Journal of Molecular Liquids 367 (2022) 120365

Contents lists available at ScienceDirect

Journal of Molecular Liquids

journal homepage: www.elsevier.com/locate/molliq

Solubility analysis of 18 active pharmaceutical ingredients and intermediates in the non-polar solvents dioxane, toluene and cyclopentyl methyl ether

Carlos Moreno-Leon [a], Corin Mack [a], Sudipta Roy [b], Joop H. ter Horst [a,c,*]

[a] EPSRC Centre for Innovative Manufacturing in Continuous Manufacturing and Crystallisation (CMAC), Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS), Glasgow G1 1RD, UK
[b] Chemical and Process Engineering, Strathclyde University, James Weir Building, Glasgow G1 IXJ, UK
[c] Univ Rouen Normandie, Laboratoire Sciences et Méthodes Séparatives (SMS), UR 3233, F-76000 Rouen, France

ARTICLE INFO

Article history:
Received 27 May 2022
Revised 28 August 2022
Accepted 10 September 2022
Available online 24 September 2022

Keyword:
Thermodynamics
Solubility
Thermodynamic models
Activity coefficient
Non-polar solvents
Small organic molecules

ABSTRACT

The temperature dependent solubility of 18 organic compounds, of which some are active pharmaceutical ingredients, was measured in 1,4-dioxane, toluene and cyclopentyl methyl ether over a temperature range between 0 and 70 °C using a turbidity method. The solubility of the studied compounds in the three solvents increases with temperature in all binary solute-solvent systems, and all compounds exhibit a higher solubility in 1,4-dioxane than in toluene and cyclopentyl methyl ether. The non-ideal behaviour of the binary systems was evaluated in terms of the activity coefficient. All the compounds in toluene presented a positive deviation from ideal solution, while both positive and negative deviations were observed for the model compounds in the other two apolar solvents. From the solubility and the activity coefficient, a number of thermodynamic models including van't Hoff, Apelblat, λh, Margules, van Laar, Wilson, and the Non-Randomness Two Liquids (NRTL) were used for the correlation of the experimental data. The NRTL and the van't Hoff equations present closer overall correlations, while the Margules and the λh models show deviations >5 % for several compounds in the three solvent systems. The mixing thermodynamic properties of the several solute-solvent systems were further analyzed by the van 't Hoff and Wilson equations for a better understanding of the solution behaviour. The analysis revealed systems with apparent ideal behaviour where counteracting non-ideality effects cancel each other's.

1. Introduction

In the manufacturing of highly regulated products such as pharmaceuticals and fine chemicals, crystallization from solution is commonly employed, since this separation technology enables the formation of a crystalline particulate product with a purity close to 100 % in a single step [1–3]. The temperature dependent solubility of crystal compounds in a specific solvent strongly influences design process parameters such as productivity, yield, and purity of the crystallization process [4,5]. The maximum yield of a batch-wise cooling crystallization, for instance, can be defined as the fractional approach towards removal of all solute from the solution, which is a function of the solubility at the final batch temperature. Thermodynamic solubility information of a binary system depends on solid state properties such as heat of fusion and melting temperature and on non-ideal solution behaviour, which both are difficult to estimate and describe accurately. The design and optimization of the purification process require the knowledge of phase equilibrium thermodynamics, which demands, on the one hand, accurate solubility measurements as a function of the temperature and, on the other hand, accurate thermodynamic models to describe these solubilities.

Over the last decades, efficient experimental methods have been developed to access reliable solubility data [5]. Amongst them, a temperature variation approach has been implemented in automated and commercially available equipment, which allows for faster measurements while minimizing manual labour [5]. As solubility generally increases with temperature, by increasing the temperature of a suspension, the solids gradually dissolve and at some temperature the system turns into a clear solution. This change of the system from suspension to clear solution can be captured by turbidity measurements, resulting in a clear point tem-

* Corresponding author at: EPSRC Centre for Innovative Manufacturing in Continuous Manufacturing and Crystallisation (CMAC), Strathclyde Institute of Pharmacy and Biomedical Sciences (SIPBS), Glasgow G1 1RD, UK.
E-mail address: joop.terhorst@strath.ac.uk (J.H. ter Horst).

https://doi.org/10.1016/j.molliq.2022.120365
0167-7322/

perature for a given sample composition, which can be assumed to be equal to the saturation temperature of the system if the temperature increase is sufficiently slow and the equilibrium between solid and solution is rapidly reached [6]. Solubility data can then be correlated, interpolated and to a certain extend extrapolated with temperature by using thermodynamic models such as van't Hoff or the activity coefficient model of Wilson [7].

Isolator solvents enable the use of alternative energy sources as a process intensification tool in crystallization that may maximize the phase separation efficiency. For instance, Weiwei *et al.*, [8] demonstrated that a strong inhomogeneous electric field induces particle movement in a suspension, and the particles collect at the surface of one of the two electrodes, depending on the solid compound. The combination of electric fields with crystallization allows for instance the separation of organic crystalline compounds from their mixed suspension in isolator solvents such as 1,4-dioxane. Therefore, building a large solubility database of organic compounds in isolator solvents has significant benefits for the development of such alternative purification processes. Here, the temperature dependent solubility behaviour of 18 organic compounds of which some are active pharmaceutical ingredients (API) is assessed in the non-polar solvents 1,4-dioxane, toluene, and cyclopentyl methyl ether (CPME) at a temperature range between 0 and 70 °C. Following solubility analysis, the activity coefficients of the experimental solute-solvent systems are calculated from experimental data and the ideal solubility. Then, six thermodynamic models were used to correlate the measured and calculated activity coefficient. The gained information alongside the thermodynamic properties of the equilibrium are used to rationalise the measured solubility and the solution behaviour, which at the same time may contribute to the development of thermodynamic predictive methods.

2. Materials and methods

2.1. Materials

Details of the 18 model compounds and 3 solvents used in this study are listed in Table 1, and their chemical structure are presented in Fig. 1. All the materials were used as received without further purification.

2.2. Differential scanning calorimetry (DSC)

The melting point (T_m) and enthalpy of fusion (ΔH_f) of each model compound was measured in a DSC214 Polyma calorimeter (Netzsch Geratebau GmbH, UK), calibrated using indium as a standard substance. Approximately 5 mg of pure compounds, weighted using an analytical balance (Sartorius Quintix-124-1S), were placed into 25 μL aluminium pans with pierced lids. The samples were subjected to a heating cycle from 20 °C to a maximum of 250 °C at a rate of 10 $Kmin^{-1}$ while being continuously purged with helium. An empty aluminium pan was used as a reference, and the retrieved data was processed by the Netzsch Proteus software (v.6.1) to obtain the melting temperature and the enthalpy of fusion of a given crystalline compound based on the onset temperature and the area under the main thermodynamic event in the generated DSC curve.

2.3. Solubility

The temperature dependent solubility of 10 APIs and intermediates commonly used in the pharmaceutical industry, 2ABA, 2HBA, ASP, CAF, CBZ, RS-NPX, S-NPX, PHZ, PIC and SSA, was measured in a Crystal16 setup (Technobis Crystallization Systems, The Netherlands) at a temperature range between 12 °C and 70 °C in pure 1,4-dioxane, and between 0 °C and 70 °C in toluene and CPME. The lower boundary temperature for dioxane was set at 12 °C to avoid solubility measurements below the melting temperature of dioxane, while the lower and higher boundary temperatures of 0 and 70 °C were set for convenience. In addition, solubility measurements in 1,4-dioxane at the same temperature range were carried out on 8 additional organic compounds: 4ABA, 4AMP, 3HBA, 4HBA, 4NBA, INA, NA, and PCT. A solubility analysis of these 8 latter compounds was attempted in toluene and CMPE under similar conditions. However, complications such as formation of fouling or solubility levels below the instrument limit (<1 mg/mL), prevented the analysis from the accurate determination of clear point temperatures.

Solubility measurements were carried out on suspensions of accurately known compositions of the model compound in the pure solvent system, individually prepared in 1.8 mL high-performance liquid chromatography (HPLC) glass vials. Both the amount of model compound and the amount of added solvent

Table 1
Details of the crystalline materials used in the present study. Alongside the CAS number of the product, the number between brackets refers to the supplier: [A] Sigma-Aldrich (Merck KGaA, Germany), [B] VWR Chemicals (Avantor, UK), and [C] Tokio Chemicals Industry Co., Ltd. (Japan).

Material	Acronym	Chemical Formula	Purity [%] ≥	Molecular Weight [g/mol]	CAS Number
2-Aminobenzoic acid	2ABA	$C_7H_7NO_2$	98.0	137.1	118-92-3 [A]
4-Aminobenzoic acid	4ABA	$C_7H_7NO_3$	99.0	137.1	150-13-0 [B]
4-Aminophenol	4AMP	$C_7H_7NO_4$	98.0	109.1	123-30-8 [A]
Aspirin	ASP	$C_9H_8O_4$	99.0	180.2	50-78-2 [B]
Caffeine	CAF	$C_8H_{10}N_4O_2$	99.0	194.2	58-08-2 [A]
Carbamazepine	CBZ	$C_{15}H_{12}N_2O$	98.0	236.3	298-46-4 [A]
2-Hydroxybenzoic acid	2HBA	$C_7H_6O_3$	99.0	138.1	69-72-7 [A]
3-Hydroxybenzoic acid	3HBA	$C_7H_6O_4$	99.0	138.1	99-06-9 [A]
4-Hydroxybenzoic acid	4HBA	$C_7H_6O_5$	99.0	138.1	99-96-7 [A]
Isonicotinamide	INA	$C_6H_6N_2O$	99.0	122.1	1453-82-3 [A]
RS-Naproxen	RS-NPX	$C_{14}H_{13}NaO_3$	98.0	230.3	23981-80-8 [C]
S-Naproxen	S-NPX	$C_{14}H_{13}NaO_3$	99.0	230.3	22204-53-1 [C]
Nicotinamide	NA	$C_6H_6N_2O$	99.5	122.1	98-92-0 [A]
4-Nitrobenzoic acid	4NBA	$C_7H_5NO_4$	98.0	167.1	62-23-7 [A]
Paracetamol	PCT	$C_8H_9NO_2$	98.0	151.2	103-90-2 [A]
Phenazine	PHZ	$C_{12}H_8N_2$	99.0	180.2	92-82-0 [B]
Picolinamide	PIC	$C_6H_6N_2O$	98.0	122.1	1452-77-3 [A]
Salycylsalicylic acid	SSA	$C_{14}H_{10}O_5$	97.5	258.2	552-94-3 [B]
Cyclopentyl methyl ether	CPME	$C_6H_{12}O$	99.9	100.2	5614-37-9 [B]
1,4-Dioxane	dioxane	$C_4H_8O_2$	99.8	88.1	123-91-1 [A]
Toluene	toluene	C_7H_8	98.0	92.1	108-88-3 [B]

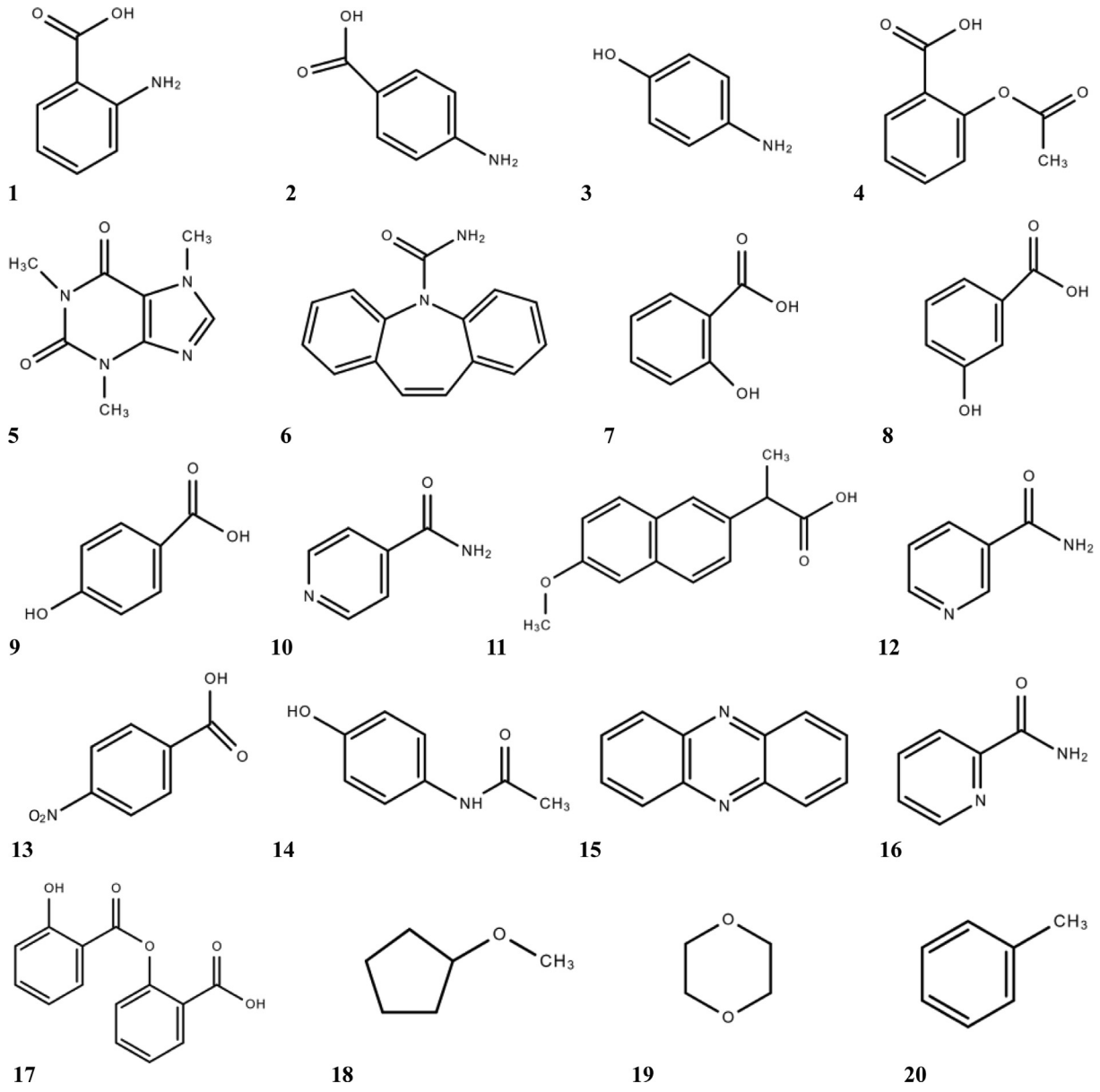

Fig. 1. Molecular structure of the experimental compounds and solvents used in this study: 1) 2ABA, 2) 4ABA, 3) 4AMP, 4) ASP, 5) CAF, 6) CBZ, 7)2HBA, 8) 3HBA, 9) 4HBA, 10) INA, 11) NPX, 12) NA, 13) 4NBA, 14) PCT, 15) PHZ, 16) PIC, 17) SSA, 18) CPME, 19) dioxane, and 20) toluene.

were weighed using an analytical balance (Mettler Toledo AG204) with a precision of ± 0.1 mg. A coated polytetrafluoroethylene (PTFE) magnetic stirrer was added to the vials before sealing, and the vials were loaded into the Crystal16 equipment (Technobis, the Netherlands). The Crystal16 allows for solubility determination of 16 samples in parallel automatically by coupling temperature control and light transmission measurements. The solute excess samples were magnetically stirred at 700 rpm while being subjected to a minimum of 3 consecutive cycles of heating and cooling, separated by an isothermal step at the lowest and highest temperature for 1 h. The heating rate was set at 0.3 °C /min, with that rate being previously reported to be sufficiently slow to accurately determine the temperature at which the suspension becomes a clear solution [6]. Cooling was carried out at a maximum rate of 1 °C /min. The clear point temperature for a given sample composition is obtained from the light transmission through the sample during the heating part of the temperature cycle and is the temperature at which the light transmission reaches 100 %. Fig. 2 presents a typical measurement outcome of the temperature and transmission as a function of time for a single experiment.

The average of three clear points temperatures measured in consecutive temperature cycles was taken as the saturation temperature of a solution composition if the difference between the clear points was lower than 3 °C. We assume that the heating rate of 0.3 °C per minute is sufficiently slow so that the saturation temperature of the sample composition can be approximated by the average of the determined clear point temperatures for a sample [6]. For each binary system, a minimum of 4 saturation temperatures were measured within the temperature range. According to Black et al.,[9] this number of data points exceed the minimum number to evaluate solubility models and design cooling crystallization process.

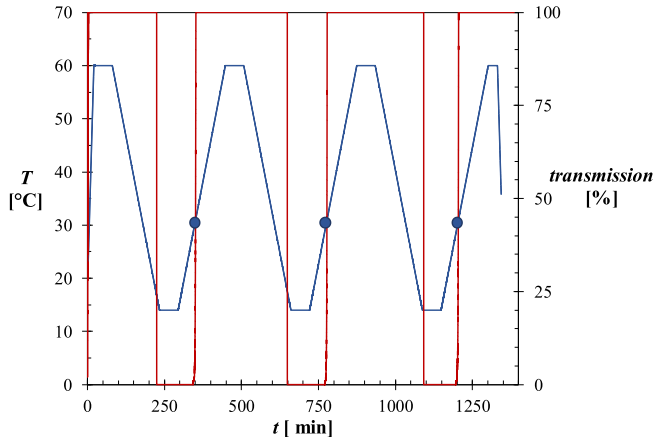

Fig. 2. Temperature cycle for picolinamide in dioxane as a function of the temperature. The blue line represents the temperature profile and the red line is the measured light transmission over the experimental period. During heating, the clear point is taken at the point where the transmission line reaches 100 %. The blue circles indicate the temperature point where the transmission reaches 100 %. Assuming that the heating rate is slow enough, the average clear point temperature is taken as the saturation temperature T_S. For the composition of picolinamide in dioxane represented in the graph, T_S = 31.1 ± 0.2 °C.

2.4. X-ray powder diffraction (XRPD)

To investigate the crystal structure of the different organic compounds in the selected isolating solvents, suspensions of approximately 5 mL of these compounds in the pure solvent were prepared and kept at room temperature under gentle stirring for a minimum of 5 days. Following equilibration for 5 days, the samples were filtered under reduced pressure over a Buchner funnel. The recovered solids and the purchased compound were separately loaded on a multi-well plate supported on a polyamide film and the crystalline phase was identified using a Bruker D8 Discover diffractometer (GX002103 - Priscilla) equipped with a Cu anode (Cu Kα1 λ = 1.54056 Å), monochromator, and a Vantec PSD detector. The XRPD spectra were recorded over a 2θ range of 4 to 35° at a scan rate of 0.015° per second.

3. Thermodynamic models

For modelling purposes, the measured solubility in mg/mL units was transformed into mole fraction solubility (x_2) using the following expression:

$$x_2 = \frac{\frac{m_2}{M_2}}{\frac{m_1}{M_1} + \frac{m_2}{M_2}} \tag{1}$$

where m_i represents the weight, M_i stands for the molecular weight, and the subscripts 1 and 2 refers to solvent and solute, respectively.

Solubility behaviour can be accurately described by thermodynamic equations, allowing for the interpolation and extrapolation of miscibility data at any given temperature. For this, the following algebraic expression enables solubility prediction for ideal binary systems based on parameters accessible from thermal analysis [10]:

$$\ln\left(x_2^{id}\right) = \frac{-\Delta H_f}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right) - \frac{1}{RT}\int_{T_m}^{T} \Delta C_p dT + \frac{1}{R}\int_{T_m}^{T} \frac{\Delta C_p}{T} dT \tag{2}$$

here, x_2^{id} stands for the ideal mole fraction solubility of the solute, ΔH_f refers to the enthalpy of fusion, R is the gas constant, and T_m is the melting temperature. The significance of the molar heat capacity ΔC_p is usually neglected compared to that of the enthalpy of fusion,[11,12] and thus equation (2) can be reduced to the following expression with minor loss of accuracy on solubility prediction for ideal systems:

$$\ln\left(x_2^{id}\right) = \frac{-\Delta H_f}{R}\left(\frac{1}{T} - \frac{1}{T_m}\right) \tag{3}$$

Ideal behaviour is associated to equal solute-solute, solvent-solvent, and solvent-solute interaction energies in the solution, and therefore a dissolution process of an ideal system proceeds without net energy change [13]. However, most systems present significant differences of their molecular interaction energies, which leads to solubility behaviour deviating from an ideal one [4]. To account for these deviations from ideality and provide a closer description of the solid-liquid equilibrium, the van't Hoff equation replaces the heat of fusion ΔH_f by the enthalpy of dissolution ΔH_{dis} and the melting temperature T_m by a temperature T_0, temperature related to the triple point temperature [13].

$$\ln(x_2) = \frac{-\Delta H_{dis}}{R}\left(\frac{1}{T} - \frac{1}{T_0}\right) \tag{4}$$

By plotting the natural logarithm $\ln(x_2)$ of the mole fraction solubility as a function of the inverse temperature, the two parameters of the van't Hoff equation can be determined from the slope and intercept of the fitted linear line to the experimental data.

Solubility data of non-ideal systems have also been extensively correlated by the use of the Apelblat equation [14]. Apelblat proposed a three parameter semi-empirical equation for the correlation of solubility data, which reported excellent fitting results for binary systems [12,15,16]:

$$\ln(x_2) = a_A + \frac{b_A}{T} + c_A \ln(T) \tag{5}$$

The last term with the parameter c_A accounts for the differences in heat capacity of the solid phase and the liquefied solute at the solution temperature. Compared to equation (4), Apelblat's equation contains an additional temperature dependent term associated to an additional fitting parameter.

In addition to the van't Hoff equation and the Apelblat model, Buchowski et al. [17] proposed another two parameter model, known as the λh model, to describe mole fraction solubilities at any given temperature:

$$\ln\left(1 + \frac{\lambda(1 - x_2)}{x_2}\right) = \lambda h\left(\frac{1}{T} - \frac{1}{T_m}\right) \tag{6}$$

Instead of describing solubility, Buchowski studied the activities of the solvent and solute as a function of temperature in order to account for the deviations from ideal solution due to associated systems. The first parameter (λ) refers to the mean association number of the solute molecules. Buchowski stated that deviations from λ = 1 reflects non-ideal behaviours due to self-association. The second factor (h) relates to the enthalpy of dissolution, and therefore the λh model accounts for the different molecular interaction energies associated to the mixing of solvent and solute.

The parameters of the Apelblat and the λh semi-empirical models can be regressed from experimental data. In this study, a non-linear least-squared algorithm implemented in Matlab 2018a was used to regress model parameters.

Solubility models based on the excess energy of the mixture offer a more rigorous thermodynamic approach for data correlation. A general expression to represent the Gibbs free energy excess of mixing in binary mixtures is given by [7]:

$$G^E = RT(n_1 \ln(\gamma_1) + n_2 \ln(\gamma_2)) \tag{7}$$

where G^E is the total excess Gibbs energy, n_i stands for the number of moles, γ_i is the activity coefficient, and the subscripts 1 and 2 refer to solvent and solute, respectively. The activity coefficient describes the extent to which a system deviates from ideal thermodynamic behaviour. It can be estimated from experimentally determined solubility data x_i and the ideal solubility x_i^{id}:

$$x_i^{id} = x_i\gamma_i \tag{8}$$

Ideal behaviour occurs when $\gamma = 1$ ($x_i^{id} = x_i$) and is well-described by equation (3). An activity coefficient lower than 1 ($\gamma < 1$) indicates a negative deviation from ideal behaviour, and positive deviations occur when $\gamma > 1$.

The derivative of equation (7) at constant pressure and temperature relates the excess Gibbs energy G^E to the activity coefficient of the solute ($i = 2$) and solvent ($i = 1$):[7].

$$\left(\frac{dG^E}{n_i}\right)_{P,T,n_j} = RT\ln(\gamma_i) \tag{9}$$

This expression allowed for the development of activity coefficient models accounting for the effect of the solute concentration on the activity coefficient. A simple relation between the mole fraction solubility and the activity coefficient was given by Margules [7]:

$$\ln\gamma_i = \frac{a_M}{RT}(1 - x_i)^2 \tag{10}$$

where a_M is a constant which can be regressed from the experimental data, and the suffix refers to the solvent ($i = 1$) or solute ($i = 2$). The one-parameter equation implies a symmetrical behaviour of the activity coefficients γ_1 and γ_2 for solvent and solute, which is a substantial assumption for a binary system of a strongly differing pharmaceutical compound and solvent. Van Laar's model introduces a second parameter to compensate these limitations and describe asymmetric systems. According to van Laar's model, the activity coefficients of solvent and solute are given by:

$$\ln\gamma_1 = \frac{a_L}{RT\left(1 + \frac{a_L x_1}{b_L x_2}\right)^2} \tag{11}$$

$$\ln\gamma_2 = \frac{b_L}{RT\left(1 + \frac{b_L x_2}{a_L x_1}\right)^2} \tag{12}$$

Despite being widely used due to their simplicity, solubility data correlation of strong non-ideal systems by the Margules and Van Laar's models are imprecise, and other Gibbs energy-based models are used to obtain more accurate representations of the solute-solvent systems. Wilson [18] proposed a two-parameter equation from a molecular interpretation of the non-ideal effect contribution of the difference of the local composition around a molecule from the overall composition, illustrated in Fig. 3 for a binary system of species i and j. According to this model, the activity coefficients of the solvent and solute are respectively given by [7]:

$$\ln(\gamma_1) = x_2\left(\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21}x_1}\right) - \ln(x_1 + \Lambda_{12}x_2) \tag{13}$$

$$\ln(\gamma_2) = -x_1\left(\frac{\Lambda_{12}}{x_1 + \Lambda_{12}x_2} - \frac{\Lambda_{21}}{x_2 + \Lambda_{21}x_1}\right) - \ln(x_2 + \Lambda_{21}x_1) \tag{14}$$

where:

$$\Lambda_{ij} = \frac{v_j}{v_i}\exp\left(-\frac{\lambda_{ij} - \lambda_{ii}}{RT}\right) \tag{15}$$

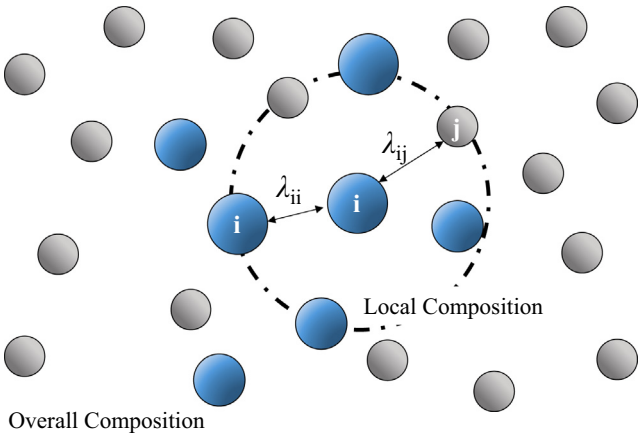

Fig. 3. Schematic diagram of the molecular local composition concept. The composition around a given molecule differs from the bulk composition due to the non-random orientation of the molecules and the dissimilar molecular interactions. The diagram shows molecules i and j of different sizes and describes their interactions from the molecule i at centre.

here, λ_{ij} refers to the average interaction energy between molecule i and j, and v_i is the molar volume of a molecule i. Equation (15) shows that this model considers the dissimilar interaction energies λ_{11}, λ_{22} and $\lambda_{12} = \lambda_{21}$ respectively among solvent molecules, among solute molecules and between solute and solvent molecules, and differentiates between the size of each species in the binary solution by introducing the molecular volume of the solvent (v_1) and the solute (v_2). The cross-interaction energies (λ_{12}-λ_{11}) and (λ_{21}-λ_{22}) are the two adjustable parameters of the Wilson equation that can be regressed from experimental data.

On the basis of the local composition concept proposed by Wilson, Renon and Prausnitz [19] derived the non-random two liquids (NRTL) equation. This model describes the excess of Gibbs energy as the sum of the different and distinct energy changes associated with the incorporation of the molecular species i and j from a pure liquid state to a mixed solution of both species. The semi-empirical equation describes the activity coefficient using two adjustable parameters (g_{12} and g_{21}) and a pre-defined constant (α) which value depend on the chemical composition of the mixture. According to the NRTL model, the activity coefficient as a function of the solubility mole fraction can be expressed as:

$$\ln(\gamma_i) = x_j^2\left[\tau_{ji}\left(\frac{G_{ji}}{x_i + G_{ji}x_j}\right)^2 + \tau_{ij}\frac{G_{ij}}{(x_j + G_{ij}x_i)^2}\right] \tag{16}$$

where:

$$\tau_{ij} = \frac{g_{ij}}{RT} \tag{17}$$

$$G_{ij} = \exp(-\alpha\tau_{ij}) \tag{18}$$

here, g_{ij} is the cross-interaction energy between the molecule j and molecule i, and α stands for the non-randomness of the distribution of the molecular species in the mixture. The non-randomness parameter is usually set between 0.2 and 0.47, however some authors assert that this parameter offers a better solubility correlation if it is used as a regression parameter as well [20].

4. Results

4.1. Solid phase characterisation

Crystalline Phase Analysis. The crystalline phase of the raw material was studied by XRPD, and the identified initial polymorphic form is listed in Table 2.

The suspension samples of the model compounds for which we report the solubility in the next section turn into a clear solution during the clear point measurement as the solids dissolve. As a transformation might occur upon introducing the initial polymorphic form into the solvent or during the clear point measurement, we have no absolute certainty about what form we measure the solubility of. Therefore, we have further performed slurry experiments at room temperature in the used solvents to identify possible transformations. The XRPD spectra of all the crystalline samples from the equilibrated systems can be consulted in the Supporting Information. Most compounds show that the initial polymorphic form does not change during the 5-day slurry experiments and for these we assume that we measure the clear point temperatures of the initial polymorphic form. The sample from the 2ABA suspension in dioxane retained the initial form I [21], with extra peaks at *ca.*11, 15, and 27° that may indicate that some metastable polymorphic form II [22,23] crystallized during the solid-liquid treatment. The compounds CBZ, 4HBA and 4NBA did show a transformation in the solvent dioxane. Fig. 4 shows the diffraction patterns of the purchased material of these four compounds compared to those after suspension equilibration in the non-polar solvent. 4NBA shows a polymorphic transformation in dioxane from form II to I and it is therefore likely we measure the solubility of form I [24] rather than form II [25]. The diffractogram of 4HBA suspended in dioxane was equivalent to a reported solvate [26], with extra peaks at approximately 18° and 25° that suggest the presence of the pure crystalline phase of 4HBA [27]. The diffraction spectra from the equilibrated CBZ crystals in the non-polar solvent differed from the polymorphic forms listed in the CCDC database. A solvate between CBZ and 1,4-dioxane has been previously reported and could explain the observed difference [28,29]. The solubility values measured are still valid for the solvates, be it that they are now solvent-free solubilities in units of mass solvent free compound per unit of solvent volume.

Thermal Analysis. The determined enthalpy of fusion (ΔH_f) and melting temperature (T_m) of the model compounds are compared to the literature data in Table 2. The experimental melting temperatures T_m of all compounds are similar to the reported melting temperature, with only a substantial temperature difference for SSA. The difference of the ΔH_f with the literature data was>5 % for 2ABA, 4ABA, 2HBA, 4AMP, ASP, CAF, INA, PCT, and SSA. An analysis repetition of a series of compounds, including 2ABA, 4ABA, 4HBA, INA, and PCT confirmed the repeatability of the thermal analysis method (Table 2). Presence of impurities or imperfections in purchased material, sample-to-sample variation and measurement errors may contribute to the data deviations between our data and that from literature.

The DSC plots, including the ΔH_f, the peak temperature, and the onset temperature for the model compounds investigated in this study are given in the Supporting Information. Most compounds have a single thermal event, indicating the melting temperature. A small thermal event preceding the major one was detected during the DSC measurement for 2ABA, CAF, CBZ, INA, and PIC, indicated with an asterisk in Table 2, which shows for these compounds that a polymorphic transformation before the melting takes place. The melting temperature and heat of fusion of the melting polymorphic form are determined using the DSC and shown in Table 2.

4.2. Solubility

Fig. 5 shows the temperature dependent solubility in mg/ml of the compounds in 1,4-dioxane. The solubility magnitude of the solutes at 45 °C decreases in the following order: SSA > 2HBA > 2ABA > PIC > ASP > S-NPX > 3HBA > RS-NPX > 4HBA > 4ABA > 4NBA > PHZ > CBZ > NA > 4AMP > INA > CAF. The solubility of PCT was only high enough above 60 °C to determine. According to the results from the crystalline phase analysis, it is likely that we measured the solubility of the forms resulting from the transition rather than the form in the input material for 2ABA, CBZ, 4HBA and 4NBA in 1,4-dioxane. Since the thermodynamic properties of the input material and the solids in solution may differ, these systems were not considered for further analysis. We estimate the error in the concentration is not larger than the symbol in Fig. 5 while the three clear point temperatures of each sam-

Table 2
Enthalpy of fusion (ΔH_f) and melting temperature (T_m) obtained by DSC analysis for 18 APIs and intermediates compounds used in this work. Where there are 2 values in a cell, this indicates the values of two independent measurements. The thermodynamic values of the organic compounds reported in the literature are tabulated alongside the measured data. The polymorphic form information was obtained by XRPD analysis carried out on the purchased material (provided in the Support Information). The * symbol indicates that two thermal events were observed during DSC heating for the particular compound.

Compound	Initial Polymorphic Form	ΔH_f [kJ/mol]	T_m [°C]	Literature		
				ΔH_f [kJ/mol]	T_m [°C]	Ref.
2ABA*	I	23.5; 23.5	145.4; 145.5	20.7	145.0	[30]
4ABA	I	26.3; 26.5	187.8; 187.8	23.2; 25.5	187.2; 187.6	[31]
4AMP		32.6	189.4	26.0	189.4	[32]
ASP	I	31.8	139.1	29.8	140.9	[33]
CAF*	β	20.8	235.8	21.9; 19.9	237.1; 236.4	[34,35]
CBZ*	III	26.7	189.9	25.6	192.1	[36]
2HBA		25.4	158.7	27.1; 24.6	161.0; 158.9	[32,36]
3HBA	I	36.9	201.2	36.5	203.3	[37]
4HBA		33.7; 33.9	214.9; 215.1	31.0; 30.9	214.9; 215.0	[37,38]
INA*	I	24.7; 25.0	156.2; 156.1	26.8	157.9	[39]
NA	I	24.7	128.3	23.8; 25.5	130.6; 128.5	[36,40]
4NBA	II	35.2	239.6	36.9	239.6	[32]
RS-NPX		35.8	156	33.2	155.8	[41]
S-NPX		34.0	155.9	31.7	156.2	[41]
PCT	I	29.3; 29.3	168.9; 168.9	27.6	170.1	[42]
PHZ		24.8	174.6	24.9	174.8	[43]
PIC*	II	19.4	106.3	19.4	106.4	[44]
SSA		32.6	143.9	29.0	157.1	[45]

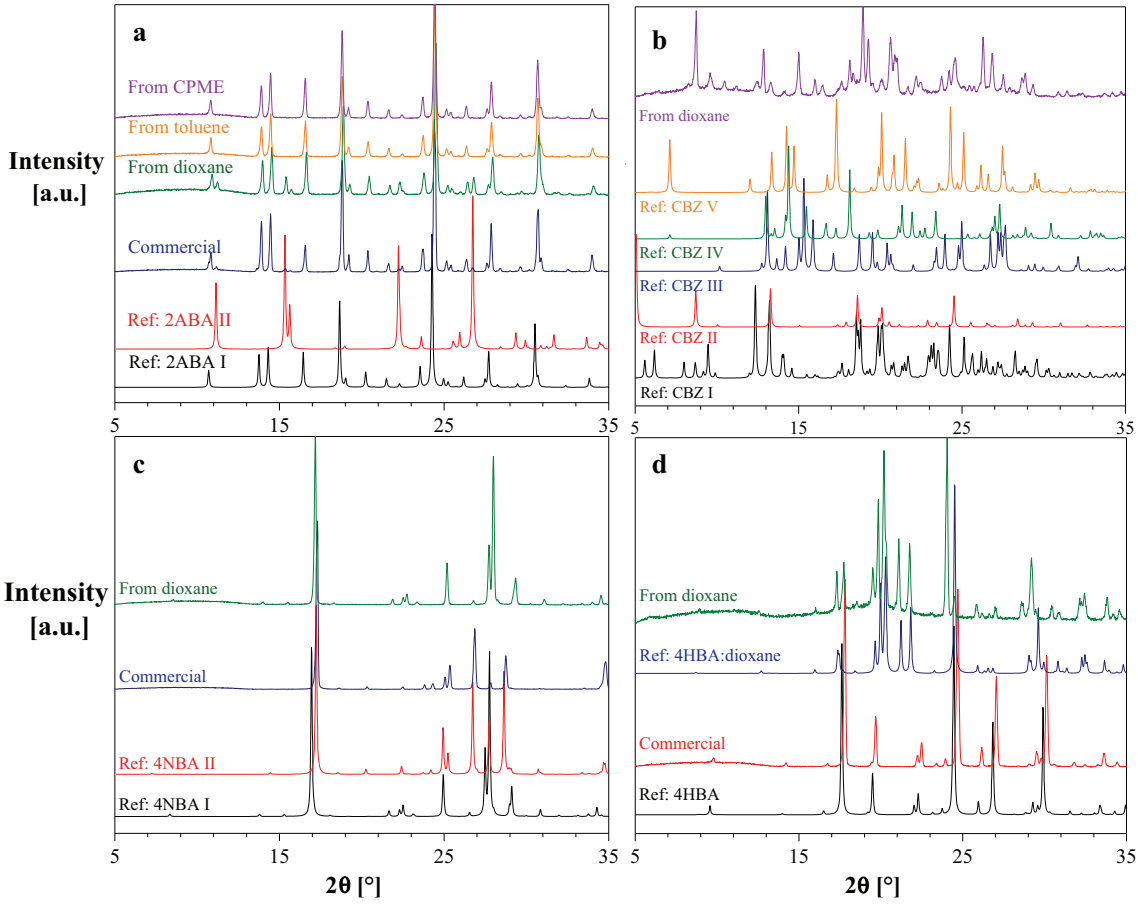

Fig. 4. a: The diffraction patterns retrieved from 2ABA after equilibration showed that the compound retained the polymorphic form I (CCDC – AMBACO01) in toluene and CPME. Additional peaks at *ca.* 7 and 15° indicates partial transformation to form II (CCDC – AMBACO03) when the compound is suspended in dioxane. **b:** The pattern of the CBZ crystals in 1,4-dioxane differs from all known CBZ forms; CBZ I (CCDC – CBMZPN11), CBZ II (CCDC – CBMZPN03), CBZ IV (CCDC – CBMZPN12), and CBZ V (CCDC - CBMZPN16). **c:** NBA in 1,4-dioxane shows a transition from the polymorphic structure II (CCDC - NBZOAC14) to form I (CCDC - NBZOAC15). **d:** 4HBA in 1,4-dioxane shows a transition from 4HBA (CCDC – JOZZIH01) to the solvate 4HBA:dioxane (CCDC - PATGEZ).

ple did not deviate >3 °C among each other, and in most cases not >1 °C.

The ortho isomers of hydoxybenzoic acid (2HBA), aminobenzoic acid (2ABA), and pyridinecarboxamide (PIC) present greater solubility values in dioxane than the analogous meta and para isomer molecules. The ortho isomer of pyridinecarboxilamide (PIC) presents the greatest temperature dependence, depicted by the larger slope of the estimated trendline to the solubility data. In the solvent CPME, the solubility order of the studied solutes follows a similar trend to that in dioxane, although, as can be seen in Fig. 5, the solubility of the compounds is lower in CPME. In toluene, the studied compounds, with the exception of CAF and PHZ, present a considerably lower solubility. For instance, at approximately 45 °C the solubility of SSA in dioxane and CPME are respectively 37 and 14 times higher than that in toluene. CAF and PHZ are only slightly less soluble in CPME.

4.3. Thermodynamic model regression

4.3.1. Semi-empirical solubility models

Table 3 presents the results from the regressions for the three empirical models from the experimental mole fraction solubility of the studied solutes across the three apolar solvents. The goodness of the fit to the experimental data is discussed later using the relative standard deviation, which are listed in Table 5.

A small difference between the enthalpy of dissolution ΔH_{dis} regressed from the van't Hoff equation and the enthalpy of fusion ΔH_f measured from the thermal analysis indicates a solution behaviour close to ideality. For instance, the binary systems of PHZ in the three non-polar solvents present similar values of ΔH_{dis} and ΔH_f, that may suggest a closer ideal behaviour than that of the other compounds in the same solvent systems.

4.3.2. Activity coefficient-based models

To evaluate the activity coefficient γ_2 of the studied organic compounds, first the ideal solubility of the solute x_2^{id} was calculated from equation (3) using the data obtained from the thermal analysis. Then, equation (8) was solved for γ_2 from the calculated x_2^{id} and the experimental data x_2. To determine the activity coefficient from the solubility data, we need access to the melting temperature and heat of fusion of the polymorphic form we measure the clear point temperature of. For most compounds these were determined using DSC, see Table 2. However, some DSC patterns (2ABA, CAF, CBZ, INA and PIC) show thermal events before melting which indicate a polymorphic transition. In absence of more accurate data we nevertheless chose to use the measured melting temperature and heat of fusion of the high melting form for the calculation of the activity coefficients.

Fig. 6 shows the temperature dependent activity coefficients γ_2 for each compound in the different organic solvents as a function of the temperature. Within the experimental temperature range, the

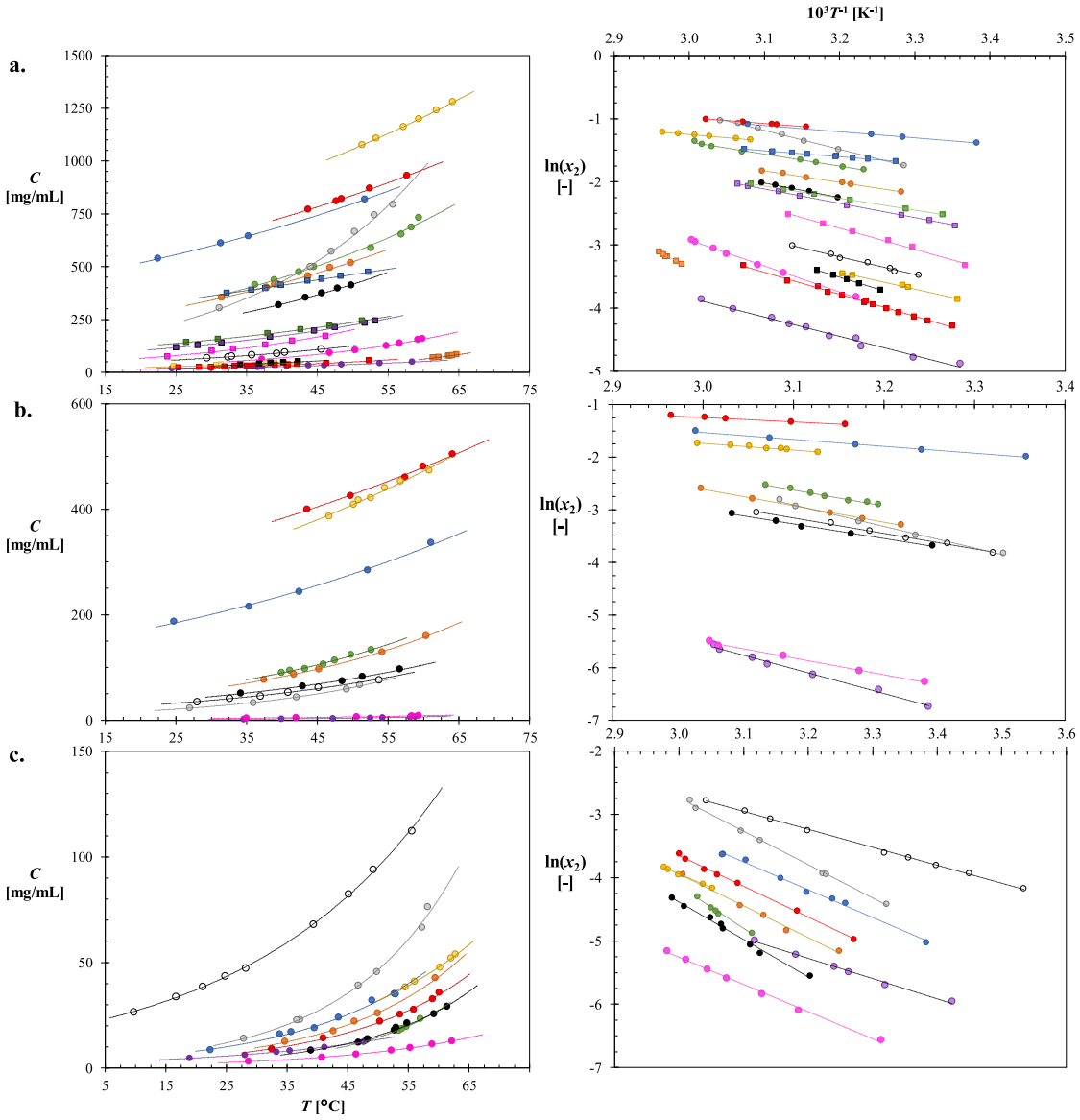

Fig. 5. Temperature dependent solubility in mg/mL (left) and the van 't Hoff plot (right) of 2ABA (), 2HBA (), ASP (), CAF (), CBZ (), RS-NPX (), S-NPX (), PHZ (), PIC (), SSA (), 3HBA (), 4ABA (),4AMP (), 4HBA (), INA (), NA (), 4NBA (), and PCT () in 1,4-dioxane (a.), CPME (b.) and toluene (c.) over a the experimental temperature range. The lines are added as a guide to the eye.

binary systems in general deviate from ideality ($\gamma_2 = 1$), with a clear tendency towards ideal solutions at increasing temperature. By defining an arbitrary boundary value, such as $-0.25 < \ln \gamma_2 < 0.25$, we can identify the close-to-ideal systems. For instance, the activity coefficients of CAF, PHZ, and PIC in 1,4-dioxane, 2ABA, ASP and NPX in CPME, and PHZ in toluene are (partially) within the defined boundaries. The activity coefficients of all the studied compounds are positive in toluene, which corresponds to a positive deviation from ideality. In the two other apolar solvents, the pyridinecarboxilamide compounds, CBZ, CAF, PHZ, and PCT show a positive deviation, while the activity coefficient of the other compounds is lower than one, showing a negative deviation.

Interestingly, the enantiopure and racemic compounds of the chiral compound NPX present dissimilar behaviour in the three solvent systems. The activity coefficients deduced from the racemic compound (RS-NPX) and S-enantiomer solubilities are similar in 1,4-dioxane over the experimental temperature range. In CPME the racemic and enantiopure solutions of NPX show close to ideal behaviour. The racemic and enantiopure solutions of NPX behave differently in toluene. NPX in the apolar solvent toluene shows positive deviations, with racemic solutions further away from ideality than the enantiopure solutions. This seems to indicate that the interactions between the two different enantiomers (R and S) differs from that between equal enantiomers (e.g., S and S) in toluene while it is close to equal in dioxane and CPME. Perhaps the self-association of NPX in toluene is promoted due to the solvent apolarity, while in dioxane and CPME, NPX is less prone to self-association as it can hydrogen bond with the solvent. Then, in cases such as in toluene, racemic solution of RS-NPX should be consider as three-component solutions to properly model the activity coefficient of such systems.

Similar activity coefficients were calculated for PHZ across the three solvent systems over the experimental temperature range. The magnitude of γ_2 for the binary systems of PHZ, are positive and showed a trend towards slightly larger values at increasing temperature. These may relate to measurement errors either in the solubility or in the thermal analysis. Since PHZ systems show close-to-ideal behaviours having activity coefficients (γ_2) close to

Table 3
Regressed parameters of the van't Hoff (ΔH_{dis} and T_0), Apelblat (a_A, b_A, and c_A), and the λh model (λ and h) obtained from the correlation to the experimental solubility data.

Solvent	Solute	van't Hoff		Apelblat			λh	
		ΔH_{dis} [kJ/mol]	T_0 [°C]	$10^{-1}a_A$ [–]	$10^{-3}b_A$ [K^{-1}]	c_A [–]	λ [–]	$10^{-3}h$ [K^{-1}]
dioxane	2HBA	7.80	240	−0.05	−0.82	0.35	0.11	2.24
	ASP	16.0	164	−30.7	12.8	45.9	0.54	3.04
	CAF	24.0	327	−16.6	4.89	25.4	0.31	8.50
	RS-NPX	18.9	179	−0.02	−2.01	0.77	0.78	3.01
	S-NPX	15.1	205	−1.50	−0.94	2.78	0.44	3.53
	PHZ	23.5	209	−9.76	1.93	15.4	0.58	4.78
	PIC	25.0	96.6	−4.77	−0.36	8.25	1.64	2.03
	SSA	8.82	276	−6.55	2.22	9.92	0.01	4.10
	4ABA	19.0	187	−0.70	−1.73	1.77	1.02	2.25
	4AMP	22.2	249	−0.51	−2.18	1.50	0.73	4.11
	3HBA	7.86	390	4.31	−2.89	−6.17	0.17	2.79
	INA	28.9	200	−20.8	6.49	31.9	0.55	6.68
	NA	30.7	172	−21.2	6.50	32.7	0.65	6.73
	PCT	53.8	130	0.05	−5.60	2.23	2.91	2.05
CPME	2ABA	11.0	269	−13.8	5.24	20.8	0.13	4.81
	2HBA	7.27	361	−13.2	5.54	19.7	0.002	3.57
	ASP	24.8	177	−39.7	16.1	59.7	0.43	6.47
	CAF	38.8	279	−0.01	−4.04	1.15	0.76	6.65
	CBZ	26.4	516	−0.05	−3.23	0.81	0.13	31.1
	RS-NPX	23.0	248	−48.0	20.0	71.8	0.41	7.66
	S-NPX	26.2	187	−25.1	9.07	38.0	0.54	5.65
	PHZ	24.4	222	6.92	−5.87	−9.37	0.45	6.19
	PIC	33.9	146	−51.3	20.2	77.5	0.27	13.7
	SSA	10.8	330	35.2	−18.1	−51.5	0.01	7.58
toluene	2ABA	36.5	172	15.3	−10.9	−21.3	0.68	6.99
	2HBA	40.3	172	−32.9	11.3	50.2	0.77	6.42
	ASP	55.5	147	−0.17	−5.96	2.66	0.89	7.79
	CAF	26.5	373	−75.6	31.7	113	0.40	8.66
	CBZ	35.6	291	0.06	−4.45	1.29	0.37	13.5
	RS-NPX	49.1	170	28.3	−18.7	−39.7	0.79	7.81
	S-NPX	42.0	177	−40.7	14.9	61.7	0.66	7.95
	PHZ	23.6	213	−4.07	−0.75	6.94	0.58	4.83
	PIC	43.9	130	−31.9	10.29	49.1	0.39	13.2
	SSA	36.2	204	−186	86.8	275	0.76	7.97

1, less accuracy on the regression of the activity coefficient is expected as experimental deviations in solubility and thermal analysis have a relatively large influence on the accuracy compared to more strongly non-ideal activity coefficients.

Following determination of γ_2, a series of thermodynamic models based on excess Gibbs energy of solution (Margules, Van Laar, Wilson, and NRTL) were used to correlate the activity coefficients. Table 4 displays the parameters for these models regressed using non-linear least squared methods using the activity coefficients determined from the experimental data. The regression of the NRTL parameters was conducted using two different approaches. On the basis of Renon's work [19], who assigned an α value ranging from 0.20 to 0.47 according to the chemical nature of the components binary systems. The non-randomness constant was set as 0.3, which is the value suggested by Renon and Prausnitz for mixtures of non-polar liquids and mixtures of apolar and polar species. Additionally, the activity coefficient was correlated from the NRTL model using α as an adjustable parameter. Thus, we can compare the best fitting results between the two-parameters and the three-parameters NRTL models.

4.4. Accuracy

The goodness of fit of the analysed models was determined by the relative standard deviation of the fitted data (x_i^{cal}) to the experimental mole fraction solubility (x_i) at a given temperature. The relative standard deviation (s) was calculated using the following expression:

$$s[\%] = \sqrt{\frac{1}{(N-1)}\sum_{i=1}^{N}\left(\frac{x_i - x_i^{cal}}{x_i}\right)^2} \cdot 100 \quad (19)$$

where N is the number of experimental data points. The relative standard deviation gives the average percentile deviation between a measurement solubility x_i and the model estimated solubility x_i^{cal}. Understanding that, and considering the general low solubility values of the studied compounds in CPME and toluene, a standard deviation lower than 5 % would be a reasonable upper error to describe the binary systems. If a model gives an error larger 5 %, we considered that the model does not describe the experimental data with sufficient accuracy. The relative standard deviations for each system are shown in Table 5. In general, there is a good agreement between the correlated and experimental solubilities obtained from each thermodynamic model across the three solvent systems, with poorer correlations in CPME. The calculated standard deviation was lower than 5 % for all the studied binary systems using the van't Hoff equation, which alongside the Apelblat and the three parameters NRTL models presented the best overall correlations.

The Apelblat model showed the lowest average relative standard deviation s over all compounds in dioxane and CPME, with values of 0.97 % and 1.62 %, respectively. However, the calculated overall s value in toluene for the semi-empirical model was larger than the s magnitudes obtained from the binary systems in toluene using van't Hoff or the three parameters NRTL models. The lowest calculated average s for the toluene systems (1.75 %) was given by the three parameters NRTL model. The use of an additional adjus-

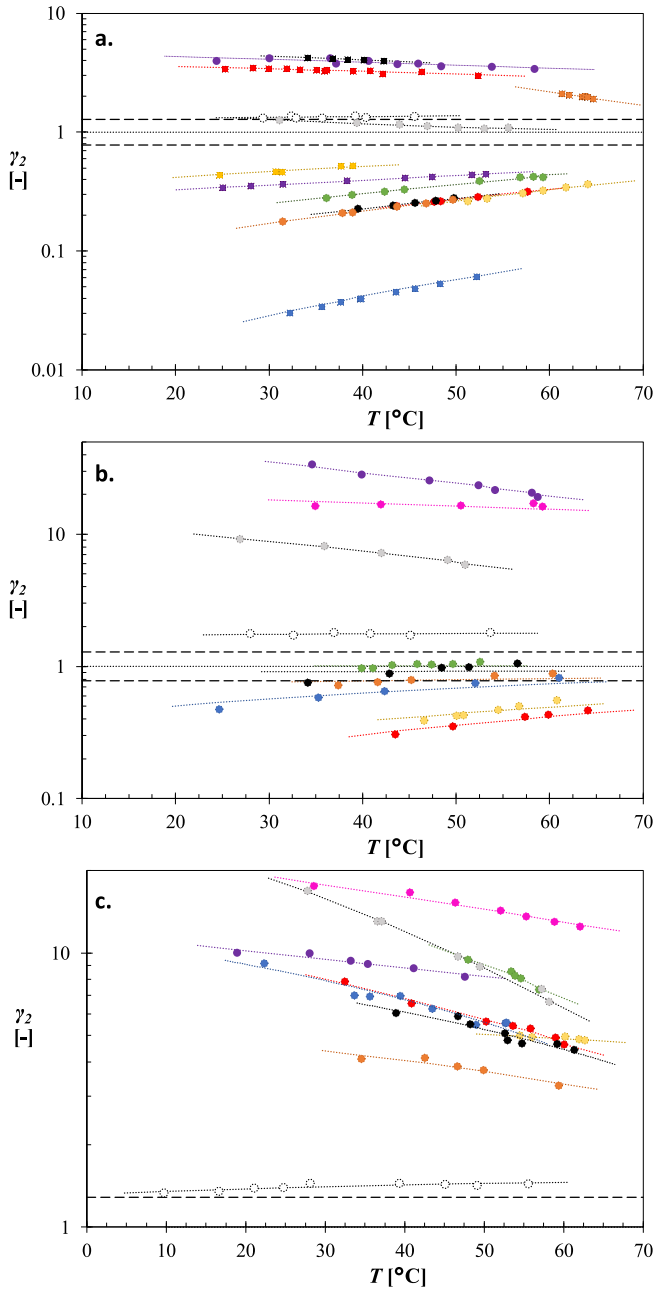

Fig. 6. Activity coefficient as a function of the absolute temperature of 2ABA (●), 2HBA (●), ASP (●), CAF (●), CBZ (●), RS-NPX (●), S-NPX (●), PHZ (○), PIC (●), SSA (●), 3HBA (■), 4ABA (■),4AMP (■), INA (■), NA (■), and PCT (■) in 1,4-dioxane (a.), CPME (b.) and toluene (c.) over the experimental temperature range. The lines for each solute were drawn using the parameters from the regression with the Wilson model. The broken lines represent an arbitrary boundary values used to visually identify the close-to-ideal binary systems.

table constant (α) considerably reduced the overall *s* for all the studied systems compared to that of the two parameters NRTL model. The Margules model showed poorer fitted solubility data across the three non-polar solvent systems, with an average deviation from the experimental data over 5 % for the binary systems in CPME and toluene. A maximum error of *ca.* 16.5 % was obtained from the Margules model for PIC in toluene.

5. Discussion

The analysis of the temperature dependent solubility of a number of APIs and intermediates in three non-polar solvents showed that the compounds are significantly more soluble in 1,4-dioxane and present lower solubilities in toluene. Zou et al. [46] state that molecular interactions between a solute and a non-polar solvent takes places through van der Waals forces. However, the ability of 1,4-dioxane to form hydrogen bonding has been confirmed by dielectric spectroscopy methods [47,48]. Indeed, the two oxygen atoms of the ether groups in the heterocyclic diethyl ether molecule of dioxane offers hydrogen acceptor sites that may promote hydrogen bond interactions with the organic model compounds which increase the solubility. The lack of hydrogen donor or acceptor sites of toluene may explain the low solubility values measured in this solvent of the compounds with hydrogen bonding capacity. The steric hindrance around the ether group of CPME (see Figure 3.1 for the molecular structure), probably contributes to, on average, a lower solubility compared to that in 1,4-dioxane.

The measured temperature dependent solubility was correlated by a series of thermodynamic models. Overall, the van't Hoff and the three parameters NRTL models provide the best fitting to the experimental data. The former model permits to compare the enthalpy and set point temperature estimated from the van't Hoff plots against the enthalpy of fusion and melting temperature. We observed a clear linear dependency of ln x_2 with T^{-1} in the relatively small temperature region for most of the studied binary systems. The fitting parameters show that this does not imply an ideal thermodynamic behaviour as the slope and intercept of the fitted line differ substantially from $-\Delta H_f/R$ and $-\Delta H_f/RT_m$ [4] for many of the systems used.

By plotting the ratio of the van't Hoff parameters to the enthalpy of fusion and melting temperature, we can identify the weight of these thermodynamic parameters on the deviations from ideal solution, i.e, the distance of the pair ($\Delta H_{dis}/\Delta H_f$, T_0/T_m) for a given solution from the origin (1, 1) in Fig. 7 can be seen as a measure of the non-ideality of a binary system. In addition, Fig. 7 permits to identify positive and negative deviations from ideal solutions. If $T_0 > T_m$ and $\Delta H_{dis} > \Delta H_f$ we have positive deviations as γ_2 is>1, that is in the upper-right quadrants of Fig. 7. Many systems, for instance all systems measured in toluene are located in that quadrant. Interestingly, in our series there are no systems that are positioned in the bottom-left quadrant of Fig. 7. That means that all negatively deviating systems are positioned in the lower-right or upper-left quadrant where the difference in heat of fusion and dissolution is countered by the difference in melting temperature and T_0. The lower-right quadrant of Fig. 7 contains more systems than the upper-left one, but in both quadrants the non-ideality due to the heat of dissolution being smaller than the heat of fusion is balanced by the temperature T_0 being larger than the melting temperature T_m. Whether this results in a positive or negative deviation from ideality depends also on the temperature *T*. It is therefore possible to have a system in these quadrants that has an apparent ideal behaviour, which seems to be the case for both RS-NPX and S-NPX in the solvent CPME.

A more detailed information of the solute-solvent interactions can be deduced from the activity coefficient of the solute γ_2, determined from the solubility and solid state properties. The ortho isomers 2HBA and PIC show greater activity coefficients in dioxane than their analogous meta and para isomer molecules, indicating a solution behaviour closer to ideality which may explain their higher solubilities in the apolar solvent. In the same context, CAF and PHZ presented similar activity coefficients in all the tested solvents, suggesting similar solubility values across the three apolar solvents.

The activity coefficient was also correlated by thermodynamic models based on the excess Gibbs free energy of mixing, for which the determined parameters describe the interaction of a molecule with its surroundings. As an example, the Wilson model parameters (λ_{ji}) refer to the interaction energies between a species *i* sur-

Table 4
Regressed parameters of the Margules (a_M), van Laar (a_L and b_L), and Wilson (λ_{12}- λ_{22} and λ_{21}- λ_{11}) models alongside the 2 parameters (g_{12}-g_{22}, and g_{21}-g_{11}) and the three parameters (α, g_{12}-g_{22}, and g_{21}-g_{11}) NRTL models for the studied solute-solvent systems.

Solvent	Solute	Margules	van Laar		Wilson		NRTL (α = 0.3)		NRTL (unrestricted α)		
		a_M [kJ/mol]	a_L [kJ/mol]	b_L [kJ/mol]	λ_{12}- λ_{11} [kJ/mol]	λ_{21}- λ_{22} [kJ/mol]	g_{12} [kJ/mol]	g_{21} [kJ/mol]	α [-]	g_{12} [kJ/mol]	g_{21} [kJ/mol]
dioxane	2HBA	−8.03	−7.27	−11.7	−5.17	−2.35	−1.60	−5.68	0.48	−2.11	−4.93
	ASP	−4.37	−3.40	−5.69	−1.81	−1.72	0.35	−3.80	4.04	−1.65	−1.73
	CAF	3.61	2.18	3.67	1.80	1.94	12.0	−4.56	0.17	24.1	−9.81
	RS-NPX	−4.71	−2.88	−6.02	−1.34	−1.62	1.56	−4.38	0.18	3.83	−6.15
	S-NPX	−5.26	−3.10	−7.66	−1.82	−1.35	0.43	−0.04	0.88	−0.71	−2.66
	PHZ	0.83	17.5	0.77	5.66	−2.50	−2.15	5.56	0.43	−1.44	5.39
	PIC	0.69	0.51	1.81	−1.61	4.83	0.54	0.17	0.75	2.87	−0.86
	SSA	−6.18	−5.26	−8.12	−4.96	−0.36	0.42	−5.49	0.60	4.06	−7.44
	4ABA	−2.97	−2.33	−3.19	−1.30	−1.16	−0.33	−2.17	0.88	−1.07	−1.31
	4AMP	−1.96	−0.35	−2.77	−3.61	6.10	8.73	−5.46	0.13	19.2	−11.5
	3HBA	−12.8	−9.63	−17.3	−5.59	−4.23	−3.92	−6.08	0.12	1.14	−11.1
	INA	3.21	4.40	3.17	2.42	1.62	1.41	2.36	0.85	2.67	7.21
	NA	3.88	4.28	3.86	2.61	2.25	2.22	2.15	0.29	2.21	2.14
	PCT	2.11	0.40	3.89	−1.74	7.82	12.0	−4.83	0.24	14.2	−6.09
CPME	2ABA	−1.70	−1.21	−15.7	−4.14	19.8	5.93	−4.51	2.80	0.97	−1.31
	2HBA	−4.89	5.09	−92.9	−6.46	2.92	1.18	−5.02	0.82	3.11	−6.67
	ASP	0.04	0.06	0.04	−1.43	2.22	0.02	0.02	0.30	−2.78	16.1
	CAF	8.66	0.54	9.31	−1.36	11.6	10.8	−1.68	0.84	9.56	−0.13
	CBZ	7.50	4263	7.46	36.7	3.79	5.05	17.2	0.03	−16.1	33.1
	RS-NPX	−0.24	−0.07	−9.75	−1.00	1.96	−0.005	−0.23	0.02	58.0	−40.5
	S-NPX	−0.70	1.28	−0.59	−1.03	1.57	−0.14	−0.53	0.03	43.8	−31.4
	PHZ	1.56	45.3	1.47	7.77	−1.32	−1.56	6.28	0.17	1.03	76.7
	PIC	5.65	3.24	6.00	1.46	5.36	5.93	0.11	0.56	5.44	11.6
	SSA	−3.04	−1.85	−4.75	−3.27	7.10	1.32	−3.40	2.17	1.35	−2.07
toluene	2ABA	5.12	1.31	5.66	−1.18	7.24	11.5	−3.63	0.99	5.20	4.80
	2HBA	4.89	1.05	5.58	−1.73	8.07	12.8	−4.24	1.13	5.08	3.60
	ASP	5.88	0.86	6.76	−1.03	8.56	15.3	−4.83	0.13	30.7	−12.4
	CAF	5.72	17.4	5.68	3.08	3.58	3.64	2.95	0.70	4.46	3.53
	CBZ	7.23	2.36	7.32	2.22	5.27	19.6	−5.88	0.31	18.9	−5.63
	RS-NPX	4.49	0.62	5.00	−0.59	6.63	16.1	−5.63	0.28	17.1	−6.11
	S-NPX	3.64	1.00	3.85	0.37	3.72	11.8	−4.57	0.18	20.3	−8.72
	PHZ	0.91	21.3	0.85	6.68	−2.03	−2.11	6.09	0.03	−11.8	15.4
	PIC	6.77	2.29	7.57	0.48	7.56	9.77	−1.72	0.68	6.54	4.58
	SSA	4.54	489	4.38	7.60	0.62	1.24	6.84	0.40	3.36	18.9

rounding by molecules *j* (Fig. 3), leading to a difference between overall and local concentrations depending on the size of both interaction energies. Orye and Prausnitz [49] expressed that, although they do not have a rigorous definition, these parameters provide information of the molecular interactions in the solution. Fig. 8 shows that the compounds with activity coefficient values $\gamma_2 < 1$ have negative λ_{12}-λ_{11}. Compounds having positive activity coefficients are bound to at least one positive Wilson parameter, indicating that the energy interactions between the different molecular species in a binary system define the solution behaviour. For instance, for SSA the Wilson regression indicates that the solute does not experience a substantially different interaction whether it is surrounded by dioxane, toluene or itself. Therefore, the difference in behavior in dioxane and toluene is entirely due to the difference in interaction between the solvent surrounded by solute and itself. On the contrary, the Wilson regression of 2HBA suggests that the solution behaviour is predominantly influenced by the different interactions between the solute with the solvent and itself, while the effect of the solvent system seems less significant.

The temperature dependent solubility is an important property, for instance in the design of crystallization processes [1]. Unfortunately, solubility prediction methods are still highly inaccurate. On the one hand, this is due to the lack of information or inaccuracy of the solid state properties like heat of fusion and melting temperature. It is for instance still not possible to accurately predict crystal structures. On the other hand, this is due to the difficulty of predicting and measuring solute behaviour, which is quantified in the activity coefficient. The solubilities and regression parameters determined here will lead to large databases which, with database analysis techniques, can be analysed and used as a basis for solubility prediction as well as understanding of solute behaviour.

6. Conclusions

The temperature dependent solubility of 18 APIs and intermediates has been measured in three non-polar solvents, including 1,4-dioxane, toluene and cyclopentyl methyl ether over a temperature range between 0 and 70 °C. The organic compounds showed a variation of solubility from *ca.* 1 mg/mL to solubilities >1 g/mL, presenting greater solubility values in 1,4-dioxane which may be explained by the greater hydrogen bond propensity of this solvent. The experimental data was modelled by three semi-empirical equations: van't Hoff, Apelblat, and the λh model for which the relative standard deviation showed that the van't Hoff equation provides the best overall fitting. In addition, a series of thermodynamic models based on the excess Gibbs energy of solution, Margules, Van Laar, Wilson, and NRTL, were used to describe the activity coefficients and herewith the solubility. According to the relative standard deviation, the three parameters NRTL provides the best overall fitting results for the studied systems, and therefore can be used along with the van't Hoff equation to accurately describe the studied organic compounds in the non-polar solvents and hence for the design of crystallization processes. The solution behaviour of the model compounds in the apolar sol-

Table 5
Relative standard deviation in percentage units of the calculated mole fraction solubilities by the thermodynamic models (x^{cal}) with respect to the experimental data (x).

Solvent	Solute	s (%)							
		van't Hoff	Apelblat	λh	Margules	van Laar	Wilson	NRTL (α = 0.3)	NRTL (unrestricted α)
dioxane	2HBA	0.20	0.20	0.21	0.42	0.04	0.65	1.14	0.60
	ASP	1.43	1.16	1.36	5.40	2.65	2.58	2.54	2.09
	CAF	4.88	3.34	5.73	2.34	2.44	3.52	2.81	1.82
	RS-NPX	0.54	0.55	0.80	3.06	0.97	0.98	0.93	0.92
	S-NPX	0.51	0.51	0.50	6.25	1.08	1.11	2.35	1.02
	PHZ	1.60	1.59	1.60	2.35	2.10	1.52	1.54	1.54
	PIC	1.17	1.29	1.15	3.07	0.98	0.95	3.00	0.91
	SSA	0.11	0.08	0.46	6.72	5.51	2.61	4.33	0.70
	4ABA	0.38	0.38	0.38	1.80	0.62	0.60	0.56	0.54
	4AMP	1.19	1.23	2.49	5.86	2.42	2.21	2.71	1.47
	3HBA	0.32	0.31	0.37	10.4	9.17	3.93	4.84	1.27
	INA	1.85	1.86	2.75	1.82	1.79	1.74	1.75	1.65
	NA	0.27	0.23	2.40	0.29	0.26	0.26	0.25	0.25
	PCT	0.89	0.91	1.16	2.64	0.56	0.51	0.44	0.42
Average		*1.10*	*0.97*	*1.52*	*3.75*	*2.18*	*1.65*	*2.08*	*1.09*
CPME	2ABA	1.59	0.39	0.90	15.5	4.90	8.64	11.8	8.62
	2HBA	0.52	0.35	0.90	9.31	7.90	4.26	6.41	2.30
	ASP	0.92	0.73	1.01	4.08	4.06	4.02	4.08	1.53
	CAF	3.30	4.64	5.09	10.3	2.88	3.72	9.20	3.43
	CBZ	2.07	4.29	10.0	10.6	10.0	6.37	4.81	2.05
	RS-NPX	1.73	1.15	4.92	12.3	8.74	12.5	12.3	0.63
	S-NPX	0.98	0.28	0.96	7.10	8.21	6.80	7.04	0.12
	PHZ	2.01	1.91	2.16	3.84	3.02	2.02	2.04	1.94
	PIC	2.83	1.97	2.60	5.34	1.09	1.17	1.10	0.91
	SSA	0.62	0.49	0.74	8.83	7.04	6.27	7.69	4.69
Average		*1.66*	*1.62*	*2.93*	*8.72*	*5.79*	*5.58*	*6.64*	*2.62*
toluene	2ABA	3.55	3.72	6.30	8.91	3.77	3.38	3.73	3.37
	2HBA	2.15	3.15	2.52	10.3	2.02	1.73	2.18	1.73
	ASP	2.25	2.30	2.48	6.15	1.10	1.13	0.74	0.27
	CAF	2.41	5.83	5.38	2.37	2.38	2.21	2.28	2.17
	CBZ	2.80	7.08	9.57	3.12	2.29	2.22	0.93	0.88
	RS-NPX	3.73	4.32	4.15	7.31	3.03	2.95	2.75	2.75
	S-NPX	3.53	3.00	4.03	4.66	2.57	2.69	1.83	1.35
	PHZ	1.70	1.35	1.81	5.61	4.64	1.57	1.58	1.43
	PIC	2.86	2.57	2.81	16.5	3.51	2.26	3.32	2.83
	SSA	0.84	0.85	5.89	1.46	0.91	0.77	0.79	0.76
Average		*2.58*	*3.42*	*4.49*	*6.64*	*2.62*	*2.09*	*2.01*	*1.75*
Overall Average		**1.70**	**1.88**	**2.81**	**6.06**	**3.37**	**2.94**	**3.40**	**1.73**

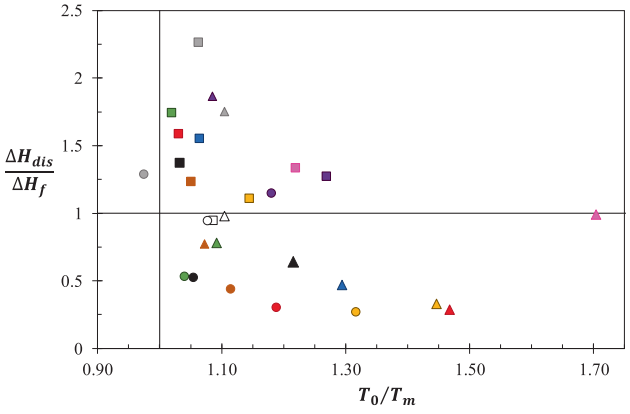

Fig. 7. Deviation of the solute-solvent systems with respect to the ideal solubility based on the ratio $\Delta H_{dis}/\Delta H_f$ of the regressed enthalpy of dissolution ΔH_{dis} and the compound's enthalpy of fusion ΔH_f and the ratio T_0/T_m of the regressed temperature T_0 and the compound's melting temperature T_m. These deviations were estimated for 2ABA (blue), 2HBA (red), ASP (green), CAF (purple), CBZ (pink), RS-NPX (black), S-NPX (orange), PHZ (white), PIC (grey), and SSA (yellow) in the three experimental apolar solvents: 1,4-dioxane (○), CPME (Δ), and toluene (□).

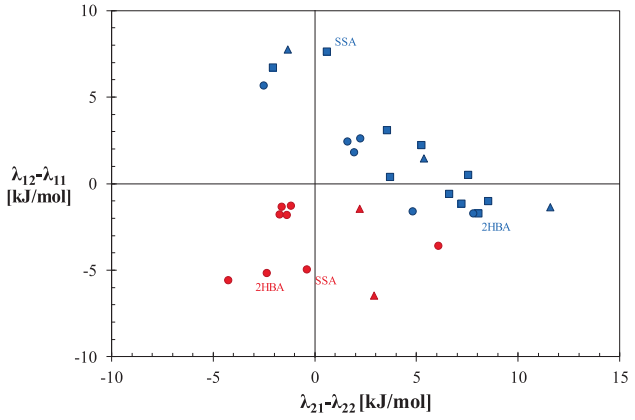

Fig. 8. Wilson interaction energy parameters λ_{12}-λ_{22} against λ_{21}-λ_{11} estimated on compounds with positive (blue) and negative (red) deviations from ideality in dioxane (○), CPME (△), and toluene (□). Only the compounds that show s values lower than 5 % for the regression with the Wilson equation (Table 5) are included, in order to have the parameters represent the binary systems with sufficient accuracy. In the figure, the different solution behaviour of 2HBA and SSA in dioxane (red) and in toluene (blue) are represented. In dioxane, 2HBA and SSA present a negative deviation from ideality, while the deviation is positive for these compounds in toluene.

vents was analyzed in terms of the activity coefficient, which permits to identify positive and negative deviations from ideality. All the studied binary systems deviate from ideality, with a closer ideal behaviour at higher temperatures. The activity coefficient of all the tested systems in toluene exhibits a positive deviation from ideal solutions, whilst the studied organic compounds in 1,4 dioxane and CPME present both positive and negative deviations from ideal solutions. An interesting behaviour was observed for the chiral compound of NPX in toluene, which suggests that solutions of the two enantiomers behave differently when compared to solutions of one enantiomer. Further analysis of the behaviour of the binary systems was carried out using the thermodynamic parameter of the van 't Hoff equation, from which the relative deviation with respect to the ideal behaviour can be visually assessed by plotting $\Delta H_{dis}/\Delta H_f$ vs T_0/T_m. The solution behaviour was further elucidated using the regressed parameters of the Wilson equation. According to the results of this analysis, the energy interactions between the different molecules in solutions are responsible of the positive or negative deviation of the binary systems from ideality.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge support from the UK Engineering and Physical Sciences Research Council (Grant Ref: EP/K503289/1) and the Doctoral Training Centre in Continuous Manufacturing and Crystallization (CMAC). We also thank Leila Keshavarz for her help with the Matlab script.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.molliq.2022.120365.

References

[1] J.H. ter Horst, C. Schmidt, J. Ulrich, Fundamentals of Industrial Crystallization, Second Edi, Elsevier B.V., 2015. https://doi.org/10.1016/B978-0-444-63303-3.00032-8.
[2] J. Ulrich, M. Jones, Industrial crystallization: Developments in research and technology, Chem. Eng. Res. Des. 82 (2004) 1567–1570, https://doi.org/10.1205/cerd.82.12.1567.58038.
[3] B.Y. Shekunov, P. York, Crystallization processes in pharmaceutical technology and drug delivery design, J. Cryst. Growth. 211 (2000) 122–136, https://doi.org/10.1016/S0022-0248(99)00819-2.
[4] J.W. Mullin, Crystallization, Fourth Edition., Butterworth-Heinemann, 2001.
[5] M.A. Reus, A.E.D.M. Van Der Heijden, J.H. ter Horst, Solubility Determination from Clear Points upon Solvent Addition, Org. Process Res. Dev. 19 (2015) 1004–1011, https://doi.org/10.1021/acs.oprd.5b00156.
[6] J.H. ter Horst, M.A. Deij, P.W. Cains, Discovering New Co-Crystals, Cryst. Growth Des. 9 (2009) 1531–1537, https://doi.org/10.1021/cg801200h.
[7] B.E. Poling, J.M. Prausnitz, J.P. O'Connell, The Properties of Gases and Liquids (Fifth Edition), McGraw-Hill, 2001. https://doi.org/10.1036/0070116822.
[8] W.W. Li, N. Radacsi, H.J.M. Kramer, E.D.M. Van Der Heijden, H. Joop, Solid Separation from a Mixed Suspension through Electric-Field-Enhanced Crystallization, Angew. Chemie Int. Ed. (2016) 1–5.
[9] S. Black, L. Dang, C. Liu, H. Wei, On the measurement of solubility, Org. Process Res. Dev. 17 (2013) 486–492, https://doi.org/10.1021/op300336n.
[10] J.M. Prausnitz, R.N. Lichtenthaler, E.G. Azevedo, Molecular Thermodynamics of Fluid Phase Equilibria, 3rd ed., Prentice-Hall Englewood Cliffs, New York, 1999.
[11] J. Wang, A. Xu, R. Xu, Solubility and solution thermodynamics of 4-hydroxybenzaldehyde in twelve organic solvents from T = (278.15 to 318.15) K, J. Mol. Liq. 237 (2017) 226–235. https://doi.org/10.1016/j.molliq.2017.04.082.
[12] Y. Wan, H. He, P. Zhang, Z. Huang, R. Zhao, J. Sha, T. Li, B. Ren, Solid-liquid equilibrium solubility and thermodynamic properties of cis-5-norbornene-endo-2,3-dicarboxylic anhydride in fourteen pure solvents and three binary solvents at various temperatures, J. Mol. Liq. 297 (2020) 111396, https://doi.org/10.1016/j.molliq.2019.111396.
[13] P. Augustijns, M. Brewster, Solvent Systems and Their Selection in Pharmaceutics and Biopharmaceutics, Springer-Verlag, New York, First, 2007.
[14] A. Apelblat, E. Manzurola, 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. 31 (1999) 85–91, https://doi.org/10.1006/jcht.1998.0424.
[15] K. Anwer, M. Muqtader, R. Ali, A. Alalaiwe, F. Shakeel, Estimating the Solubility, Solution Thermodynamics, and Molecular Interaction of Aliskiren Hemifumarate in Alkylimidazolium Based Ionic Liquids, Molecules. 1–13 (2019).
[16] W. Li, Y. Ma, Y. Yang, S. Xu, P. Shi, S. Wu, Solubility measurement, correlation and mixing thermodynamics properties of dapsone in twelve mono solvents, J. Mol. Liq. 280 (2019) 175–181, https://doi.org/10.1016/j.molliq.2019.02.023.
[17] H. Buchowski, A. Ksiazczak, S. Pietrzyk, Solvent activity along a saturation line and solubility of hydrogen-bonding solids, J. Phys. Chem. 84 (1980) 975–979, https://doi.org/10.1021/j100446a008.
[18] G.M. Wilson, X.I. Vapor-Liquid Equilibrium, A New Expression for the Excess Free Energy of Mixing, J. Am. Chem. Soc. 86 (1964) 127–130, https://doi.org/10.1021/ja01056a002.
[19] H. Renon, J.M. Prausnitz, Local Compositions in Thermodynamic Excess Functions for Liquid Mixtures, AIChE J. 14 (1968) 135–144.
[20] B. Long, J. Li, R. Zhang, L. Wan, Solubility of benzoic acid in acetone, 2-propanol, acetic acid and cyclohexane: Experimental measurement and thermodynamic modeling, Fluid Phase Equilib. 297 (2010) 113–120, https://doi.org/10.1016/j.fluid.2010.06.021.
[21] C.J. Brown, The Crystal Structure of Anthranilic Acid, Proc. R. Soc. A. (1968) 185–199.
[22] C.D.G. Boone, J.L. Derissen, J.C. Schoone, Anthranilic Acid II (o-Aminobenzoic Acid), Acta Crystallogr. Sect. B. B33 (1977) 3205–3206. https://doi.org/https://doi.org/10.1107/S0567740877010541.
[23] S. Jiang, P.J. Jansens, J.H. ter Horst, Control over polymorph formation of o-aminobenzoic acid, Cryst. Growth Des. 10 (2010) 2541–2547, https://doi.org/10.1021/cg901257s.
[24] Q. Wang, CCDC 1499185: Experimental Crystal Structure Determination, CSD Commun. (2016).
[25] M.E. Light, CCDC 1476879: Experimental Crystal Structure Determination, CSD Commun. (2016).
[26] C. Kingsbury, C.J. Commons, R.W. Sanchez-Arlt, B.F. Abrahams, R. Robson, CCDC 1557139: Experimental Crystal Structure Determination, CSD Commun. (2017).
[27] J.J. Du, S.A. Stanton, P.A. Williams, J.A. Ong, P.W. Groundwater, J. Overgaard, J.A. Platts, D.E. Hibbs, Using Electron Density to Predict Synthon Formation in a 4-Hydroxybenzoic Acid: 4,4′-Bipyridine Cocrystal, Cryst. Growth Des. 18 (2018) 1786–1798, https://doi.org/10.1021/acs.cgd.7b01676.
[28] R. Hilfiker, J. Berghausen, F. Blatter, A. Burkhard, S.M. De Paul, B. Freiermuth, A. Geoffroy, U. Hofmeier, C. Marcolli, B. Siebenhaar, M. Szelagiewicz, A. Vit, M. Von Raumer, Polymorphism - Integrated approach from high-throughput screening to crystallization optimization, J. Therm. Anal. Calorim. 73 (2003) 429–440, https://doi.org/10.1023/A:1025409608944.
[29] A.J. Florence, A. Johnston, S.L. Price, H. Nowell, A.R. Kennedy, N. Shankland, An automated parallel crystallisation search for predicted crystal structures and packing motifs of carbamazepine, J. Pharm. Sci. 95 (2006) 1918–1930, https://doi.org/10.1002/jps.20647.
[30] W.E. Acree, Thermodynamic Properties of Organic Compounds: Enthalpy of Fusion and Melting Point Temperature Compilation, Thermochim. Acta. 189 (1991) 37–56, https://doi.org/10.1016/0040-6031(91)87098-H.
[31] S. Gracin, Å.C. Rasmuson, Polymorphism and crystallization of p-aminobenzoic acid, Cryst. Growth Des. 4 (2004) 1013–1023, https://doi.org/10.1021/cg049954h.
[32] E.S. Domalski, E.D. Hearing, Heat Capacities and Entropies of Organic Compounds in the Condensed Phase Volume II, J. Phys. Chem. Ref. Data. 19 (1990) 881–1047, https://doi.org/10.1063/1.555876.
[33] D.R. Kirklin, Enthalpy of combustion of acetylsalicylic acid, J. Chem. Thermodyn. 32 (2000) 701–709, https://doi.org/10.1006/jcht.1999.0650.
[34] S.S. Pinto, H.P. Diogo, Thermochemical study of two anhydrous polymorphs of caffeine, J. Chem. Thermodyn. 38 (2006) 1515–1522, https://doi.org/10.1016/j.jct.2006.04.008.
[35] J.X. Dong, Q. Li, Z.C. Tan, Z.H. Zhang, Y. Liu, The standard molar enthalpy of formation, molar heat capacities, and thermal stability of anhydrous caffeine, J. Chem. Thermodyn. 39 (2007) 108–114, https://doi.org/10.1016/j.jct.2006.05.009.
[36] D.J. Good, R.H. Naír, Solubility advantage of pharmaceutical cocrystals, Cryst. Growth Des. 9 (2009) 2252–2264, https://doi.org/10.1021/cg801039j.
[37] S.S. Pinto, H.P. Diogo, R.C. Guedes, B.J. Costa Cabral, M.E. Minas Da Piedade, J.A. Martinho Simões, Energetics of hydroxybenzoic acids and of the corresponding carboxyphenoxyl radicals. Intramolecular hydrogen bonding in 2-hydroxybenzoic acid, J. Phys. Chem. A. 109 (2005) 9700–9708, https://doi.org/10.1021/jp054220g.
[38] N.A. Armstrong, K.C. James, C.K. Wong, Inter -relationships between solubilities, distribution coefficients and melting points of some substituted benzoic and phenylacetic acids, J. Pharm. Pharmacol. 31 (1979) 627–631.
[39] H. Negoro, T. Miki, S. Ueda, T. Sanada, R. Okada, Solubility phenomena of pyridine-and pyrazinemonocarboxamides. II. Heats of fusion of picolinamide,

nicotinamide, isonicotinamide, and pyrazinecarboxamide, Yakugaku Zasshi. 80 (1960) 670–673.
[40] S. Nicoli, S. Bilzi, P. Santi, M.R. Caira, J. Li, R. Bettini, Ethyl-paraben and nicotinamide mixtures: Apparent solubility, thermal behavior and x-ray structure of the 1:1 co-crystal, J. Pharm. Sci. 97 (2008) 4830–4839, https://doi.org/10.1002/jps.21341.
[41] D.E. Braun, M. Ardid-Candel, E. D'Oria, P.G. Karamertzanis, J.B. Arlin, A.J. Florence, A.G. Jones, S.L. Price, Racemic naproxen: A multidisciplinary structural and thermodynamic comparison with the enantiopure form, Cryst. Growth Des. 11 (2011) 5659–5669, https://doi.org/10.1021/cg201203u.
[42] F.L. Mota, A.P. Carneiro, A.J. Queimada, S.P. Pinho, E.A. Macedo, Temperature and solvent effects in the solubility of some pharmaceutical compounds: Measurements and modeling, Eur. J. Pharm. Sci. 37 (2009) 499–507, https://doi.org/10.1016/j.ejps.2009.04.009.
[43] R.D. Chirico, A.F. Kazakov, W.V. Steele, Thermodynamic properties of three-ring aza-aromatics. 1. Experimental results for phenazine and acridine, and mutual validation of experiments and computational methods, J. Chem. Thermodyn. 42 (2010) 571–580, https://doi.org/10.1016/j.jct.2009.11.010.
[44] A.O.L. Évora, R.A.E. Castro, T.M.R. Maria, M.T.S. Rosado, M.R. Silva, J. Canotilho, M.E.S. Eusébio, Resolved structures of two picolinamide polymorphs. Investigation of the dimorphic system behaviour under conditions relevant to co-crystal synthesis, CrystEngComm. 14 (2012) 8649–8657, https://doi.org/10.1039/c2ce26244d.
[45] J.J.M. Ramos, H.P. Diogo, M.H. Godinho, C. Cruz, K. Merkel, Anomalous thermal behavior of salicylsalicylic acid and evidence for a monotropic transition to a nematic phase, J. Phys. Chem. B. 108 (2004) 7955–7962, https://doi.org/10.1021/jp049903v.
[46] F. Zou, W. Zhuang, J. Wu, J. Zhou, Q. Liu, Y. Chen, J. Xie, C. Zhu, T. Guo, H. Ying, Experimental measurement and modelling of solubility of inosine-5′-monophosphate disodium in pure and mixed solvents, J. Chem. Thermodyn. 77 (2014) 14–22, https://doi.org/10.1016/j.jct.2014.04.023.
[47] R.J. Sengwa, S. Sankhla, V. Khatri, Dielectric characterization and molecular interaction behaviour in binary mixtures of amides with dimethylsulphoxide and 1,4-dioxane, J. Mol. Liq. 151 (2010) 17–22, https://doi.org/10.1016/j.molliq.2009.10.011.
[48] S. Schrödle, G. Hefter, R. Buchner, Dielectric spectroscopy of hydrogen bond dynamics and microheterogenity of water + dioxane mixtures, J. Phys. Chem. B. 111 (2007) 5946–5955, https://doi.org/10.1021/jp0713413.
[49] R.V. Orye, J.M. Prausimitz, Multicomponent equilibria: With the wilson equation, Ind. Eng. Chem. 57 (1965) 18–26, https://doi.org/10.1021/ie50665a005.