Experimental and theoretical solubility advantage screening of bi-component solid curcumin formulations

Maciej Przybyłeka, Łukasz Reckia,b, Karina Mroczyńskaa, Tomasz Jelińskia, Piotr Cysewskia

a Department of Physical Chemistry, Pharmacy Faculty, Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Toruń, Kurpińskiego 5, 85-092, Bydgoszcz, Poland
b Pharmacy Faculty, Collegium Medicum in Bydgoszcz, Nicolaus Copernicus University in Toruń, Poland
c Research Laboratory, Faculty of Chemical Technology and Engineering, University of Technology and Life Sciences in Bydgoszcz, Seminarjynia 3, 85-326, Bydgoszcz, Poland

ARTICLE INFO

Keywords:
Curcumin
Solubility
Nutraceuticals
Excipients screening
Solubilizing agents
QSPR

ABSTRACT

A comprehensive experimental and theoretical screening was performed for identification of curcumin solubilizers. Experimental data led to formulation of a non-linear QSPR model correlating molecular descriptors with measured solubilities. The majority of synthesized binary systems exhibited a moderate enhancement of curcumin solubility, which was found to be the highest in the case of curcumin cocryrstallized with pyrogallol. New excipients for curcumin were found by utilization of the model within its applicability domain. It appears that a five-fold rise of curcumin solubility is the upper limit for this kind of formulations what was inferred from the screened 230 thousand compounds. Finally, theoretical analysis was extended on naturally occurring curcumin analogues, including demethoxycurcumin, bisdemethoxycurcumin, α-, β- and Ar-turmerones. In all cases, a list of coformers suitable for binary solids preparation with potential enhanced solubility was obtained. They can be treated as first choice lists for further experimental exploration of solubility enhancement of curcumin and its analogues.

1. Introduction

Curcumin, known also as Curcuma, is a hydrophobic polyphenol with various pharmacological effects including anti-inflammatory, antioxidant, antiproliferative and antiangiogenic activities. This compound belongs to a wide class of nutraceuticals, termed as curcuminoids which occur naturally in turmeric (Curcuma longa). This herb has been widely used in traditional Asian medicine [1,2]. It is well known that curcumin and its analogues have many beneficial effects on humans’ health. These compounds have been studied for their various anti-cancer activities [2,3], antioxidant actions [2,4–13], anti-inflammatory properties [2,14,15], anti-acidogenic [2,16] and neuroprotective potential [3,17–21]. The latter aspect, associated with treatment of neurological diseases, deserves particular attention. It is worth mentioning that in vitro tests and studies in rodents have shown that curcumin induces memory improvement [2,22], protects from brain mitochondrial dysfunction [2,23] and offers Alzheimer disease treatment [2,24–27]. The above-mentioned examples indicate that the phenomenon of curcumin lies in a very wide range of its potential applications. However, curcuminoids are generally poorly bioavailable and the low plasma and tissue levels of curcumin appear to be due to its poor solubility, rapid metabolism and rapid systemic elimination [2,28–30].

Many attempts have been made to improve or modify curcumin properties including solubility and bioavailability. One of the most frequently used methods is preparation of curcumin mixtures with various additives and excipients. A number of reports dealing with new curcumin compositions appeared recently, such as solid dispersions (eutectic mixtures, cocryrstals, co-amorphous mixtures, soluble polymer composites) with improved solubility and bioavailability [31–37], nanoparticles, nanocomposites and encapsulated systems exhibiting high antioxidant, anticancer and anti-diabetic potentials [38–42], nanofibers used for wound treatment [43,44], and metal complexes with enhanced ability for reducing proteins causing Alzheimer disease and enhanced antioxidant potential [45–47]. Interestingly, although curcumin exhibits high ability for metal ions complexation, it is rather hardly miscible with non-polymeric organic compounds in the solid state, since many more eutectic systems have been reported in the literature when compared to cocryrstals. Furthermore, theoretical cocrysralt screening studies showed that the miscibility of curcumin with a
popular class of coformers, i.e. phenolic acids, is very low [48]. Since curcumin cocrystals and eutectics can be applied for solubility enhancement [35,37], curcumin-based solid dispersions containing other organic compounds can be potentially useful for bioavailability enhancement. Indeed, it has been shown that addition of several phytochemicals, ferulic acid, xanthohumol, sesamin and naringenin can improve oral bioavailability of curcuminoids up to 88-fold [30].

There is a great need from the pharmaceutical industry to develop fast, efficient and high throughput screening procedures and theoretical models for solid dose improvement. Solid dispersions deserve particular attention due to the wide variety of excipients that can be used for tuning pharmaceutically relevant properties including solubility [49–58]. In recent years, there appeared several new binary mixtures models allowing for evaluation of cocrystal formation ability [48,59–61], thermodynamic properties prediction [62–64], dissolution rate [65] and solubility improvement assessment [66–68]. Our previous studies showed that simple molecular descriptors, easily computable from SMILES strings, can be very useful for such purposes [48,65]. The main goals of this paper are to perform screening studies on various curcumin solubility enhancers and to develop a theoretical model based on 1D and 2D descriptors allowing for fast excipient selection. In addition, in order to gain better insight into curcumin miscibility with different classes of compounds, enthalpy of mixing based on the post-quantum COSMO-RS methodology was calculated.

2. Materials and methods

2.1. Chemicals

All chemicals were purchased at the highest available purity and used as received. Curcumin, CAS: 458-37-7 and the following solid dispersions’ components, namely phenols (resorcinol, CAS: 108-46-3, pyrogallol, CAS: 87-66-1, hydroquinone, CAS: 123-31-9), phenolic acids (salicylic acid, CAS: 69-72-7, 3-hydroxybenzoic acid, CAS: 99-06-9, 4-hydroxybenzoic acid CAS: 99-96-7, 2,4-dihydroxybenzoic acid, CAS: 89-86-1, 2,5-dihydroxybenzoic acid, CAS: 490-79-9, 2,6-dihydroxybenzoic acid, CAS: 303-07-1, 3,4-dihydroxybenzoic acid CAS: 99-50-3, 2-hydroxy-1-naphthoic acid, CAS: 2283-08-1), dicarboxylic and hydroxycarboxylic acids (oxalic acid, CAS: 144-62-7, fumaric acid, CAS: 110-17-8, maleic acid, CAS: 110-16-7, glycolic acid, CAS: 79-14-1, citric acid, CAS: 77-92-9, DL-malic acid, CAS: 6915-15-7, l-tartaric acid, CAS: 110-17-8, maleic acid, CAS: 6915-15-7, l-tartaric acid, CAS: 87-69-4), methylxanthines (caffeine, CAS: 58-08-2, theophylline CAS: 58-55-9) saccharin, CAS: 81-07-2 and nicotinamide, CAS: 98-92-0 were obtained from Sigma-Aldrich. Other excipients namely glycine, CAS: 56-40-6 and urea, CAS: 57-13-6, as well as solvents, methanol CAS: 67-56-1 and ethanol, CAS: 64-17-5 were purchased from Avantor (Gliwice, Poland).

2.2. Curcumin binary solid mixtures preparation

In this study a mechanochemical method based on the approach described by Sanphui et al. [37] was applied. The protocol utilized liquid assisted co-grinding of 0.3 g of curcumin with equimolar amounts of excipients and 50 μL of ethyl alcohol. This procedure was carried out for 25 min using Retsch MM 200 ball mill equipped with 5 ml grinding jars and 5 mm stainless balls. The vibrational frequency was set to 25 Hz and two balls were used for each sample.

2.3. Solubility measurements

The equilibrium (thermodynamic) solubility of obtained powder samples was measured at 20°C using “shake-flask” approach. Firstly, the samples (0.05 g) were suspended in 10 ml of 40% (v/v) methanol-water system in 14 ml sealed glass vials and then mixed on a roller stirrer. After 48 h, the mixtures were centrifuged at 4000 rpm for 15 min and filtered using polytetrafluoroethylene (PTFE) membrane filters (0.20 μm pore size). The solutions were diluted ten times with methanol and measured against methyl alcohol using BIOSENS UV-VIS spectrophotometer. Each “Shake-Flask” experiment was repeated three times. The concentration of curcumin was determined from standard curve (λmax = 422 nm) prepared for curcumin methanol solutions.

2.4. ATR-FTIR measurements

The Attenuated total reflectance-Fourier-transform infrared spectra (ATR-FTIR) were measured using Alpha-PFT-IR equipment (Bruker Optik, Ettingen).

2.5. Computational protocol

Theoretical model of solubility was constructed using molecular descriptors derived directly from information encoded in SMILES strings, so that no real geometries of considered compounds were necessary. Since a variety of SMILES strings can be defined for a given compound, only the canonical ones were used, as provided by PubChem database [69]. Automatic data search was conducted using PubChemPy 1.0.4 [70]. This free of charge Python wrapper around the PubChem offers flexible scripting for retrieving information from the online database. Coformers were characterized by molecular descriptors computed using PaDEL 2.22 software [71] with interests restricted only to 1D and 2D indices. In total, 1444 parameters are offered by this package, which can be generated in seconds. This is an important aspect from the perspective of further screening, allowing for an exhaustive search for new compounds.

All computed descriptors underwent a simple preselection data reduction for eliminating those not computable for all compounds, interrelated to each other and with low variability. The remaining 996 parameters, considered as independent variables in the QSPR model, were directly used for the model formulation. For this purpose Multivariate Adaptive Regression Splines (MARSplines) [72] was used, as implemented in Statistica 12 [73]. This approach relies on a non-parametric procedure that makes no assumption about the underlying functional relationship between the dependent and independent variables. MARSplines solves regression-type problems based on a set of arbitrary defined basis functions. This algorithm is based on the gradual addition of new predictors or their combinations, so that the error of the model is minimized. The next stage is a reverse procedure, namely the equation is simplified by removing descriptors which give the smallest improvement of the model. This step is performed based on the Generalized Cross Validation (GCV) algorithm. The resulting model is defined as an ordinary QSPR regression equation with new descriptors resulting from splitting functions. Therefore, it is possible to use typical criterions for model quality assessments, as internal validation. In this study a leave many out (LMO) cross-validation parameters, Rm2 (adjusted determination coefficient), Kx (global correlation among descriptors) [74,75], RMSE (root-mean-square error), MAE (mean absolute error), F (Fisher ratio) and leave one out (LOO) determination coefficient Qm2 were calculated. In case of LMO approach, 30% of dataset was used for the test set. The model cross-validation step was performed automatically using QSARINS (http://www.qsar.in/) which is a comprehensive QSR/QSPR software developed by Gramatica et al. [76,77].

The mixing enthalpy was used for characterizing co-crystallization propensity using a procedure already applied for theoretical cocrystals screening [78]. It relies on computing the relative values of enthalpy of binary mixtures with respect to single components, presuming a hypothetical state of supercooled liquids at ambient condition. Such procedure requires conformational analysis of each monomer, which was done using COSMO conf software [79]. As a result, a series of structures were obtained, the geometries of which were optimized both in the gas and liquid states using RI-BP functional with TZVPD basis set. This step was followed by single point calculations using the same DFT
approach augmented with def2-TZVPD basis set at fine grid cavity. These quantum chemistry computations were performed using Turbomole 7 package [80] interfaced with TmolX 4.3 [81]. The values of enthalpies were computed using COMSOtherm [82] and BP-TZVPD-FINE parameter set.

3. Results and discussion

During the first stage of the study, the solubility measurements of newly synthesized binary systems formed by curcumin and 24 co-formers were reported. Additionally, two known cocrystals of curcumin were obtained for which dissolution profiles were measured. In the second step, a theoretical QSAR model was formulated. Next, the model was applied for screening of new potential solubility enhancers of curcumin. Also, the solubilities of curcuminoids sharing the same molecular core structure were screened using a QSPR model within its applicability domain. Finally, the possibility of extending applicability of the formulated model on turmerones, co-occurring in natural products with curcumin, was offered.

3.1. Experimental solubility advantage

In this study, solid mixtures of curcumin with different compounds were prepared and examined in terms of solubility improvement. The results of curcumin solid dispersions solubility measurements were presented in Table 1. As one can see, significant solubility advantage (> 1.5) can be observed in the case of two phenolic compounds (pyrogallol, catechol), two methylxanthines (theophylline, caffeine) and one amide (nicotinamide). Interestingly, there is no direct correlation between hydrophilicity or aqueous solubility of excipients and curcumin amide (nicotinamide). Interestingly, there is no direct correlation between hydrophilicity or aqueous solubility of excipients and curcumin solubility advantage [88–90]. Other important issues are the defects of crystalline structure, metastable polymorphs and incomplete conversion of starting components [88,89]. However, despite of these artifacts, the usability of the mechnochemical approach has been confirmed for cocrystal screening of various systems including curcumin [35,37,88,91].

Numerous studies have shown that vibrational spectroscopy, and especially FTIR-ATR technique, is an effective and reliable way of classification of pharmaceutical solid dispersions into homogeneous and inhomogeneous ones [92–100]. Indeed, any kind of strong interactions results in significant band shifts. It has been found that in the case of curcumin only pyrogallol and resorcinol can adopt cocrystal forms [37]. This was also confirmed by our measurements, as documented by absorption spectra presented in Fig. 1. The curcumin-pyrogallol system is characterized by absorption at 3533 cm\(^{-1}\), 3322 cm\(^{-1}\), 3226 cm\(^{-1}\), which can be assigned to OH stretching modes, \(\nu(OH)\). In the case of the molecular complex, OH groups in pyrogallol are involved in formation of new hydrogen bonds with curcumin, which can be inferred from the moderate blue-shift effect leading to the appearance of a new band at 3389 cm\(^{-1}\). In the case of curcumin-resorcinol system, cocrystal formation can be confirmed by a red-shift of curcumin absorption band at 3510 cm\(^{-1}\) to 3435 cm\(^{-1}\) coming from phenolic OH stretching mode in curcumin molecule [101]. In other cases, summarized in supplementary materials in Figs. S1–S24, the formation of new intermolecular interactions was not observed on the spectra and all absorption bands characterizing binary mixtures are undistinguishable from those recorded for pure components. It is worth noting, that immiscibility of nicotinamide, hydroquinone, 4-hydroxybenzoic acid, salicylic acid and \(\alpha\)-tartaric acid with curcumin was already reported by other authors [33,35] who have found that these compounds form simple eutectic systems.

3.2. FTIR-ATR solid dispersions characteristics

Both cocrystals and eutectics can be successfully used for the solubility enhancement [87]. However, the information about the solid-state miscibility of API and excipient is useful from a technological viewpoint, since not all methods suitable for cocrystal preparation, e.g. slow solvent evaporation, can be used for eutectic mixtures preparation. Fortunately, co-grinding technique can be applied both for cocrystals and eutectics synthesis. In general, three types of solid dispersions are expected to be formed as a result of co-grinding, namely cocrystals, simple eutectic systems and co-amorphous mixtures. The cocrystals are considered as new chemical compounds bonded via chemical interactions. Eutectics, which are stabilized by weaker intermolecular interactions, are less homogenous than cocrystals and co-amorphous mixtures. In contrary to co-amorphous dispersions, both eutectics and cocrystals are relatively stable systems. However, the possibility of amorphous phase formation ought to be taken into account when analyzing multicomponent solids obtained via co-grinding methods [88–90].

Table 1

<table>
<thead>
<tr>
<th>Excipient</th>
<th>(S \pm SD)</th>
<th>(SA)</th>
<th>Excipient</th>
<th>(S \pm SD)</th>
<th>(SA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pyrogallol</td>
<td>0.109 ± 0.002</td>
<td>2.27</td>
<td>nicotinamide</td>
<td>0.082 ± 0.001</td>
<td>1.71</td>
</tr>
<tr>
<td>2,6-dihydroxybenzoic acid</td>
<td>0.062 ± 0.001</td>
<td>1.28</td>
<td>saccharin</td>
<td>0.076 ± 0.002</td>
<td>1.58</td>
</tr>
<tr>
<td>4-hydroxybenzoic acid</td>
<td>0.061 ± 0.001</td>
<td>1.26</td>
<td>catechol</td>
<td>0.073 ± 0.003</td>
<td>1.51</td>
</tr>
<tr>
<td>oxalic acid</td>
<td>0.061 ± 0.001</td>
<td>1.26</td>
<td>2,4-dihydroxybenzoic acid</td>
<td>0.072 ± 0.001</td>
<td>1.49</td>
</tr>
<tr>
<td>2,5-dihydroxybenzoic acid</td>
<td>0.060 ± 0.001</td>
<td>1.24</td>
<td>nicotinic acid</td>
<td>0.072 ± 0.001</td>
<td>1.50</td>
</tr>
<tr>
<td>resorcinol</td>
<td>0.058 ± 0.002</td>
<td>1.20</td>
<td>urea</td>
<td>0.070 ± 0.002</td>
<td>1.46</td>
</tr>
<tr>
<td>fumaric acid</td>
<td>0.057 ± 0.001</td>
<td>1.19</td>
<td>2-hydroxy-1-naphthoic acid</td>
<td>0.068 ± 0.001</td>
<td>1.42</td>
</tr>
<tr>
<td>citric acid</td>
<td>0.057 ± 0.001</td>
<td>1.19</td>
<td>glycolic acid</td>
<td>0.068 ± 0.001</td>
<td>1.42</td>
</tr>
<tr>
<td>D-(\alpha)-malic acid</td>
<td>0.057 ± 0.002</td>
<td>1.18</td>
<td>glycine</td>
<td>0.067 ± 0.001</td>
<td>1.39</td>
</tr>
<tr>
<td>3-hydroxybenzoic acid</td>
<td>0.056 ± 0.001</td>
<td>1.16</td>
<td>3,4-dihydroxybenzoic acid</td>
<td>0.064 ± 0.001</td>
<td>1.33</td>
</tr>
<tr>
<td>(\alpha)-tartaric acid</td>
<td>0.048 ± 0.001</td>
<td>0.99</td>
<td>maleic acid</td>
<td>0.064 ± 0.001</td>
<td>1.33</td>
</tr>
<tr>
<td>theophylline</td>
<td>0.085 ± 0.002</td>
<td>1.76</td>
<td>hydroquinone</td>
<td>0.051 ± 0.001</td>
<td>1.06</td>
</tr>
<tr>
<td>caffeine</td>
<td>0.083 ± 0.002</td>
<td>1.72</td>
<td>salicylic acid</td>
<td>0.047 ± 0.001</td>
<td>0.98</td>
</tr>
</tbody>
</table>

\(a\) Solubility (S) [mg/ml] with standard deviations (SD).

\(b\) Solubility advantage (SA) defined as a ratio between solubility of binary mixtures with respect of curcumin, which was found to be equal to 0.048 ± 0.001 mg/ml.

3.3. Solubility enhancement modeling

Taking advantage of the measured solubility data of curcumin solid dispersions containing various additives, an attempt of formulating a theoretical model was undertaken. It is worth mentioning that this is not a trivial task since many efforts have been made for predicting
solubility using QSPR methods [102–108] with varying success. The major difficulty comes from the expectation that the dissolution process is driven by some linear trend in relation with some molecular parameters what is surely not the case due to intrinsic thermodynamic relationship between activity of solute and activity of solvent [109]. That is why application of methods inherently accounting for non-linearity, as for example neural networks, attracted much attention. In this paper, another strategy was adopted taking advantage of nonlinear MARS-plines methodology [72] which has been successfully applied for physicochemical properties and biological activities [110–115]. This approach allows for formulation of models more accurate than traditional multivariable QSPR method [113,114] and in some cases more accurate than artificial neural networks [111,112,115]. In this study, the chemical information encoded in SMILES strings was used and molecular descriptors were computed using PaDEL software [71].

Chenical information encoded in SMILES strings was used and molecular descriptors were computed using PaDEL software [71]. The binary solids were characterized by a API-coformer similarity measure [48] defined simply as the absolute difference between a given molecular descriptor of curcumin and the particular coformer. These values were normalized and after elementary data reduction, conducted by elimination of less significant non-orthogonal descriptors, used for model generation. Based on the above methodology, the following equation was obtained:

\[
S_{[\text{mg/ml}]} = 0.0797 - 0.1744 \cdot \text{minwHBa} + 0.1607 \cdot \text{mindssC} - 0.0735 \cdot \text{mindssC AATSC6i} + 0.1796 \cdot \text{mindssC AATSC6i} + 0.1757 \cdot \text{mindssC GATS1m AATSC6i} - 0.0969 \cdot \text{mindssC AATSC6i GATS5 p} + 0.0284 \cdot \text{minaasC} \cdot \text{ATSC6 m} - 0.0145 \cdot \text{ATSC6m}
\]

(1)

Statistics:

- Fitting criteria: N = 26, R² = 0.987, \( R_{adj}^2 = 0.981 \), Kxx = 0.509, RMSE = 0.002, MAE = 0.001, F = 166.12
- Internal validation criteria: \( Q_{LOO}^2 = 0.965 \), RMSE = 0.003, MAE = 0.002

Despite slightly worse statistics, it also offers quite reasonable accuracy. The graphical comparison of both models is provided in Fig. 2. The common set of descriptors in both models reveals the main structural origin of curcumin solubility advantage. The ability to form weak hydrogen bonds is associated with inclusion of the minwHBa descriptor in the model with high contribution. The presence of double bonded ionization potential (AATSC6i), by mass (ATSC6m, GATS1m), by Sanderson electronegativity (MATS3e) and by polarizability (GATS5p) [119]. Since normalized values of all descriptor distributions were used for model formulation, the values of descriptors directly reflect their importance in the model. Interestingly, when the parameters of the lowest contributions are removed, it is possible to obtain even a simpler model. Indeed, after exclusion of GATS1m, GATSSp and MATS3e the re-optimization of the remaining weights leads to a much simpler model with only slight loss of accuracy. The new simplified equation is defined as follows:

\[
S_{[\text{mg/ml}]} = 0.0823 - 0.2017 \cdot \text{minwHBa} + 0.1874 \cdot \text{mindssC} - 0.0486 \cdot \text{mindssC AATSC6i} + 0.033 \cdot \text{minaasC} \cdot \text{ATSC6 m} - 0.021 \cdot \text{ATSC6m}
\]

(2)

Statistics:

- Fitting criteria: N = 26, R² = 0.923, \( R_{adj}^2 = 0.904 \), Kxx = 0.351, RMSE = 0.003, MAE = 0.003, F = 48.07
- Internal validation criteria: \( Q_{LOO}^2 = 0.833 \), RMSE = 0.016, MAE = 0.005

The above equation is characterized by a good fitting quality, expressed by several criteria such as \( R_{adj}^2 \) and MAE. However, from the reproducibility viewpoint, the model can be evaluated using a validation procedure. Therefore the obtained equation was analyzed using internal validation criteria. The high value of \( Q_{LOO}^2 \) parameter expressing the predictive ability, as well as low values of error measures (RMSE and MAE) confirm its usefulness. As it can be seen from eq. (1), the model involves only eight descriptors belonging to two classes of molecular indices. The first one comprises electrotropological state (E-state) atom type descriptors including minwHBa (minimum E-state value corresponding to weak hydrogen bonding acceptors), mindssC (minimum E-State for dssC atom type) and minaasC (minimum E-state for aasC atom type) [116–118]. The second set encompasses Broto-Moreau, Moran and Geary Autocorrelation descriptors weighted by first
and aromatic-type carbon atoms is emphasized by $\text{minasC}$ and $\text{minmaxC}$, respectively. The other two descriptors involved in the simplified model characterize autocorrelation function computed with different weightings. These two descriptors have smaller contributions to the formula characterizing solubility of curcumin-containing systems. In general, both models quite accurately describe the trend of solubility, except for the lowest solubility region in the case of the simplified model.

3.4. Application of QSRR model for screening of new coformers

The encouraging correlation between experimentally obtained solubilities and the ones resulting from the formulated QSRR model allows for screening of new coformers, suitable for enhancing curcumin dissolution. For this purpose, the PubChem database was used for searching of analogues of coformers with the highest experimentally proved solubility advantage. Potential new coformers were searched in related records section, which comprises a variety of potential sources including medications, literature, 3D structures, bioactivities and also literature related species. The search was quite comprehensive since it involved similarities in connectivity, parent compound and conformers. The wide-ranging analogues were retrieved for pyrogallol (22273), caffeine and theophylline (41403), nicotinamide (62890), saccharin (14723) and pyrocatechol (71050). The numbers of potential structures taken for consideration were provided in parenthesis. All kinds of isopropomers, radicals and charged species were removed from this initial list. Finally, the list of potential coformers was also augmented by 1149 structures that constitute EAFUS/GRAS list and 15807 common coformers of cocrystals found in CSD. In total, the list of initial structures comprised almost 230 thousand chemicals. This step was automated using a simple Python script. For all structures, the values of molecular descriptors were computed using PaDEL software [71]. Despite the fact that the list of compounds was very extensive, this step was not tedious since the time required for generation of the whole set of 1444 PaDEL descriptors is in the order of a fraction of a second. The whole procedure is easily automatable for retrieving the required subset of descriptors. The application of the model provided by equation (1) relies only on eight molecular descriptors which were collected for the whole set of initial coformers. It is obvious that not all of these compounds can be directly used for the screening purposes, since limitations imposed by the training set are to be enforced also on all potential ingredients of new binary solids with curcumin. The major restriction comes from the range of values used in the normalization procedure. Hence, hypothetical binary mixtures were excluded from further analysis if their characteristics were outside of minimal and maximal values for any of the eight descriptors included in the model. This exclusion criterion, often termed as applicability domain, is important from the methodological point of view since the inference should be based on matching the new information to the one encompassed by the training set. This was a very restricting exclusion criterion as only 4% of compounds fulfilled this requirement, reducing the length of the screening list down to 9100 compounds. The highest restrictions were imposed by $\text{minssC}$ and $\text{minmaxHBA}$ for which only about 4% of structures were included. On the contrary, the least restrictive was the $\text{minasC}$ parameter for which as many as 80% of structures have values within the training set interval. Finally, identification of potential coformers enhancing curcumin solubility was assessed by application of eq. (1) and compared to experimental values. Despite the fact that only 3.5% of the initial population was pointed out as potential curcumin solubility enhancers, still as many as 492 compounds were found to be potentially applicable as more effective than pyrogallol. The total number of coformers worth considering, i.e. for which $\text{SA}>2$, comprises 965 compounds. The results of the performed screening procedure are schematically documented in Fig. 3, which shows that the applied procedure was successful in identification of many coformers potentially suitable for enhancing curcumin solubility. This is shown by collection of points above the red line characterizing solubility of curcumin-pyrogallol cocrystal. Potentially effective enhancers can be found practically in every group of analogues with only one exception. It is interesting to notice that no compound belonging to the EAFUS/GRAS subset was found as a potential solubility enhancer of curcumin. The list of top 5% of potential solubility enhancers is provided in Table 2. Taking this list as a starting point, it is possible to prolong the search for curcumin solubility enhancers by performing series of new experiments. However, there is a serious limit for solubility advantage since the highest value was found to be only 2.3 times higher than for the cocrystal with pyrogallol, which stands for 5.2 times higher solubility with respect to curcumin. This value seems to be the maximum that can be expected from binary solid formulations.

Application of the model given by eq. (1) within its application domain allowed for identification of many potentially effective enhancers of curcumin solubility. It is also interesting to check the nature of potential binary solids because there is a serious advantage of co-crystallization over preparation of simple physical mixtures of the drug and excipient [120]. For this distinction, the $H_{\text{mix}}$ values were used as a criterion. As it was documented beforehand [78], the classification of binary solids into either eutectic mixtures or cocrystals might be obtained assuming the model of mixing components in a hypothetical supercooled state at room temperature. Despite ignoring the fusion energetics, this simple model can be used as a first aid classification tool. The proper selection of the threshold value of $H_{\text{mix}}$ can minimize the number of eutectics classified as cocrystals and vice versa. However, it was already pointed out [48] that the main shortcoming of this approach is the rejection of many potential cocrystals from the analysis. Taking advantage of the simplicity of this model, the values of $H_{\text{mix}}$ were computed for selected pairs and used for classification of potential binary systems comprising curcumin. As it is presented in Fig. 4, another challenge of using $H_{\text{mix}}$ is clearly visible. The inaccuracy of the method results in misclassification of negative cases as positive ones. However, it is worth emphasizing that the systems studied here were classified as physical mixtures based only on a limited number of instrumental measurements. For proper ensuring that the studied binary systems involving curcumin are really simple binary eutectic systems their whole phase diagrams should be measured. On the other hand, the lack of cocrysalization is generally the first sign of inability to form
homogeneous solid dispersions. The distribution presented in Fig. 4 still provides useful information and advises to select for further experiments those systems which are above the gray line (potential solubility enhancers), irrespectively of their position with respect to the red line (classification into physical mixtures or cocrystals). This spectacular failure of the $H_{mix}$ as a cocrystals classification tool is the first published sight of the necessity of a serious revision of this criterion.

### 3.5. Application of QSPR model for screening of solubility enhancement of curcuminoids

Apart from curcumin itself also its analogues occurring in turmeric, including α-, β- and Ar-Turmerone. They also suffer from low solubility in including α-, β- and Ar-Turmerone. They also suffer from low solubility in.

### 3.6. Application of QSPR model for screening of solubility enhancement of turmerones

The biological activities of Turmeric are not only limited to curcuminoids since also turmerones can be derived from essential oils. The minoids since also turmerones can be derived form of essential oils. The

### Table 2

The characteristics of top 5% of potential solubility enhancers of curcumin identified during the comprehensive screening procedure. The values of solubility and solubility advantage were computed using eq. (1).

<table>
<thead>
<tr>
<th>group</th>
<th>Chemical name</th>
<th>CID</th>
<th>$S_{[mg/ml]}$</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>3-fluorobenzene carbothioamide</td>
<td>2782800</td>
<td>0.249</td>
<td>5.18</td>
</tr>
<tr>
<td>3</td>
<td>3-fluorobenzencarbothiohydrazide</td>
<td>55281616</td>
<td>0.221</td>
<td>4.58</td>
</tr>
<tr>
<td>5</td>
<td>4-ethenyl-3-((2-methylprop-1- enyl)benzene-1,2-diol)</td>
<td>909091291</td>
<td>0.218</td>
<td>4.53</td>
</tr>
<tr>
<td>3</td>
<td>4-fluorobenzencarbothiohydrazide</td>
<td>18534968</td>
<td>0.212</td>
<td>4.40</td>
</tr>
<tr>
<td>6</td>
<td>5-((3-fluorophenyl)-3,4-dihydro 2H-pyrole</td>
<td>-</td>
<td>0.208</td>
<td>4.32</td>
</tr>
<tr>
<td>3</td>
<td>3-(6-fluoro-5-methylpyridin-3-yl)-4,5-dihydro 1,2-oxazole</td>
<td>22719492</td>
<td>0.204</td>
<td>4.24</td>
</tr>
<tr>
<td>2</td>
<td>5-(difluoromethoxy)-4,6-dimethyl-1,3-dihydrobenzimidazole-2-thione</td>
<td>20189582</td>
<td>0.204</td>
<td>4.24</td>
</tr>
<tr>
<td>3</td>
<td>6-ethyl 2-amino-4-fluoro-5-methylbenzenecarbothioate</td>
<td>53946600</td>
<td>0.203</td>
<td>4.21</td>
</tr>
<tr>
<td>5</td>
<td>4-(1(1E,3Z)-2-methylpent-1- enyl)benzene-1,2-diol</td>
<td>89013525</td>
<td>0.202</td>
<td>4.19</td>
</tr>
<tr>
<td>3</td>
<td>4-(2-fluoro-6-methoxyphenyl)-1,2,3,6-tetrahydropyridine</td>
<td>69094214</td>
<td>0.201</td>
<td>4.18</td>
</tr>
<tr>
<td>5</td>
<td>(2E)-2-ethylidene-5-fluorobenzimidazole</td>
<td>58549958</td>
<td>0.194</td>
<td>4.03</td>
</tr>
<tr>
<td>3</td>
<td>1-ethylsulfonyl-2-[[4-fluorophenyl]methylsulfanyl]-4,5-dihydroimidazole</td>
<td>2148588</td>
<td>0.194</td>
<td>4.03</td>
</tr>
<tr>
<td>5</td>
<td>(6-fluoro-2H-1,4-benzoxazin-3-yl) methanamine</td>
<td>70010484</td>
<td>0.193</td>
<td>4.02</td>
</tr>
<tr>
<td>3</td>
<td>N-[4-(fluorophenyl)thiyl]-1-pyrindin-3-ylethanimine</td>
<td>58215735</td>
<td>0.192</td>
<td>3.98</td>
</tr>
<tr>
<td>6</td>
<td>4-(Fluorophenyl)thiourea</td>
<td>693061</td>
<td>0.191</td>
<td>3.97</td>
</tr>
<tr>
<td>3</td>
<td>4-fluoro-N-[[1(1)-pyridin-3-ylethyldeneamino]janline</td>
<td>5875641</td>
<td>0.19</td>
<td>3.95</td>
</tr>
<tr>
<td>5</td>
<td>7-fluoro-1-methyl-3,4-dihydroimidazoline</td>
<td>18613806</td>
<td>0.19</td>
<td>3.94</td>
</tr>
<tr>
<td>7</td>
<td>1,4,5,6,7-pentadaeterio-2-deuterioybenzimidazole</td>
<td>59638556</td>
<td>0.188</td>
<td>3.90</td>
</tr>
<tr>
<td>7</td>
<td>N-[4-(fluorophenyl)thiyl]-3a,4,5,6,7a-hexahydro-1H-benzimidazol-2-amine</td>
<td>68560894</td>
<td>0.187</td>
<td>3.88</td>
</tr>
<tr>
<td>1</td>
<td>2-(3-methylbut-2-enyl)benzene-1,3-diol</td>
<td>12335852</td>
<td>0.183</td>
<td>3.80</td>
</tr>
<tr>
<td>4</td>
<td>6-fluoro-1,1-dioxo-1H-1$\tilde{S}$ (6) 2-benzothiadiazine-3-thione</td>
<td>53729998</td>
<td>0.182</td>
<td>3.79</td>
</tr>
<tr>
<td>4</td>
<td>5-prop-1-en-2-ylypyridine-2-sulfonic acid</td>
<td>18639054</td>
<td>0.182</td>
<td>3.78</td>
</tr>
<tr>
<td>3</td>
<td>2-fluorobenzene carbothioamide</td>
<td>2734821</td>
<td>0.178</td>
<td>3.69</td>
</tr>
<tr>
<td>5</td>
<td>5-fluoro-1-methyl-3,4-dihydroimidazoline</td>
<td>58416485</td>
<td>0.175</td>
<td>3.63</td>
</tr>
<tr>
<td>3</td>
<td>2-(4-fluoro-2-methylphenyl)-3-pyridin-3-ylyprop-2-enal</td>
<td>54114351</td>
<td>0.169</td>
<td>3.51</td>
</tr>
</tbody>
</table>

Fig. 4. Classification of binary solids comprising curcumin using the criterion of solubility ($S$) enhancement and the probability of cocrystallization characterized by $H_{mix}$ values. Upper shadowed region collects solubility enhancers. Black full circles represent experimentally studied systems.

3.6. Application of QSPR model for screening of solubility enhancement of turmerones

The biological activities of Turmeric are not only limited to curcuminoids since also turmerones can be derived form of essential oils. The treatment effectiveness of these sesquiterpenes, specific for the Curcuma genus and responsible for the aroma of this spice, was documented in numerous studies [122–124]. Since the core of turmerones structure is analogous to curcuminoids, it seems to be valuable to expect that intrinsic solubility will be very low. This is justified not only by qualitative observations during extraction from natural sources but also by the estimated values (group contribution method) [121] provided by ChemAxon (http://www.chemaxon.com), as they are equal to 0.020 mg/ml and 0.017 mg/ml for Demethoxycurcumin and Bisdemethoxycurcumin, respectively. This is a reasonable estimation, at least in the order of magnitude, since the corresponding value for curcumin equals 0.024 mg/ml what is acceptable if confronted with our experimental measurement providing value 0.048 mg/ml. Moreover, the solubilities of all three curcuminoids are to be expected as quite similar and it is also reasonable to expect that binary solids of all three curcuminoids will also have similar properties. Application of our QSPR model for predicting solubilizers of Demethoxycurcumin and Bisdemethoxycurcumin resulted in quite rich distributions documented in Fig. 5. The list of potential coformers seems to be as extended as in the case of curcumin. It is therefore safe to assume that future experimenters can take items from the obtained list as the first choice for new solid formulations.

### Image 130

Fig. 4. Classification of binary solids comprising curcumin using the criterion of solubility ($S$) enhancement and the probability of cocrystallization characterized by $H_{mix}$ values. Upper shadowed region collects solubility enhancers. Black full circles represent experimentally studied systems.

4. Application of QSPR model for screening of solubility enhancement of turmerones

The biological activities of Turmeric are not only limited to curcuminoids since also turmerones can be derived form of essential oils. The treatment effectiveness of these sesquiterpenes, specific for the Curcuma genus and responsible for the aroma of this spice, was documented in numerous studies [122–124]. Since the core of turmerones structure is analogous to curcuminoids, it seems to be valuable to extend our analysis to the set of three naturally accruing compounds including α-, β- and Ar-Turmerone. They also suffer from low solubility in...
polar media including water, so dissolution enhancement is valuable from the perspective of their bioavailability. Using the same set of potential coformers the values of molecular descriptors characterizing binary mixtures were computed using previously formulated QSPR model. Again, a careful analysis was conducted and exclusion criteria were imposed on all considered pairs. As expected, several possible candidates for solubility enhancers were found. The distributions of theoretical screening are provided in Fig. 6. Unfortunately, the exact solubilities of these compounds are unknown and only their very low dissolution was observed. The estimated values using group contribution method [121] approach provided by ChemAxon (http://www.chemaxon.com) suggest the values 0.029 mg/ml, 0.043 mg/ml 0.018 mg/ml for AR, α- and β-turmerone, respectively. These values suggest slightly higher solubility of α-turmerone compared to the other two but still solubility advantage offered by binary solid formulation might be beneficial. In general, the applied screening procedure resulted in a much lower number of candidates identified for this set of curcumin analogues, although there still is a non-trivial selection of coformers. In the case of aromatic Turmerone 104 compounds were identified as potentially better solubilizers compared to pyrogallol. For α- and β- Turmerones these numbers are slightly lower and equal to 80 and 74, respectively. The obtained distributions are provided in Fig. 6.

4. Conclusion

The effective utilization of various pharmacologically positive effects of curcumin documented in many reports encounters one major obstacle, which is its poor solubility. This issue was directly addressed in this study by both experimental and theoretical screening. It was proven that the majority of experimentally synthesized binary mixtures have some potential of enhancing the concentration of curcumin in methanol-water solutions. Unfortunately, it is very difficult to prepare cocrystals with curcumin and among 26 studied binary mixtures only two were identified as homogeneous systems. These two were already known and their structures were described in detail [37]. According to our experiments, they offer significant solubility advantage doubling the concentration of curcumin in solution. The rest of new binary systems were classified by our experiments as simple binary eutectics. Despite this fact, they also offer some solubility advantage and only one system was found as being unable to elevate the concentration of curcumin. Theoretical screening began with formulating the mathematical relationship between molecular descriptors used for definition of orthogonal factors. This non-linear formula was able to reproduce experimental solubility with high accuracy. After a careful definition of the applicability domain, an exhaustive screening was performed for finding new coformers suitable for enhancing the solubility of curcumin. It appears that the five-fold solubility advantage was found as the upper limit based on the screening of 230 thousand chemicals. The applicability of the provided model for screening of solubilizers of curcumin analogues was also discussed. The obtained lists of potential solubility enhancers formulate to-do lists as the aid for further experimental investigations.
Conflicts of interest

The authors declare no conflict of interest.

Acknowledgement

Technical assistance during data collection and manuscript preparation of Tomasz Miernik and Anna Cieślińska is acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jddst.2019.01.023.

Fig. 6. Distributions of predicted values of solubility (S) of a given turmerone forming hypothetical binary solids with one of constituents from the screening set. Full circles (dataset 0) denote experimentally measured solubilities. The black solid line represents level of solubility measured in the case of curcumin-pyrogallol cocrystal, while dashed black line stands for experimental solubility of curcumin. The gray lines characterize solubility according to ChemAxon estimation for a given turmerone (solid line) and curcumin (dashed line).

References


