and absorbance, as well as the obtained parameters of the curve, can be found in Supplementary Materials Table S1a,b, respectively.

3.3. Preparation of Samples in Organic Solvents and Their Binary Mixtures with Water

Solvent mixtures were prepared by mixing the appropriate amounts of organic solvent and water in 10 mL volumetric flasks in order to obtain different molar ratios. Further, test tubes were filled with an excess amount of theophylline for ensuring that a saturated solution is obtained. To these test tubes, the prepared solvent mixture or a pure solvent was added. Three samples were prepared for each tested solvent, both neat and binary. The samples were placed in an Orbital Shaker Incubator ES-20/60 from Biosan (Riga, Latvia) and incubated for 24 h at 25 °C. The temperature setting accuracy was 0.1 °C, and the daily deviations of temperature were at a level of ± 0.5 °C. The mixing of solutions was assured by their shaking at 60 rev/min. After incubation, the samples were filtered in a multi-step procedure with the use of a PTFE syringe filter with 0.22 μm pore size. All of the test tubes, syringes, filters, etc., were pre-heated in the same incubator in order to obtain the same temperature as the sample. From the obtained filtrate, a fixed amount of solution was transferred to other test tubes containing methanol as a diluting agent and later measured spectrophotometrically. Meanwhile, the density of the solution was measured in 10 mL volumetric flasks and used for determining the mole fractions of theophylline in the solution.

3.4. Preparation of Samples in NADES and Their Binary Mixtures with Water

Natural deep eutectic solvents (NADES) were prepared by mixing choline chloride and the second constituent in 1:1 molar ratios. Mixtures prepared in this way were placed in sealed test tubes and put in a water bath at 90 °C until a uniform solution was formed. NADES were used in their neat form and as cosolvents mixed with water in different molar ratios. An excess amount of theophylline was then added to the NADES solution, and the samples were incubated for 24 h at 25 °C, as was the case for organic solvents. Because of the increased viscosity and density, the samples were centrifuged (1000 rev/min for 5 min) using a Hettich EBA (Tuttlingen, Germany) 20 centrifuge so that the undissolved precipitate remained on the bottom of the test tube. The samples were then filtered, and their absorbances were measured after dilution, as described earlier.

3.5. Solubility Measurements

The concentration of theophylline in the samples was determined spectrophotometrically with the use of an A360 spectrophotometer from AOE Instruments (Shanghai, China). The measurements were carried out in a wavelength range from 190 nm to 700 nm with a resolution of 1 nm. Before the actual measurements, the instrument was calibrated using methanol as a reference. The whole spectra were registered for confirming that the position of the maximum was not shifted after solubility measurements. In order to ensure that the absorbance values are not outside the linearity limit, several dilutions of the samples using methanol were made, depending on the concentration of theophylline in the samples. Based on the slope coefficient and y-intercept of the calibration curve and the absorbance values measured at 270 nm, the concentration of theophylline in the samples was determined. Three samples were measured for each system, and their mean concentrations were determined and expressed as mole fractions together with the standard deviation values.

3.6. Differential Scanning Calorimetry Measurements

In order to characterise the pseudo-polymorphic behaviour of theophylline in the presence of water and organic solvents, differential scanning calorimetry (DSC) measurements were utilised for co-ground samples. The mechanochemical approach was selected due to its efficiency and previously documented application of co-grinding in successful theophylline hydrate synthesis [97].

Theophylline samples (0.4 g) were co-ground in a ball mill with appropriate amounts of pure water, water-methanol mixture and pure methanol. This step was performed using a Retsch MM 200 mill, Haan, Germany (25 Hz, 30 min.). The samples were ground in 5 mL stainless steel jars with two stainless steel balls. Then, the obtained powders were analysed using the DSC method. For this purpose, the DSC 6000 (PerkinElmer, Waltham, MA, USA) calorimeter was applied. The heating rate was set to 10 K/min. The inert atmosphere was provided by nitrogen flow (5.0 grade, Linde, 20 mL/min). The samples were measured in standard aluminium pans. Prior to the measurements, the calorimeter was calibrated using indium and zinc melting standards provided by the DSC manufacturer (PerkinElmer).

3.7. COSMO-RS Solubility Computations

The amount of solute dissolved at a given temperature in any solvent is governed by the thermodynamics of the solid-liquid phase equilibrium, in which the value of the chemical potential of the i -th solute in its saturated state is identical to the value of the chemical potential of the pure solid phase:

$$\mu_i^s = \mu_i^{liq} \quad (1)$$

$$\mu_i^{o,s} + RT \ln(a_i^s) = \mu_i^{o,liq} + RT \ln(\gamma_i^{sat} x_i^{sat}) \quad (2)$$

where γ_i^{sat} is the activity coefficient of the solute in the saturated solution of solute concentration equal to x_i^{sat} ; T and R represent the temperature and the gas constant, respectively. For practical applications, the above equation requires the definition of the thermodynamic reference states, $\mu_i^{o,s}$ and $\mu_i^{o,liq}$. Here a commonly accepted convention is assumed [104] in which the reference state for the solid phase is the solid itself, and consequently, the activity of the solid equals the activity of the solute, a_i^s , in the saturated solution. The thermodynamic reference state for the solute in the solution is defined as a pure compound supercooled melt at a given temperature corresponding to the conditions of the solubility measurements, $a_i^{sat} = \gamma_i^{sat} x_i^{sat}$. The activity of the solid phase depends on the fusion properties:

$$\ln(a_i^s) = \ln(\gamma_i^{sat} x_i^{sat}) = -\frac{\Delta_{fus} \bar{G}_i^m}{RT} \quad (3)$$

where $\Delta_{fus} \bar{G}_i^m$ is the partial molar Gibbs energy of melting at the solubility measurements conditions, which by definition is equal to zero for the pure solute at its melting point. The COSMO-RS is a theory [64,105,106] of the liquid state and, as such, offers a direct calculation of chemical potentials of components in the bulk phase by calculating molecular interactions as local pair interactions of segments of molecular COSMO-surfaces. Practical solubility computations within the COSMO-RS approach are done by iteratively solving Equation (3), assuming that the values of Gibbs free energy of fusion are provided from external sources. Hence, the working equation has the following form:

$$\ln(\gamma_i^{sat,i+1} x_i^{sat,i+1}) = \frac{1}{RT} (\mu_i^{o,liq} - \mu_i^{(i)} (\gamma_i^{sat,i} x_i^{sat,i})) + \max(0, \Delta_{fus} \bar{G}_i^m) \quad (4)$$

In the above equation, superscripts i and $i + 1$ denote the values obtained in two subsequent iterations. The iterative cycle is repeated until convergence is achieved, which means that the computation is supposed to be completed if the difference in the computed solubility drops below a defined threshold value.

The quantum chemical part of the COSMO calculations provides information regarding the interactions of discrete polarisation charge densities of the contacting surface segments representing a molecule embedded in a virtual conductor. Such microscopic state properties are related to macroscopic thermodynamic properties of a liquid by statistical thermodynamics through the analysis of density probability distributions termed σ -profiles and σ -potentials. Integration of the latter over the surface leads to the residual contribution

to the chemical potential allowing for predictions of almost all thermodynamic properties, including activity coefficients, excess properties or partition coefficients and solubility:

$$\mu_s(\sigma) = -\frac{RT}{a_{aff}} \ln \left[\int P_s(\sigma') \exp \left\{ \frac{a_{aff}}{RT} \left[\mu_s(\sigma') - E_{misfit}(\sigma, \sigma') - E_{HB}(\sigma, \sigma') \right] \right\} d\sigma' \right] \quad (5)$$

The actual calculations of these properties require a proper representation of the molecular structure. This is typically done by an adequate exploration of the conformational space using COMSOconf for generating the most energetically favourable structures. All computations were performed with the aid of TURBOMOLE rev. V7.5.1, Karlsruhe, Germany interfaced with TmoleX 2021 (version 21.0.1, BIOVIA, San Diego, CA, USA) using RI-DFT BP86 (B88-VWN-P86) functional and def-TZVP basis set for geometry optimisation and def2-TZVPD basis set for single-point calculations with the inclusion of the fine grid tetrahedron cavity and the inclusion of parameter sets with hydrogen bond interaction and van der Waals dispersion term based on the "D3" method of Grimme et al. [107]. All of the solubility calculations were performed using COSMOtherm (version 20.0.0, revision 5273M, BIOVIA, San Diego, CA, USA) [108] with BP_TZVPD_FINE_20.ctd parametrization.

3.8. Molecular Descriptors

According to COSMO-RS theory, any bulk system is modelled as an ensemble of closely packed ideally screened molecules enclosed by a virtual conductor. The origin of Coulomb interactions comes from the screening of electrostatics of two different contacting segments by their surroundings and the back-polarisation of the solute molecule. The specific interaction energy per unit area resulting from this "misfit" stands for the electrostatic contribution to the total energy:

$$E_{MF}(\sigma, \sigma') = a_{eff} \frac{\alpha'}{2} (\sigma + \sigma')^2 \quad (6)$$

where a_{eff} is the effective contact area between two surface segments and α' is an adjustable parameter. Furthermore, the hydrogen bonding (H-Bond) is similarly described by screening the charge density of the two adjacent centres of strongly negative and positive densities of donors and acceptors, respectively. It is assumed that such interactions take place if two sufficiently polar pieces of the surface of opposite polarity are in contact and defined by a mathematical formula in the following form:

$$E_{HB}(\sigma, \sigma') = a_{eff} c_{HB} \min(0; \min(0; \sigma_{don} + \sigma_{HB}); \min(0; \sigma_{acc} - \sigma_{HB})) \quad (7)$$

where σ_{HB} and σ'_{HB} are adjustable parameters. Additionally, the formalism of COSMO-RS also takes into account Van der Waals (VdW) interactions between surface segments. Such energetic contributions are defined via:

$$E_{vdW}(\sigma, \sigma') = a_{eff} c_{vdW}(\sigma, \sigma') = a_{eff} (\tau_{vdW} + \tau'_{vdW}) \quad (8)$$

where τ_{vdW} , τ'_{vdW} , and c_{vdW} are element specific adjustable parameters. The vdW energy is dependent only on the element type of the atoms that are involved in surface contact. Since any prediction applied within the COSMO-RS framework generate these energetic contributions, they are directly available from outputs of different properties, including solubility. Hence, the energetic information is taken as a set of universal descriptors characterising the properties of any system. For practical purposes, the values of misfit (E_i^{misfit}), H-Bond (E_i^{HB}) and van der Waals (E_i^{VdW}) mean interaction energies computed for the pure i -th components were augmented with a set of relative values defined as follows:

$$e_i^j = \frac{E_i^j}{E_i^{misfit} + E_i^{HB} + E_i^{VdW}} \quad (9)$$

where j stands for misfit, HB or VdW contribution. In the case of multicomponent systems, the descriptor values were computed as a weighted sum with molar fraction.

3.9. Machine Learning Protocol

The non-linear modelling was used for the development of an ensemble of artificial neural networks. The flexibility of computations offered by STATISTICA software (version 13, TIBCO Software Inc., Palo Alto, CA, USA) was used for Statistica Automated Neural Networks (SANNs) growth. Simple one-layer infrastructure was assumed. The input layer comprised the sets of seven molecular descriptors depicting intermolecular interactions, and the only expected value in the output layer was the estimated solubility. The default settings of SANN generation were assumed, allowing for the automatic selection of initial and output neurons to be in the functional form of linear, logistic, Tanh, exponential or sin functions. Only algorithms utilising multilayer perceptrons (MLP) were allowed during the machine learning phase. The model was constructed using 160 points, and the dataset was automatically split into training (70%), test (15%) and validation (15%) subsets. The sum-squares error function (SOS) was used as an error function for the training of the neural networks. The ensemble of networks was constructed by a successful accumulation of networks fulfilling two conditions, namely $RMDS < 0.04$ and the largest number of outliers not exceeding four cases. The applicability domain of each model was computed using the hat value according to standard methodology [109–111].

3.10. Solvents Selection for Solubility Screening

The list of neat solvents was taken from the literature [112,113], where classification into ten chemically intuitive classes was suggested. Hence, the screening of this set of neat solvents might provide clues about solvent types offering higher solubility of theophylline. The second group of solvents was taken from an in house database of binary solvents effectively used for solubility measurements. In principle, binary mixtures might be constructed by all possible combinations of neat solvents, but this is not recommended since it is not obvious which pairs of solvents are miscible. Hence, the list was restricted only to such mixtures, which were already utilised as solvents for solubility measurements. The last set was formulated using replacers for glycerol in NADES formulations with choline chloride. For this purpose, the PubChem database was searched for analogues of glycerol and other sugars used in this study. After removing duplicates, the initial list was inspected for systems containing atoms other than C, H and O, since the search was narrowed only to polyol analogues. These were removed from the list along with radicals, anions and isotopomers. Finally, 1050 compounds were used for the NADES screening in compositions mimicking prescriptions of real NADES systems. For every solvent, the values of molecular descriptors were computed and used for theophylline solubility computations with an aid to the developed SANNs ensemble and averaged.

4. Conclusions

Solubility of theophylline is quite poor in water and the majority of neat organic solvents. The rare exception is DMSO, which as an aprotic highly dipolar solvent, is commonly used for the solubility enhancement of many APIs. Likewise, a serious solubility gain is often observed for binary mixtures of some organic solvents with water. These observations are intuitive and agree with expectations arising from chemical experiences. However, the performed experiments provided additional less obvious characteristics. Namely, a non-trivial synergistic effect was measured for aqueous solutions with five different organic solvents. In each case, the highest solubility was found to be corresponding to a unimolar proportion of organic solvent and water. The 1:1 mixture of DMSO with water is characterised by 1.27 times higher solubility compared to pure DMSO. A similar effect was observed for mixtures of such organic solvents like DMF, 1,4-dioxane and some aliphatic alcohols. For further improvements of theophylline dissolution, several NADES compositions were tested, and many of them not only offer better solubility at room tem-

perature but also exhibit an even stronger synergistic effect. The unimolar mixtures of choline chloride with glycerol and other polyols have been found as efficient solubilisers preserving solubility enhancement after dilution with water. The solubility measurements revealed that the most effective designed solvent comprises a unimolar composition of choline chloride with glycerol in an aqueous formulation comprising 0.2 molar fractions of water. It is quite interesting to see that the water dilution of NADES promotes the solubility of theophylline since typical NADES possess very strong hygroscopic properties and their physicochemical properties are often very sensitive to the amount of contained water. Here it was proved that for solubility enhancement of theophylline, the hygroscopicity of NADES is not a constrain but rather promotes hydrotropic features.

The experimental screening of theophylline solubility was extended with theoretical analysis using non-linear modelling, which resulted in an ensemble of SANNs enabling for very accurate mole fraction back-computations and also for extensive screening of neat solvents, their binary mixtures and a variety of polyols in NADES formulations containing choline chloride. Among 153 neat solvents grouped into ten classes, the solvents of the first choice are aprotic highly dipolar solvents and such ones as NMP, morpholine and sulfolane are alternatives to DMSO. This conclusion also holds for aqueous binary formulations as inferred from the screening of theophylline solubility in 252 binary solvent mixtures. The model formulated based on the neural networks ensemble was also effective in the optimisation of NADES constituents. Several replacers of glycerol were found offering significant solubility enhancement, among which hydroxy-analogues of tetrahydrofuran, cyclopropane and cyclopropene were potentially more efficient NADES than the one with glycerol. Future experiments for theophylline solubility enhancement might be directed by these observations.

Additionally, the performed non-linear analysis also provides methodological clues for the further development of accurate models for solubility predictions. It is not advised to use a single neural network even if it offers high accuracy in back computed values and matches to the experimental ones. Occasionally, even for systems within the applicability domain of the model, the predicted values do not guarantee mole fractions within the formal requirement belonging to the range between zero and unity. This problem is not very serious since it has been encountered in less than 0.4% of cases. A more severe problem is posed by the occasional inconsistencies between predictions of different SANNs, even with comparable RMSDs. Fortunately, taking into account a sufficiently extended ensemble of neural networks guarantees the convergence of predicted solubility values provided that each individual network is satisfactorily accurate. Although depending on the considered solvent, the size of ensemble necessary for convergence varies, using at least 30 networks was found to be generally sufficient. From the perspective of practical implementations, this is not a serious restriction since each step can be easily automated.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/ijms22147347/s1>.

Author Contributions: P.C. (Piotr Cysewski), conceptualization, methodology, writing—original draft preparation, writing—review and editing, formal analysis, supervision, project administration, resources and investigation; P.C. (Patryk Cymerman), investigation; T.J., investigation, writing—original draft preparation, writing—review and editing, project administration and resources. M.P., investigation, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data supporting reported results are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Andreeva, E.Y.; Dmitrienko, S.G.; Zolotov, Y.A. Methylxanthines: Properties and determination in various objects. *Russ. Chem. Rev.* **2012**, *81*, 397–414. [[CrossRef](#)]
2. Spiller, G.A. *Caffeine*, 1st ed.; CRC Press: Boca Raton, FL, USA, 1998.
3. Craig, C.R.; Stitzel, R.E. *Modern Pharmacology with Clinical Applications*, 6th ed.; Lippincott Williams and Wilkins: Philadelphia, PA, USA, 2003.
4. Satoskar, R.S.; Rege, N.; Bhandarkar, S.D. *Pharmacology and Pharmacotherapeutics*, 24th ed.; Elsevier India: New Delhi, India, 2015.
5. Essayan, D.M. Cyclic nucleotide phosphodiesterases. *J. Allergy Clin. Immunol.* **2001**, *108*, 671–680. [[CrossRef](#)] [[PubMed](#)]
6. Müller, C.E.; Hide, I.; Daly, J.W.; Rothenhausler, K.; Eger, K. 7-Deaza-2-phenyladenines: Structure-Activity Relationships of Potent A1 Selective Adenosine Receptor Antagonists. *J. Med. Chem.* **1990**, *33*, 2822–2828. [[CrossRef](#)]
7. Roche, N. Systemic Medications in Chronic Obstructive Pulmonary Disease: Use and Outcomes. *Clin. Chest Med.* **2020**, *41*, 485–494. [[CrossRef](#)]
8. Barnes, P.J. Therapy of Airway Disease: Epigenetic Potential. In *Epigenetics in Human Disease*, 1st ed.; Elsevier Inc.: London, UK, 2012.
9. Rabe, K.F.; Hurd, S.; Anzueto, A.; Barnes, P.J.; Buist, S.A.; Calverley, P.; Fukuchi, Y.; Jenkins, C.; Rodriguez-Roisin, R.; Van Weel, C.; et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am. J. Respir. Crit. Care Med.* **2007**, *176*, 532–555. [[CrossRef](#)]
10. Rowe, D.J.F.; Watson, I.D.; Williams, J.; Berry, D.J. The Clinical Use and Measurement of Theophylline. *Ann. Clin. Biochem. Int. J. Lab. Med.* **1988**, *25*, 4–26. [[CrossRef](#)]
11. Devereux, G.; Cotton, S.; Barnes, P.; Briggs, A.; Burns, G.; Chaudhuri, R.; Chrystyn, H.; Davies, L.; De Soyza, A.; Fielding, S.; et al. Use of low-dose oral theophylline as an adjunct to inhaled corticosteroids in preventing exacerbations of chronic obstructive pulmonary disease: Study protocol for a randomised controlled trial. *Trials* **2015**, *16*, 267. [[CrossRef](#)] [[PubMed](#)]
12. McDonald, C.J.; Holgate, S.T. The role of theophylline in the management of chronic asthma in adults. *Clin. Exp. Allergy* **1996**, *26*, 42–46. [[CrossRef](#)]
13. Crapnell, R.D.; Banks, C.E. Electroanalytical overview: The electroanalytical detection of theophylline. *Talanta Open* **2021**, *3*, 100037. [[CrossRef](#)]
14. Ku, M.S.; Dulin, W. A biopharmaceutical classification-based Right-First-Time formulation approach to reduce human pharmacokinetic variability and project cycle time from First-in-Human to clinical Proof-of-Concept. *Pharm. Dev. Technol.* **2012**, *17*, 285–302. [[CrossRef](#)] [[PubMed](#)]
15. Takagi, T.; Ramachandran, C.; Bermejo, M.; Yamashita, S.; Yu, L.X.; Amidon, G.L. A Provisional Biopharmaceutical Classification of the Top 200 Oral Drug Products in the United States, Great Britain, Spain, and Japan. *Mol. Pharm.* **2006**, *3*, 631–643. [[CrossRef](#)]
16. Williams, H.D.; Trevaskis, N.L.; Charman, S.A.; Shanker, R.M.; Charman, W.N.; Pouton, C.W.; Porter, C.J.H. Strategies to address low drug solubility in discovery and development. *Pharmacol. Rev.* **2013**, *65*, 315–499. [[CrossRef](#)] [[PubMed](#)]
17. Scholz, A.; Abrahamsson, B.; Diebold, S.M.; Kostewicz, E.; Polentarutti, B.I.; Ungell, A.-L.; Dressman, J.B. Influence of hydrodynamics and particle size on the absorption of felodipine in labradors. *Pharm. Res.* **2002**, *19*, 42–46. [[CrossRef](#)]
18. Merisko-Liversidge, E.; Liversidge, G.G. Nanosizing for oral and parenteral drug delivery: A perspective on formulating poorly-water soluble compounds using wet media milling technology. *Adv. Drug Deliv. Rev.* **2011**, *63*, 427–440. [[CrossRef](#)] [[PubMed](#)]
19. Van Eerdenbrugh, B.; Van den Mooter, G.; Augustijns, P. Top-down production of drug nanocrystals: Nanosuspension stabilization, miniaturization and transformation into solid products. *Int. J. Pharm.* **2008**, *364*, 64–75. [[CrossRef](#)]
20. Hancock, B.C.; Parks, M. What is the true solubility advantage for amorphous pharmaceuticals? *Pharm. Res.* **2000**, *17*, 397–404. [[CrossRef](#)]
21. Huang, L.; Tong, W.-Q. Impact of solid state properties on developability assessment of drug candidates. *Adv. Drug Deliv. Rev.* **2004**, *56*, 321–334. [[CrossRef](#)] [[PubMed](#)]
22. Shegokar, R.; Müller, R.H. Nanocrystals: Industrially feasible multifunctional formulation technology for poorly soluble actives. *Int. J. Pharm.* **2010**, *399*, 129–139. [[CrossRef](#)]
23. Brewster, M.E.; Loftsson, T. Cyclodextrins as pharmaceutical solubilizers. *Adv. Drug Deliv. Rev.* **2007**, *59*, 645–666. [[CrossRef](#)] [[PubMed](#)]
24. Rajewski, R.A.; Stella, V.J. Pharmaceutical Applications of Cyclodextrins. 2. In Vivo Drug Delivery. *J. Pharm. Sci.* **1996**, *85*, 1142–1169. [[CrossRef](#)] [[PubMed](#)]
25. Janssens, S.; Van den Mooter, G. Review: Physical chemistry of solid dispersions. *J. Pharm. Pharmacol.* **2009**, *61*, 1571–1586. [[CrossRef](#)]
26. Kranz, H.; Guthmann, C.; Wagner, T.; Lipp, R.; Reinhard, J. Development of a single unit extended release formulation for ZK 811 752, a weakly basic drug. *Eur. J. Pharm. Sci.* **2005**, *26*, 47–53. [[CrossRef](#)] [[PubMed](#)]
27. Tatavarti, A.S.; Hoag, S.W. Microenvironmental pH Modulation Based Release Enhancement of a Weakly Basic Drug from Hydrophilic Matrices (This work was presented in-part at the 31st Annual Controlled Release Society meeting in Honolulu, Hawaii (June 2004)). *J. Pharm. Sci.* **2006**, *95*, 1459–1468. [[CrossRef](#)] [[PubMed](#)]
28. Korn, C.; Balbach, S. Compound selection for development—Is salt formation the ultimate answer? Experiences with an extended concept of the “100 mg approach”. *Eur. J. Pharm. Sci.* **2014**, *57*, 257–263.

29. Serajuddin, A.T.M. Salt formation to improve drug solubility. *Adv. Drug Deliv. Rev.* **2007**, *59*, 603–616. [[CrossRef](#)]
30. Chadha, R.; Bhalla, Y.; Vashisht, M.K.; Chadha, K. Cococrystallization in Nutraceuticals. In *Recrystallization in Materials Processing*; Glebovsky, V., Ed.; InTech: London, UK, 2015.
31. Vishweshwar, P.; McMahon, J.A.; Bis, J.A.; Zaworotko, M.J. Pharmaceutical Co-Crystals. *J. Pharm. Sci.* **2006**, *95*, 499–516. [[CrossRef](#)]
32. Blanco-Márquez, J.H.; Ortiz, C.P.; Cerquera, N.E.; Martínez, F.; Jouyban, A.; Delgado, D.R. Thermodynamic analysis of the solubility and preferential solvation of sulfamerazine in (acetonitrile + water) cosolvent mixtures at different temperatures. *J. Mol. Liq.* **2019**, *293*, 111507. [[CrossRef](#)]
33. Blanco-Márquez, J.H.; Caviedes Rubio, D.I.; Ortiz, C.P.; Cerquera, N.E.; Martínez, F.; Delgado, D.R. Thermodynamic analysis and preferential solvation of sulfamethazine in acetonitrile + water cosolvent mixtures. *Fluid Phase Equilib.* **2020**, *505*, 112361. [[CrossRef](#)]
34. Smith, P.E.; Mazo, R.M. On the theory of solute solubility in mixed solvents. *J. Phys. Chem. B* **2008**, *112*, 7875–7884. [[CrossRef](#)] [[PubMed](#)]
35. Jouyban, A. Review of the cosolvency models for predicting solubility of drugs in water-cosolvent mixtures. *J. Pharm. Pharm. Sci.* **2008**, *11*, 32–58. [[CrossRef](#)] [[PubMed](#)]
36. Jouyban, A. *Handbook of Solubility Data for Pharmaceuticals*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2010.
37. Smith, E.L.; Abbott, A.P.; Ryder, K.S. Deep Eutectic Solvents (DESs) and Their Applications. *Chem. Rev.* **2014**, *114*, 11060–11082. [[CrossRef](#)]
38. Choi, Y.H.; van Spronsen, J.; Dai, Y.; Verberne, M.; Hollmann, F.; Arends, I.W.C.E.; Witkamp, G.-J.; Verpoorte, R. Are natural deep eutectic solvents the missing link in understanding cellular metabolism and physiology? *Plant Physiol.* **2011**, *156*, 1701–1705. [[CrossRef](#)]
39. Dai, Y.; van Spronsen, J.; Witkamp, G.-J.; Verpoorte, R.; Choi, Y.H. Natural deep eutectic solvents as new potential media for green technology. *Anal. Chim. Acta* **2013**, *766*, 61–68. [[CrossRef](#)] [[PubMed](#)]
40. Dai, Y.; Verpoorte, R.; Choi, Y.H. Natural deep eutectic solvents providing enhanced stability of natural colorants from safflower (*Carthamus tinctorius*). *Food Chem.* **2014**, *159*, 116–121. [[CrossRef](#)]
41. Espino, M.; de los Ángeles Fernández, M.; Gomez, F.J.V.; Silva, M.F. Natural designer solvents for greening analytical chemistry. *TrAC Trends Anal. Chem.* **2016**, *76*, 126–136. [[CrossRef](#)]
42. Mamajanov, I.; Engelhart, A.E.; Bean, H.D.; Hud, N.V. DNA and RNA in Anhydrous Media: Duplex, Triplex, and G-Quadruplex Secondary Structures in a Deep Eutectic Solvent. *Angew. Chem. Int. Ed.* **2010**, *122*, 6454–6458. [[CrossRef](#)]
43. Paiva, A.; Craveiro, R.; Aroso, I.; Martins, M.; Reis, R.L.; Duarte, A.R.C. Natural Deep Eutectic Solvents—Solvents for the 21st Century. *ACS Sustain. Chem. Eng.* **2014**, *2*, 1063–1071. [[CrossRef](#)]
44. Gutiérrez, M.C.; Ferrer, M.L.; Yuste, L.; Rojo, F.; del Monte, F. Bacteria Incorporation in Deep-eutectic Solvents through Freeze-Drying. *Angew. Chem. Int. Ed.* **2010**, *49*, 2158–2162. [[CrossRef](#)] [[PubMed](#)]
45. Francisco, M.; van den Bruinhorst, A.; Kroon, M.C.; Song, J.; Fan, H.; Han, B.; Itoh, T.; Verpoorte, R. New natural and renewable low transition temperature mixtures (LTTMs): Screening as solvents for lignocellulosic biomass processing. *Green Chem.* **2012**, *14*, 2153. [[CrossRef](#)]
46. Durand, E.; Lecomte, J.; Baréa, B.; Dubreucq, E.; Lortie, R.; Villeneuve, P.; Saheed, O.K. Evaluation of deep eutectic solvent–water binary mixtures for lipase-catalyzed lipophilization of phenolic acids. *Green Chem.* **2013**, *15*, 2275–2282. [[CrossRef](#)]
47. Bakirtzi, C.; Triantafyllidou, K.; Makris, D.P. Novel lactic acid-based natural deep eutectic solvents: Efficiency in the ultrasound-assisted extraction of antioxidant polyphenols from common native Greek medicinal plants. *J. Appl. Res. Med. Aromat. Plants* **2016**, *3*, 120–127. [[CrossRef](#)]
48. Erlund, I.; Kosonen, T.; Alfthan, G.; Mäenpää, J.; Perttunen, K.; Kenraali, J.; Parantainen, J.; Aro, A. Pharmacokinetics of quercetin from quercetin aglycone and rutin in healthy volunteers. *Eur. J. Clin. Pharmacol.* **2000**, *56*, 545–553. [[CrossRef](#)] [[PubMed](#)]
49. Radošević, K.; Curko, N.; Gaurina Srček, V.; Cvjetko Bubalo, M.; Tomašević, M.; Kovačević Ganić, K.; Radojčić Redovniković, I. Natural deep eutectic solvents as beneficial extractants for enhancement of plant extracts bioactivity. *LWT* **2016**, *73*, 45–51. [[CrossRef](#)]
50. Faller, B.; Ertl, P. Computational approaches to determine drug solubility. *Adv. Drug Deliv. Rev.* **2007**, *59*, 533–545. [[CrossRef](#)] [[PubMed](#)]
51. Skyner, R.E.; McDonagh, J.L.; Groom, C.R.; Van Mourik, T.; Mitchell, J.B.O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, *17*, 6174–6191. [[CrossRef](#)] [[PubMed](#)]
52. Bergström, C.A.S.; Norinder, U.; Luthman, K.; Artursson, P. Experimental and computational screening models for prediction of aqueous drug solubility. *Pharm. Res.* **2002**, *19*, 182–188. [[CrossRef](#)]
53. Wilson, G.M. Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *J. Am. Chem. Soc.* **1964**, *86*, 127–130. [[CrossRef](#)]
54. Renon, H.; Prausnitz, J.M. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE J.* **1968**, *14*, 135–144. [[CrossRef](#)]
55. Apelblat, A.; Manzurolo, E. Solubilities of o-acetylsalicylic, 4-aminosalicylic, 3,5-dinitrosalicylic, and p-toluic acid, and magnesium-DL-aspartate in water from T = (278 to 348) K. *J. Chem. Thermodyn.* **1999**, *31*, 85–91. [[CrossRef](#)]
56. Svärd, M.; Rasmuson, Å.C. (Solid + liquid) solubility of organic compounds in organic solvents—Correlation and extrapolation. *J. Chem. Thermodyn.* **2014**, *76*, 124–133. [[CrossRef](#)]

57. Buchowski, H.; Ksiazczak, A.; Pietrzyk, S. Solvent activity along a saturation line and solubility of hydrogen-bonding solids. *J. Phys. Chem.* **1980**, *84*, 975–979. [\[CrossRef\]](#)
58. Ran, Y.; Yalkowsky, S.H. Prediction of drug solubility by the general solubility equation (GSE). *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 354–357. [\[CrossRef\]](#) [\[PubMed\]](#)
59. Bradley, J.-C.; Abraham, M.H.; Acree, W.E.; Lang, A.S. Predicting Abraham model solvent coefficients. *Chem. Cent. J.* **2015**, *9*, 12. [\[CrossRef\]](#) [\[PubMed\]](#)
60. Barzegar-Jalali, M.; Jouyban, A.; Mohammadian, E.; Martinez, F.; Rahimpour, E. Budesonide solubility in some non-aqueous mono-solvents at different temperatures: Measurements and mathematical correlation with Abraham model. *J. Mol. Liq.* **2018**, *269*, 461–466. [\[CrossRef\]](#)
61. Rahimpour, E.; Jouyban, A. Utilizing Abraham and Hansen solvation parameters for solubility prediction of meloxicam in cosolvency systems. *J. Mol. Liq.* **2021**, *328*, 115400. [\[CrossRef\]](#)
62. Mutelet, F.; Baker, G.A.; Zhao, H.; Churchill, B.; Acree, W.E. Development of Abraham model correlations for short-chain glycol-grafted imidazolium and pyridinium ionic liquids from inverse gas-chromatographic measurements. *J. Mol. Liq.* **2020**, *317*, 113983. [\[CrossRef\]](#)
63. Klamt, A.; Eckert, F. COSMO-RS: A novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilibria* **2000**, *172*, 43–72. [\[CrossRef\]](#)
64. Eckert, F.; Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **2002**, *48*, 369–385. [\[CrossRef\]](#)
65. Oliveira, G.; Wojcickowski, J.P.; Farias, F.O.; Igarashi-Mafra, L.; de Pelegrini Soares, R.; Mafra, M.R. Enhancement of biomolecules solubility in aqueous media using designer solvents as additives: An experimental and COSMO-based models' approach. *J. Mol. Liq.* **2020**, *318*, 114266. [\[CrossRef\]](#)
66. Zhu, R.; Lei, Z. COSMO-based models for predicting the gas solubility in polymers. *Green Energy Environ.* **2021**, *6*, 311–313. [\[CrossRef\]](#)
67. Jeliński, T.; Cysewski, P. Screening of ionic liquids for efficient extraction of methylxanthines using COSMO-RS methodology. *Chem. Eng. Res. Des.* **2017**, *122*, 176–183. [\[CrossRef\]](#)
68. Khan, A.S.; Ibrahim, T.H.; Rashid, Z.; Khamis, M.I.; Nancarrow, P.; Jabbar, N.A. COSMO-RS based screening of ionic liquids for extraction of phenolic compounds from aqueous media. *J. Mol. Liq.* **2021**, *328*, 115387. [\[CrossRef\]](#)
69. Padaszyński, K.; Domańska, U. COSMO-RS screening for ionic liquid to be applied in extraction of 2-phenylethanol from aqueous solutions. *J. Mol. Liq.* **2018**, *271*, 305–312. [\[CrossRef\]](#)
70. Song, Z.; Wang, J.; Sundmacher, K. Evaluation of COSMO-RS for solid–liquid equilibria prediction of binary eutectic solvent systems. *Green Energy Environ.* **2021**, *6*, 371–379. [\[CrossRef\]](#)
71. Dezhang, S.; Huiheng, F.; Feng, X.; Wenxiu, L.; Zhigang, Z. Evaluation of COSMO-RS model for the LLE prediction of benzene plus cyclohexane plus ionic liquid system. *J. Chem. Thermodyn.* **2020**, *145*, 106032. [\[CrossRef\]](#)
72. Torcal, M.; Langa, E.; Pardo, J.I.; Mainar, A.M.; Urieta, J.S. Isobaric VLE of the mixture {1,8-cineole + ethanol}. EOS analysis and COSMO-RS modeling. *J. Chem. Thermodyn.* **2016**, *97*, 88–92. [\[CrossRef\]](#)
73. Eckert, F.; Diedenhofen, M.; Klamt, A. Towards a first principles prediction of pKa: COSMO-RS and the cluster-continuum approach. *Mol. Phys.* **2010**, *108*, 229–241. [\[CrossRef\]](#)
74. Jeliński, T.; Przybyłek, M.; Cysewski, P. Solubility advantage of sulfanilamide and sulfacetamide in natural deep eutectic systems: Experimental and theoretical investigations. *Drug Dev. Ind. Pharm.* **2019**, *45*, 1120–1129. [\[CrossRef\]](#)
75. Cysewski, P.; Jeliński, T. Optimization, thermodynamic characteristics and solubility predictions of natural deep eutectic solvents used for sulfonamide dissolution. *Int. J. Pharm.* **2019**, *570*, 118682. [\[CrossRef\]](#)
76. Xu, Y.; Zhou, Y.; Sekula, P.; Ding, L. Machine learning in construction: From shallow to deep learning. *Dev. Built Environ.* **2021**, *6*, 100045. [\[CrossRef\]](#)
77. Tang, S.; Yang, Y. Why neural networks apply to scientific computing? *Theor. Appl. Mech. Lett.* **2021**, 100242, in press. [\[CrossRef\]](#)
78. Can, E.; Jalal, A.; Zirhlioglu, I.G.; Uzun, A.; Yildirim, R. Predicting water solubility in ionic liquids using machine learning towards design of hydro-philic/phobic ionic liquids. *J. Mol. Liq.* **2021**, *332*, 115848. [\[CrossRef\]](#)
79. Huang, X.; Wang, H.; Luo, W.; Xue, S.; Hayat, F.; Gao, Z. Prediction of loquat soluble solids and titratable acid content using fruit mineral elements by artificial neural network and multiple linear regression. *Sci. Hortic.* **2021**, *278*, 109873. [\[CrossRef\]](#)
80. Fu, L.; Hu, J.; Zhang, Y.; Li, Q. Investigation on sulfur solubility in sour gas at elevated temperatures and pressures with an artificial neural network algorithm. *Fuel* **2020**, *262*, 116541. [\[CrossRef\]](#)
81. Lv, H.; Tian, D. Designing and optimizing a parallel neural network model for predicting the solubility of diosgenin in n-alkanols. *Chin. J. Chem. Eng.* **2021**, *29*, 288–294. [\[CrossRef\]](#)
82. Rostamian, H.; Lotfollahi, M.N. Statistical modeling of aspirin solubility in organic solvents by Response Surface Methodology and Artificial Neural Networks. *Phys. A Stat. Mech. Its Appl.* **2020**, *540*, 123253. [\[CrossRef\]](#)
83. Cysewski, P.; Jeliński, T.; Procek, D.; Dratwa, A. Solubility of Sulfanilamide and Sulfacetamide in neat solvents: Measurements and interpretation using theoretical predictive models, first principle approach and artificial neural networks. *Fluid Phase Equilibria* **2021**, *529*, 112883. [\[CrossRef\]](#)
84. Zhong, J.; Tang, N.; Asadzadeh, B.; Yan, W. Measurement and Correlation of Solubility of Theobromine, Theophylline, and Caffeine in Water and Organic Solvents at Various Temperatures. *J. Chem. Eng. Data* **2017**, *62*, 2570–2577. [\[CrossRef\]](#)

85. Cárdenas, Z.J.; Jiménez, D.M.; Almanza, O.A.; Jouyban, A.; Martínez, F.; Acree, W.E. Solubility and Preferential Solvation of Caffeine and Theophylline in {Methanol + Water} Mixtures at 298.15 K. *J. Solut. Chem.* **2017**, *46*, 1605–1624. [[CrossRef](#)]
86. Liu, C.; Dang, L.; Bai, W.; Wang, R.; Wei, H. Solid-liquid equilibrium of theophylline in solvent mixtures. *J. Chem. Eng. Data* **2014**, *59*, 263–268. [[CrossRef](#)]
87. Suzuki, E.; Shimomura, K.; Sekiguchi, K. Thermochemical Study of Theophylline and Its Hydrate. *Chem. Pharm. Bull.* **1989**, *37*, 493–497. [[CrossRef](#)]
88. Seton, L.; Khamar, D.; Bradshaw, I.J.; Hutcheon, G.A. Solid state forms of theophylline: Presenting a new anhydrous polymorph. *Cryst. Growth Des.* **2010**, *10*, 3879–3886. [[CrossRef](#)]
89. Seton, L.; Khamar, D.; Bradshaw, I.; Hutcheon, G. Processing induced transformations: Phase impurities introduced during hydration/dehydration. *Chem. Eng. Sci.* **2012**, *77*, 57–64. [[CrossRef](#)]
90. Lange, L.; Sadowski, G. Polymorphs, Hydrates, Cocrystals, and Cocrystal Hydrates: Thermodynamic Modeling of Theophylline Systems. *Cryst. Growth Des.* **2016**, *16*, 4439–4449. [[CrossRef](#)]
91. Szterner, P.; Legendre, B.; Sghaier, M. Thermodynamic properties of polymorphic forms of theophylline. Part I: DSC, TG, X-ray study. *J. Therm. Anal. Calorim.* **2010**, *99*, 325–335. [[CrossRef](#)]
92. Matsuo, K.; Matsuoka, M. Solid-state polymorphic transition of theophylline anhydrate and humidity effect. *Cryst. Growth Des.* **2007**, *7*, 411–415. [[CrossRef](#)]
93. Zhu, H.; Yuen, C.; Grant, D.J.W. Influence of water activity in organic solvent + water mixtures on the nature of the crystallizing drug phase. 1. Theophylline. *Int. J. Pharm.* **1996**, *135*, 151–160. [[CrossRef](#)]
94. Suihko, E.; Ketolainen, J.; Poso, A.; Ahlgren, M.; Gynther, J.; Paronen, P. Dehydration of theophylline monohydrate—A two step process. *Int. J. Pharm.* **1997**, *158*, 47–55. [[CrossRef](#)]
95. Nunes, C.; Mahendrasingam, A.; Suryanarayanan, R. Investigation of the multi-step dehydration reaction of theophylline monohydrate using 2-dimensional powder X-ray diffractometry. *Pharm. Res.* **2006**, *23*, 2393–2404. [[CrossRef](#)]
96. Korang-Yeboah, M.; Rahman, Z.; Shah, D.A.; Khan, M.A. Spectroscopic-Based Chemometric Models for Quantifying Low Levels of Solid-State Transitions in Extended Release Theophylline Formulations. *J. Pharm. Sci.* **2016**, *105*, 97–105. [[CrossRef](#)]
97. González-González, J.S.; Zúñiga-Lemus, O.; del Carmen Hernández-Galindo, M. Hydrated Solid Forms of Theophylline and Caffeine Obtained by Mechanochemistry. *IOSR J. Pharm.* **2017**, *7*, 28–30. [[CrossRef](#)]
98. Parinet, J. Prediction of pesticide retention time in reversed-phase liquid chromatography using quantitative-structure retention relationship models: A comparative study of seven molecular descriptors datasets. *Chemosphere* **2021**, *275*, 130036. [[CrossRef](#)] [[PubMed](#)]
99. Kobayashi, Y.; Yoshida, K. Development of QSAR models for prediction of fish bioconcentration factors using physicochemical properties and molecular descriptors with machine learning algorithms. *Ecol. Inform.* **2021**, *63*, 101285. [[CrossRef](#)]
100. Wang, H.; Jia, Y.; Jing, G.; Wu, X. A novel toxicity prediction model for hydrazine compounds based on 1D–3D molecular descriptors. *Comput. Toxicol.* **2021**, *18*, 100169. [[CrossRef](#)]
101. Cysewski, P.; Przybyłek, M. Selection of effective cocrystals former for dissolution rate improvement of active pharmaceutical ingredients based on lipoaffinity index. *Eur. J. Pharm. Sci.* **2017**, *107*, 87–96. [[CrossRef](#)]
102. Przybyłek, M.; Recki, Ł.; Mroczynska, K.; Jeliński, T.; Cysewski, P. Experimental and theoretical solubility advantage screening of bi-component solid curcumin formulations. *J. Drug Deliv. Sci. Technol.* **2019**, *50*, 125–135. [[CrossRef](#)]
103. Bergström, C.A.S.; Larsson, P. Computational prediction of drug solubility in water-based systems: Qualitative and quantitative approaches used in the current drug discovery and development setting. *Int. J. Pharm.* **2018**, *540*, 185–193. [[CrossRef](#)]
104. Nordström, F.L.; Rasmuson, Å.C. Determination of the activity of a molecular solute in saturated solution. *J. Chem. Thermodyn.* **2008**, *40*, 1684–1692. [[CrossRef](#)]
105. 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. [[CrossRef](#)]
106. Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J.C.W. Refinement and parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, *102*, 5074–5085. [[CrossRef](#)]
107. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **2010**, *132*, 154104. [[CrossRef](#)] [[PubMed](#)]
108. COSMOtherm, version 20.0.0, revision 5273M; Dassault Systèmes: Vélizy-Villacoublay, France; Biovia: San Diego, CA, USA, 2020.
109. Tetko, I.V. Associative neural network. *Methods Mol. Biol.* **2008**, *458*, 185–202. [[PubMed](#)]
110. Minovski, N.; Župerl, Š.; Drgan, V.; Novič, M. Assessment of applicability domain for multivariate counter-propagation artificial neural network predictive models by minimum Euclidean distance space analysis: A case study. *Anal. Chim. Acta* **2013**, *759*, 28–42. [[CrossRef](#)] [[PubMed](#)]
111. Liu, R.; Wang, H.; Glover, K.P.; Feasel, M.G.; Wallqvist, A. Dissecting Machine-Learning Prediction of Molecular Activity: Is an Applicability Domain Needed for Quantitative Structure-Activity Relationship Models Based on Deep Neural Networks? *J. Chem. Inf. Model.* **2019**, *59*, 117–126. [[CrossRef](#)] [[PubMed](#)]
112. Durand, M.; Molinier, V.; Kunz, W.; Aubry, J.-M. Classification of Organic Solvents Revisited by Using the COSMO-RS Approach. *Chem. Eur. J.* **2011**, *17*, 5155–5164. [[CrossRef](#)]
113. Moity, L.; Durand, M.; Benazzouz, A.; Pierlot, C.; Molinier, V.; Aubry, J.M. Panorama of sustainable solvents using the COSMO-RS approach. *Green Chem.* **2012**, *14*, 1132–1145. [[CrossRef](#)]