International Journal of Pharmaceutics 426 (2012) 29–43

Contents lists available at SciVerse ScienceDirect

International Journal of Pharmaceutics

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

A new expanded solubility parameter approach

Emmanuel Stefanis, Costas Panayiotou*

Department of Chemical Engineering, Aristotle University of Thessaloniki, 54024 Thessaloniki, Greece

ARTICLE INFO

Article history:
Received 8 December 2011
Received in revised form 1 January 2012
Accepted 2 January 2012
Available online 9 January 2012

Keywords:
Partial solubility parameter
Hydrogen-bonding
COSMO-RS theory
Group-contribution method

ABSTRACT

The partial or Hansen solubility parameters (HSP) are important properties of the various substances and very useful tools for the selection of their solvents or the prediction of their behaviour in numerous applications. Their design and evaluation relies on the basic rule of "similarity matching" for solubility. The present work attempts to enhance the capacity of HSPs by incorporating into their evaluation the other basic rule of solubility, namely, the rule of "complementarity matching". This is done in a simple and straightforward manner by splitting the hydrogen bonding HSP into its acidic or proton donor component and its basic or proton acceptor one. The splitting is based on the third σ-moments of the screening charge distributions or sigma profiles of the quantum-mechanics based COSMO-RS theory. The whole development and application does not involve any sophisticated calculations or any strong specific background. The new method has been applied to a variety of solubility data for systems of pharmaceutical interest in order to verify the significant improvement over the classical HSP approach. The application of the new method requires, of course, the knowledge of the HSPs. For this reason, in Appendix A is presented an updated version of a robust and reliable group-contribution method for the calculation of the HSPs. The key features of this combined tool are critically discussed.

1. Introduction

The production of chemicals, such as pharmaceuticals, cosmetics, coatings, and foodstuffs, often involves multicomponent mixtures. In this regard, there is very much interest today in the development of reliable methods for the prediction of key physicochemical properties of materials, especially for their miscibility with other substances and their interaction with their environment in order to meet process and product quality specifications. Numerous multivariate linear or non-linear quantitative structure–property relationships and related methodologies have been developed. One relatively simple and of the most widely used approaches is through the calculation of solubility parameters that reflect the various contributions to the cohesion of matter. The conceptual simplicity of the solubility parameter, δ, originally introduced by Hildebrand and Scott (1962), makes it most attractive in industry and in academia. In spite its moderate success, it remains today one of the key parameters for selecting solvents in industry, for predicting polymer compatibility, chemical resistance, and permeation rates, for characterizing surfaces, and for rationally designing new processes, such as the supercritical fluid processes, the coating, and the drug design and delivery processes (Barton, 1983, 1985; Bustamante et al., 1998; Hansen, 1967, 2004, 2007; Tehrani, 1993).

The cohesive energy density (*ced*) of a liquid is the energy of vaporization per unit volume and it reflects the strength of attractive forces holding the molecules together. Equivalently, *ced* is the ratio, E/V, of the cohesive energy, E, of the system divided by its molar volume, V. In this context, E is the increase in the internal energy per mole of the system upon removal of all intermolecular interactions. The solubility parameter, δ, is, simply, the square root of *ced* (Hildebrand and Scott, 1962). The central principle behind the use of δ is the old alchemist maxim, "similia similibus solvuntur" ("like dissolves like"), probably, the oldest rule of solubility. Of course, the use of solubility parameter is not always successful and this very lack of total success stimulates continuing research.

The above rule of solubility can, indeed, be a good guide in the selection of an appropriate solvent for a given solute, as long as we can also define with sufficient precision the degree of likeness in the solute–solvent pair. This need for precision in the definition of likeness led to the division of δ into its partial components or Hansen solubility parameters (HSP) (Hansen, 2007), δ_d, δ_p, and δ_{hb}, for the dispersion, the polar, and the hydrogen-bonding contribution, respectively. According to the "similarity matching" rule, liquids with similar δ_d, δ_p, and δ_{hb}, are very likely to be miscible. Indeed, the division of δ into its HSP components has very much improved its success in solvent selection and in related applications.

* Corresponding author. Tel.: +30 2310 996223; fax: +30 2310 996232.
E-mail address: cpanayio@auth.gr (C. Panayiotou).

0378-5173/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijpharm.2012.01.001

The "similarity matching" rule is not always successful and it is now recognized that a more appropriate principle would be the "complementarity matching" of properties (Jensen, 1987). Thus, the hydrogen bonding component, δ_{hb}, has been proposed to be further subdivided into an acidic component, δ_a, and a basic component, δ_b, in order to account for the Lewis-acid and Lewis-base character of the substance (Bustamante et al., 1998; Jensen, 1987; Karger et al., 1976). These attempts have their analogues in the widely used multivariate linear free-energy relationships (LFER) or linear solvation energy relationships (LSER), based often on purely theoretical molecular descriptors, where the hydrogen-bonding scales are also divided into acidity and basicity scales (Kamlet and Taft, 1976; Kamlet et al., 1981; Joris et al., 1972; Raevsky, 1987; Sherry and Purcell, 1972; Taft and Kamlet, 1976). The experimental information for the construction of these later scales is the enthalpy of hydrogen-bond formation (Raevsky, 1987; Sherry and Purcell, 1972) or solvatochromic studies (Abraham, 1993; Kamlet and Taft, 1976; Kamlet et al., 1981; Joris et al., 1972; Taft and Kamlet, 1976) including NMR shifts (Abraham, 1993). In the case of expanded solubility parameters, Karger et al. (Karger et al., 1976, 1978) have used the corresponding acidity and basicity scales of Kamlet et al. (Kamlet and Taft, 1976; Kamlet et al., 1981; Taft and Kamlet, 1976) and developed simple linear relations for δ_a and δ_b, respectively. The adopted key equation for their relation to the Hansen hydrogen-bonding δ_{hb} is the following

$$2\delta_a\delta_b = \delta_{hb}^2 \quad (1)$$

This equation was also adopted invariably in later studies (Barra et al., 1997; Beerbower et al., 1984; Bustamante et al., 1998; Jensen, 1987). A major drawback of equation 1 arises when one of the δ_a and δ_b is zero or very small, forcing the other to be intolerably or unacceptably high. In fact, Eq. (1) cannot apply to the very common case where a compound has only acidic or only basic character. A similar problem was faced in the construction of hydrogen bonding scales in the LSER approach (Abraham, 1993). Apart from this drawback, the expanded Hansen solubility parameter approach was substantially more successful over the plain Hansen approach (Verheyen et al., 2001). Yet, the vast majority of users prefer the plain Hansen approach because there are no extensive databases available for the separate δ_a and δ_b neither are robust and simple methods established for their unequivocal determination. On the contrary, over the years, the plain Hansen partial solubility parameters were determined for a very large number of substances and led to critical compilations available in the open literature (Barton, 1983; Hansen, 2007; Abbott and Hansen, 2010). This type of compilations is a most valuable source of information for the nature of the substances and their intermolecular interactions with other substances.

As already mentioned, the HSPs may easily be integrated in modern process simulators for the rational design of new products and processes. In this respect, knowledge of the variation of HSPs with external conditions is very much useful. In a recent work (Stefanis et al., 2006) we have adopted the NRHB (Non-Random Hydrogen-Bonding) equation-of-state framework and calculated the effect of temperature, pressure, and composition on HSPs of various fluids. Group-contribution approaches for HSPs are also useful in this respect. In another recent work (Stefanis and Panayiotou, 2008) we have presented a robust and reliable new methodology for the calculation of group contributions for HSPs based on the Conjugation theory (Constantinou et al., 1993; Mavrovouniotis, 1990). Tables with the contributions of first- and second-order groups were also presented (Stefanis and Panayiotou, 2008). In a very recent interesting work (Járvás et al., 2011), a multivariate non-linear method based on artificial neural networks has used the COSMO-RS σ-moments (Klamt, 2005) as molecular descriptors for the prediction of HSPs but their success was inferior to the above group-contribution method (Stefanis and Panayiotou, 2008). The predictive conductor-like screening model for real solvents, COSMO-RS, combines in an eloquent manner the strength of quantum chemistry with concepts of dielectric continuum models and group–surface interactions leading to a powerful tool of modern chemical thermodynamics. Thus, Ikeda et al. (2005) could satisfactorily predict with COSMO-RS the solubility of a number of drugs in polar solvents. However, the implementation and performance of this kind of relatively sophisticated calculations are not highly favored methods for the majority of pharmaceutical scientists. It is worth pointing out here that the COSMO-RS theory was also successfully transformed into an LSER model (Klamt, 2005) along the lines of Abraham's model (Abraham, 1993).

The objective of the present work is twofold: In Appendix A, we summarize the essentials of our group contribution approach (Stefanis and Panayiotou, 2008), report improved updated and extensive tables with the first- and second-order group contributions, and give representative examples of calculations of HSPs. Having the HSPs either from available compilations or from reliable calculation methods, such as the group-contribution method of the Appendix, in the main part of this work, we propose a new scheme for splitting the hydrogen bonding solubility parameter, δ_{hb}, into its acidic and basic components, δ_a and δ_b, respectively. The new splitting is based on the third moments of screening charge density profiles of the COSMO-RS theory (Klamt, 2005) for their hydrogen bonding acceptor and donor parts. This 4-parameter method is compared, subsequently, with the plain 3-parameter HSP approach against a variety of experimental solubility data for systems of pharmaceutical interest and a critical discussion follows.

2. The acidic and basic partial solubility parameters

As already mentioned in the Introduction, the aim of the present section is the split of the hydrogen bonding solubility parameter, δ_{hb}, into its acidic and basic components, δ_a and δ_b, respectively, based on the third moments of screening (polarization) charge density profiles or sigma profiles of the COSMO-RS theory (Klamt, 2005) for their hydrogen bonding acceptor and donor parts. In doing this, we will provide with a straightforward and easy to use model. In what follows we will provide with the rationale of our approach.

A most important element of the COSMO-RS model is the evaluation of the sigma (σ) profiles for the charge density distribution on the surface of molecules, which enables the calculation of molecular interactions even in complex and highly non-ideal multicomponent mixtures. These profiles are unique properties of pure compounds and can be calculated rather easily via widely available quantum chemical calculation software suits such as the Dmol^3 of Accelrys^R or the TURBOMOLE (Ahlrichs et al., 1989). The calculated σ-profiles for thousands of compounds may be found in current databases either commercial (COSMObase, 2006) or free of charge (VT Sigma Profile Databases, 2006). The distribution of molecular surface charges as provided by the σ-profiles (Fig. 1) give, in pictorial manner, valuable information on the capacity of the molecule to interact with dispersion, polar, or hydrogen bonding forces. A cutoff of ± 0.01 e/nm^2 is typically set (Klamt, 2005) beyond which the surface charge may participate in hydrogen bond formation if the complementary (opposite) charge is available in the interacting system. Thus, by looking at the σ-profile of a compound, one may evaluate the acidic and/or basic character of the compound and foresee the capacity and strength of its hydrogen bonding interactions. In Fig. 1 are shown the σ-profiles of ethanol and phenol. As observed, the acidic character of phenol is very clearly depicted by the pronounced longer tale in the left hand side of the profile (the σ-charges, as screening charges, are the opposite of the real molecular surface charges).

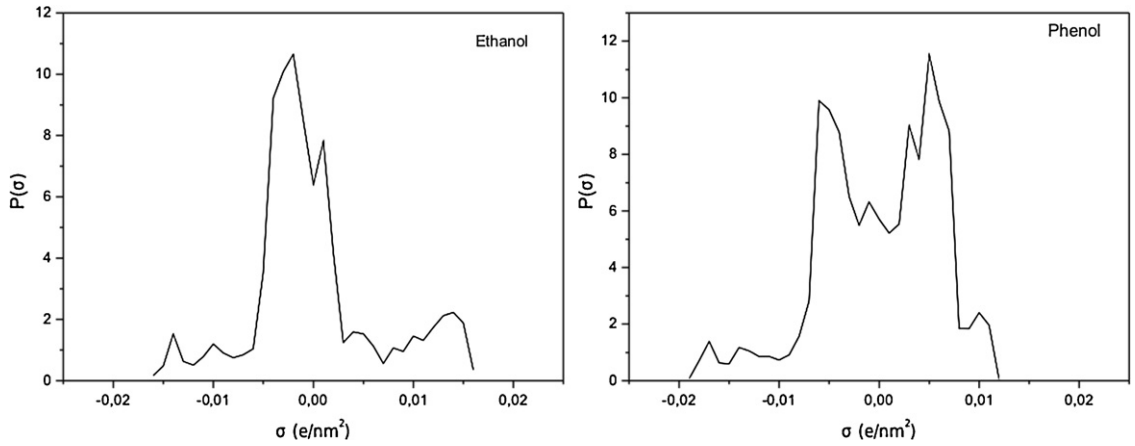

Fig. 1. The sigma profiles of ethanol and phenol clearly indicating the stronger acidic character of the latter.

From the σ-profiles one may obtain the various moments of the distribution. These σ-moments were proved very good linear descriptors in an LSER model implementation (Klamt, 2005). Of all these moments the ones that are of interest to us here are the hydrogen bonding acceptor and donor functions, *HB_acc3* and *HB_don3*, respectively (COSMObase, 2006; Járvás et al., 2011). Values of these functions for some common pharmaceuticals and solvents are reported in Table 1. The division of the hydrogen bonding solubility parameter into its acidic and basic component should be done in a way that reflects the proportion of the above donor and acceptor functions, respectively. Unfortunately, there is no absolute and universally accepted measure of this acidic or basic character, to which one had to comply, not even for a single reference substance. As an example, Beerbower et al. (1984) have adopted a scale according to which ethanol has equal acidic and basic character, as reflected by the equal values of the proposed δ_a and δ_b parameters. An inspection, however, of Fig. 1 shows that this is a rather too arbitrary assumption not corroborated by the real surface charge distribution of this compound. A proper scale would reflect the predominance of the basic character of ethanol. Even if one could question the appropriateness of the above third moments for the donor and acceptor functions to mirror δ_a and δ_b parameters, their difference (Table 1) is rather too large to justify the equality of the δ_a and δ_b parameters for ethanol.

In order to avoid the above objections regarding the acidity/basicity scale based on ethanol, we have decided to follow the common chemistry practice and adopt water as the reference "neutral" substance of equal acidity and basicity. This adoption implies, of course, a universal shift of the *HB_don3* functions in order to reflect the water neutrality. This is the basis for the rationale of our method.

As seen in Table 1, the acceptor and donor functions, *HB_acc3* and *HB_don3*, are equal to 5.758 and 3.858, respectively. They become equal by multiplying the latter by 1.492. This is the universal factor by which we must multiply the *HB_don3* values of all substances. This modified *HB_don3* value, when added to the corresponding *HB_acc3* value, give the m-SUM value for the substance and which is reported in the 4th column of Table 1. Our proposal, then, amounts to adopting the following defining equations:

$$\frac{\delta_b^2}{\delta_{hb}^2} = \frac{HB\_acc3}{m-SUM} = 1 - \frac{\delta_a^2}{\delta_{hb}^2} \tag{2}$$

In other words, our proposal implies that the splitting of the hydrogen bonding solubility parameter into its acidic and basic components obeys the following equation:

$$\delta_{hb}^2 = \delta_a^2 + \delta_b^2 \tag{3}$$

It is clear that our proposal does not have the problem mentioned in Section 1 and caused when one of the acidic or basic components is too small. In contrast to the so far adopted Eq. (1), our Eq. (3) shows that zero or negligibly small values for one of the components can be tolerated without causing the other component to adopt intolerably high values. The values of δ_a and δ_b parameters, calculated with Eqs. (2) and (3), are reported in the last two columns of Table 1.

3. Applications

In order to test the appropriateness of the estimated δ_a and δ_b parameters, we have applied the classical Hansen's radius of solubility criterion (Hansen, 2007) to a number of systems of interest to the Pharmaceutical community. Hansen's criterion for a solute 2 to be soluble in a solvent 1 is a small value (ideally equal to zero) of the following function:

$$R^2 = 4(\delta_{d1} - \delta_{d2})^2 + (\delta_{p1} - \delta_{p2})^2 + (\delta_{hb1} - \delta_{hb2})^2 \quad \text{(Hansen)} \tag{4}$$

In our case, this criterion would read:

$$R^2 = 4\left(\delta_{d1} - \delta_{d2}\right)^2 + \left(\delta_{p1} - \delta_{p2}\right)^2 + \left(\sqrt{\delta_{a1}^2 + \delta_{b1}^2} - \sqrt{\delta_{a2}^2 + \delta_{b2}^2}\right)^2 \tag{4a}$$

The criteria of Eqs. (4) or (4a) reflect the premise of "similarity" principle for solubility, as discussed previously. In the case of hydrogen bonding, however, the important feature that should be expressed by an appropriate criterion is the "complementarity" rather than the "similarity" of solute/solvent interactions. In this respect, the last term of Eq. (4) should drastically change in order to favor acid–base interactions between the solute and the solvent. There are various proposals that could be made and a most simple one is the following (Beerbower et al., 1984):

$$R^2 = 4\left(\delta_{d1} - \delta_{d2}\right)^2 + \left(\delta_{p1} - \delta_{p2}\right)^2 + 2\left(\delta_{a1} - \delta_{b1}\right)\left(\delta_{a2} - \delta_{b2}\right) \quad \text{(Extended)} \tag{5}$$

It is clear from Eq. (5) that the presence of complementary acidic and basic groups in compounds 1 and 2 does favor their miscibility as the product of the complementary components are subtracted from the rest of the terms and, thus, reduces the radius of solubility. In fact, when the solute–solvent interactions lead to the formation of strong hydrogen bonds, one may expect that this might be the most important factor and that would lead even to negative values for R^2.

Table 1
The COSMO-RS third sigma moments (COSMObase, 2006; Klamt, 2005) and the acidic and basic partial solubility parameters.

Compound	HB_acc3	HB_don3	m-SUM	δ_{hb} (Hansen, 2007; Jouyban, 2010)	δ_b	δ_a
Pharmaceuticals						
Ascorbic acid	4.390	9.255	18.203	25.50	12.52	22.21
Aspirin	1.802	4.188	8.053	9.30	4.40	8.19
Benzoic acid	1.361	4.022	7.364	10.80	4.64	9.75
Citric acid	2.622	8.968	16.007	16.72	6.77	15.29
Ibuprofen	1.340	3.942	7.223	8.89	3.83	8.02
Oxazepam	5.234	3.762	10.848	13.21	9.18	9.50
Paracetamol	4.461	5.772	13.075	13.90	8.12	11.28
Piroxicam	3.482	2.374	7.024	8.74	6.15	6.21
Salicylic acid	0.879	4.763	7.988	13.70	4.55	12.92
Stearic acid	1.901	3.687	7.403	5.50	2.79	4.74
Solvents						
Water	5.758	3.858	11.516	42.30	29.91	29.91
Methanol	4.263	1.939	7.157	22.30	17.21	14.18
Ethanol	3.983	1.540	6.281	19.40	15.45	11.74
Ethylene glycol	5.545	2.948	9.945	26.00	19.41	17.29
1-Propanol	3.848	1.578	6.203	17.40	13.70	10.72
Isopropanol	4.100	1.434	6.240	16.40	13.29	9.60
1-Butanol	3.777	1.548	6.088	15.80	12.45	9.73
2-Butanol	3.543	1.411	5.649	14.50	11.48	8.85
2-Methyl-1-propanol	3.409	1.793	6.085	15.90	11.90	10.54
2-Methyl-2-propanol	3.836	1.135	5.529	14.70	12.24	8.14
1-Pentanol	4.001	1.833	6.737	13.90	10.71	8.86
1-Hexanol	4.032	1.845	6.786	12.50	9.64	7.96
1-Heptanol	4.000	1.832	6.734	11.70	9.02	7.46
1-Octanol	3.997	1.831	6.729	11.20	8.63	7.14
1-Decanol	4.050	1.832	6.783	10.50	8.11	6.67
Phenol	0.695	3.643	6.132	14.90	5.02	14.03
Acetone	2.792	0.000	2.792	7.00	7.00	0.00
Methyl ethyl ketone	2.770	0.000	2.770	5.10	5.10	0.00
Methyl isobutyl ketone	2.518	0.000	2.518	4.10	4.10	0.00
Tetrahydrofuran	3.516	0.000	3.516	8.00	8.00	0.00
1,4-Dioxane	4.245	0.000	4.245	9.00	9.00	0.00
Butyl acetate	2.258	0.000	2.258	6.30	6.30	0.00
Ethyl acetate	2.293	0.000	2.293	7.20	7.20	0.00
Diethyl ether	2.617	0.000	2.617	4.60	4.60	0.00
Methyl acetate	2.095	0.000	2.095	7.60	7.60	0.00
Ethyl formate	1.605	0.000	1.605	8.40	8.40	0.00
Acetonitrile	1.347	0.000	1.347	6.10	6.10	0.00
Diethylamine	3.962	0.011	3.979	6.10	6.09	0.40
N,N-dimethylformamide	6.009	0.000	6.009	11.30	11.30	0.00
Dimethyl sulfoxide	9.741	0.003	9.745	10.20	10.20	0.22
Acetic acid	1.975	3.830	7.691	13.50	6.84	11.64
Dichloromethane	0.000	0.226	0.337	7.10	0.00	7.10
Chloroform	0.000	1.378	2.057	5.70	0.00	5.70
Propylene glycol	5.184	2.440	8.825	21.30	16.32	13.68
Pyridine	3.394	0.000	3.394	5.90	5.90	0.00
Formamide	5.818	3.616	11.214	19.00	13.69	13.18
Glycerol	6.459	2.639	10.398	27.20	21.44	16.74
N-methyl-2-pyrrolidone	6.948	0.000	6.948	7.20	7.20	0.00
Propionic acid	1.938	3.719	7.488	12.40	6.31	10.68
N,N-dimethylacetamide	6.870	0.000	6.870	9.40	9.40	0.00
Methyl formate	1.421	0.000	1.421	10.20	10.20	0.00
Acetophenone	2.054	0.000	2.054	3.70	3.70	0.00
Benzyl alcohol	2.837	1.752	5.452	13.70	9.88	9.49
Dichloromethane	0.000	0.226	0.337	7.10	0.00	7.10
Trichloroethylene	0.000	0.226	0.337	5.30	0.00	5.30

In Table 2 are compared the radii R^2, as calculated by Eqs. (4) and (5), for pairs of paracetamol with 24 solvents. The miscibility data for these systems and for all systems reported in Tables 2–5 are obtained from the critical compilations of Jouyban (Jouyban, 2010). For a given solubility limit (in the case of paracetamol it is set equal to 100 g/kg of solvent), the *overall* radius of solubility is selected so that it gives the maximum number of successful guesses. Thus, for paracetamol the selected overall radii R^2 for the plain Hansen criterion and for the present *Extended* criterion are 50 and 18, respectively. One guess is successful when the calculated R^2 is smaller than the overall radius of solubility for soluble systems or is bigger than the overall radius of solubility for insoluble systems. As an example, paracetamol is soluble in ethanol (solubility bigger than the solubility limit) and the calculated radii are lower than the corresponding overall radii of solubility for both criteria – Eqs. (4) and (5). This case is indicated in the last column of Table 2 by the triplet S, Y, Y (soluble system, successful guess with Hansen criterion, successful guess with *Extended* criterion) in the corresponding raw for ethanol. As another example, paracetamol is insoluble in 1-hexanol. However, the Hansen criterion fails in this case since it predicts soluble system. In contrast, the *Extended* criterion successfully predicts that this is an insoluble system. Thus, this system is indicated in the last column of Table 2 by the triplet I, N, Y (insoluble, unsuccessful guess with Hansen criterion, successful guess with *Extended* criterion). As observed in Table 2, the new *Extended* criterion is successful in 20 cases while the plain Hansen

Table 2
Evaluation of the solubility of paracetamol in various solvents (solubility limit: 100 g/kg solvent).

Solvent	Solubility (Jouyban, 2010) g/kg of solvent	R^2 HSP	R^2 *Extended*	Results
Water	17.39	857.97	51.41	I,Y,Y[a]
Methanol	371.61	112.24	22.53	S,N,N
Ethanol	232.75	49.14	−4.57	S,Y,Y
Ethylene glycol	144.3	150.74	−9.08	S,N,Y
1-Propanol	132.77	38.9	7.80	S,Y,Y
Isopropanol	135.01	41.61	12.04	S,Y,Y
1-Butanol	93.64	39.61	18.87	I,N,Y
1-Pentanol	67.82	35.6	23.89	I,N,Y
1-Hexanol	49.71	38.49	25.96	I,N,Y
1-Heptanol	37.43	44.84	30.13	I,N,Y
1-Octanol	27.47	50.5	33.76	I,Y,Y
Acetone	111.65	68.78	−23.07	S,N,Y
Methyl ethyl ketone	69.99	92.65	−17.02	I,Y,N
Methyl isobutyl ketone	17.81	140.4	18.45	I,Y,Y
Tetrahydrofuran	155.37	61.85	−23.52	S,N,Y
1,4-Dioxane	17.08	100.06	19.17	I,Y,Y
Ethyl acetate	10.73	87.93	−2.46	I,Y,N
Acetonitrile	32.83	142.09	42.70	I,Y,Y
Diethylamine	1316.9	161.72	64.91	S,N,N
N,N-dimethylformamide	1012.02	17.64	−60.54	S,Y,Y
Dimethyl sulfoxide	1132.56	49.94	−26.82	S,Y,Y
Acetic acid	82.72	49.97	80.14	I,N,Y
Dichloromethane	0.32	59.04	57.67	I,Y,Y
Chloroform	1.54	122	90.78	I,Y,Y
Overall radius of solubility (squared)		*50*	*18*	
Successful guesses		14/24	20/24	

[a] The first letter indicates whether the compound is soluble (S) or insoluble (I) in the given solvent. The second letter indicates whether the plain Hansen test is successful (Y) or unsuccessful (N). The third letter indicates whether the present COSMO test is successful (Y) or unsuccessful (N).

criterion is successful for 14 cases out of the 24 total tests. It is also observed that the overall radius of solubility is significantly smaller for the former criterion.

Similar pictures emerge from Tables 3–5. In Table 3 are compared the radii R^2 for pairs of aspirin with 23 solvents. In this case and for the set limit of solubility, the overall radii of solubility with the Hansen and *Extended* criteria are 40 and 22, respectively, while the successful guesses are 14 and 18, respectively, out of the 23 total tests with a variety of solvents. Negative values for R^2 are often associated with high solubilities and the high positive values with low solubilities, but we will come back to this point later in the discussion.

Results for the solubility of benzoic acid are reported in Table 4 where are compared the radii R^2 for pairs of benzoic acid with 29 solvents. In this case, the overall radii of solubility with the Hansen and *Extended* criteria are 100 and 53, respectively, while the successful guesses are 18 and 20, respectively, out of the 29 total tests. The corresponding results for salicylic acid are reported in Table 5 for which the overall radii of solubility with the Hansen and *Extended* criteria are 81 and 35, respectively, while the

Table 3
Evaluation of the solubility of aspirin in various solvents (solubility limit: $X_{aspirin} = 0.044$).

Solvent	Solubility (Jouyban, 2010) mole fraction of aspirin, 25 °C	R^2 HSP	R^2 *Extended*	Results
Methanol	0.0719	275.45	83.48	S,N,N
Ethanol	0.0855	147.81	17.66	S,N,Y
Isopropanol	0.05232	91.62	13.24	S,N,Y
1-Butanol	0.0453	79.06	16.26	S,N,Y
1-Pentanol	0.0395	60.09	24.89	I,Y,Y
1-Hexanol	0.0393	49.32	26.40	I,Y,Y
1-Heptanol	0.03892	43.45	25.85	I,Y,Y
1-Octanol	0.0386	42.17	27.23	I,Y,Y
Acetone	0.0828	68.73	10.38	S,N,Y
Tetrahydrofuran	0.1904	21.86	−40.47	S,Y,Y
1,4-Dioxane	0.0516	32.13	−36.18	S,Y,Y
Ethyl acetate	0.0448	47.06	−11.93	S,N,Y
Acetonitrile	0.0185	194.96	138.48	I,Y,Y
Chloroform	0.206	30.97	61.22	S,Y,N
1-Decanol	0.03652	41.05	28.64	I,Y,Y
2-Butanol	0.0536	68.81	21.84	S,N,Y
2-Methyl-1-propanol	0.03186	105.21	51.35	I,Y,Y
2-Methyl-2-propanol	0.06844	89.17	28.87	S,N,N
Butyl acetate	0.03345	58.37	1.62	I,Y,N
Diethyl ether	0.03529	116.78	59.82	I,Y,Y
Methyl acetate	0.05287	52.25	−8.25	S,N,Y
Propylene glycol	0.017 (22.55 °C)	177.80	13.77	I,Y,N
Pyridine	0.5348	16.40	−39.88	S,Y,Y
Radius of solubility (squared)		*40*	*22*	
Successful guesses		14/23	18/23	

Table 4
Evaluation of the solubility of benzoic acid in various solvents (solubility limit: $X_{benz.\ acid} = 0.178$).

Solvent	Solubility (Jouyban, 2010) mole fraction of benzoic acid, 25 °C	R^2 HSP	R^2 *Extended*	Results
Methanol	0.1632	273.77	110.55	I,Y,Y
Ethanol	0.1789	148.13	36.23	S,N,Y
Isopropanol	0.1937	102.56	33.49	S,N,Y
1-Butanol	0.2016	90.44	37.73	S,Y,Y
1-Pentanol	0.1839	77.85	49.30	S,Y,Y
1-Hexanol	0.1905	71.34	51.36	S,Y,Y
1-Heptanol	0.1946	67.37	50.60	S,Y,Y
1-Octanol	0.1987	67.77	52.33	S,Y,Y
Acetone	0.1857	107.69	21.71	S,N,Y
Tetrahydrofuran	0.0734	50.24	−39.36	I,N,N
Ethyl acetate	0.1649	86.08	−0.46	I,N,N
Acetonitrile	0.0539	233.66	149.23	I,Y,Y
Chloroform	0.1283	59.81	92.05	I,N,Y
2-Methyl-1-propanol	0.1524	123.49	83.60	I,Y,Y
Butyl acetate	0.1699	101.05	16.41	I,Y,N
Diethyl ether	0.1837	175.44	89.99	S,N,N
1,4-Dioxane	0.2853	54.25	−40.97	S,Y,Y
Acetic acid	0.1675	129.5	171.25	I,Y,Y
Acetophenone	0.1878	53.94	−34.28	S,Y,Y
Benzyl alcohol	0.1441	19.01	6.56	I,N,N
N,N-dimethylformamide	0.4909	73.53	−42.21	S,Y,Y
Dimethyl sulfoxide	0.5102	100.85	−1.50	S,N,Y
Ethylene glycol	0.0884	289.05	36.33	I,Y,N
Formamide	0.1525	471.09	398.68	I,Y,Y
Glycerol	0.0164	315.36	−1.62	I,Y,N
N-Methyl-2-pyrrolidone	0.1562 (23.8 °C)	58.12	−28.42	I,N,N
Propionic acid	0.1887	117.48	159.55	S,N,N
N,N-dimethylacetamide	0.5254	64.08	−33.95	S,Y,Y
Methyl formate	0.3428	90.97	−13.63	S,Y,Y
Radius of solubility (squared)		*100*	*53*	
Successful guesses		18/29	20/29	

Table 5
Evaluation of the solubility of salicylic acid in various solvents (solubility limit: $X_{salic.\ acid} = 0.100$).

Solvent	Solubility (Jouyban, 2010) mole fraction of salicylic acid, 25 °C	R^2 HSP	R^2 *Extended*	Results
Methanol	0.1321	158.81	34.12	S,N,Y
Ethanol	0.1100	73.49	−21.15	S,Y,Y
1-Pentanol	0.1547	48.09	17.03	S,Y,Y
1-Octanol	0.1549	58.25	26.97	S,Y,Y
Acetone	0.1792	95.85	−66.22	S,N,Y
Ethyl acetate	0.1223	96.90	−65.88	S,N,Y
Acetonitrile	0.0294	193.52	33.65	I,Y,N
Chloroform	0.0015	104.57	135.99	I,Y,Y
Diethyl ether	0.1521	201.02	41.21	S,N,Y
1,4-Dioxane	0.2979	82.93	−89.82	S,N,Y
Acetic acid	0.0549	82.04	162.33	I,Y,Y
Acetophenone	0.1527	101.60	−60.34	S,N,Y
N,N-dimethylformamide	0.3840	38.09	−156.83	S,Y,Y
Ethylene glycol	0.0851	174.82	−11.98	I,Y,N
Formamide	0.0426	336.89	300.34	I,Y,Y
Glycerol	0.0746	197.78	−63.12	I,Y,N
Propionic acid	0.0661	89.34	160.75	I,Y,Y
Ethyl formate	0.0076	77.45	−91.26	I,N,N
Propylene glycol	0.1136	79.08	−22.91	S,Y,Y
Radius of solubility (squared)		*81*	*35*	
Successful guesses		13/19	15/19	

successful guesses are 13 and 15, respectively out of the 19 total tests.

4. Discussion

The overall picture that emerges from the comparisons made in Tables 2–5 is that the split of the partial solubility parameter for hydrogen bonding, δ_{hb}, into its acidic and basic components, δ_a and δ_b, respectively, along with the *Extended* criterion of Eq. (5), leads to an improved scheme over the classical Hansen (2007) approach for the prediction of solubility and the selection of the appropriate solvents for a given solute. There is some degree of arbitrariness in the selection of the solubility limits in Tables 2–5 but the above overall picture does not change by reasonably changing these limits. It should be stressed at this point that the new scheme is far from perfect but a significant proportion of its failure may not be due to the wrong split of the hydrogen bonding parameter but rather to the restrictions that impose the very character of this parameter in the frame of Hansen's solubility parameter approach. As mentioned earlier, even the hydrogen bonding solubility parameter is evaluated on the basis of the "similarity" principle for solubility. According to this principle, out of two otherwise similar candidate solvents for a given solute, the selected one will be the solvent having δ_{hb} closer to the corresponding δ_{hb} of the solute

regardless of its acidic or basic character. In other words, an acidic solvent is not given any priority as solvent for a basic solute. This is, of course, in disaccord with common experience and since the split in δ_a and δ_b was done on the basis of these tabulated Hansen's δ_{hb}, the new acidic and basic solubility parameters inherit to an extent the concomitant drawback. As an example, we see in Table 2 that diethylamine is the best of the solvents in this table for paracetamol but, both, the classical Hansen scheme and the new scheme fail to predict it. Yet, even on the basis of these δ_{hb} parameters, the performance of the new scheme is significantly improved over the classical Hansen prediction scheme. A relevant discussion on acid–base reactions and the relationship of the HSPs of resulting organic salts to the corresponding HSPs of acid and base reactants which they are derived from is provided by Hansen (2007).

Another point that should be stressed regards the adopted *Extended* criterion of Eq. (5). This is probably an over simplistic criterion since it does not portray all relevant features of the interacting species. It favors the cross association (one interacting molecule is acidic and the other is basic) and disfavors interaction between two acids or between two bases, but it does not disfavor acid–base interactions within the same molecule (self-association). Indeed, the addition of such a term in the *Extended* criterion of Eq. (5) might improve further the new scheme. As an example, the replacement of the last term in Eq. (5) with the term

$$(\delta_{a1} - \delta_{b1})(\delta_{a2} - \delta_{b2}) + (\delta_{a1} - \delta_{a2})(\delta_{b1} - \delta_{b2}) \quad (6)$$

leads to 21/24 successful guesses (instead of 20/24 with the criterion of Eq. (5)) for paracetamol, to 22/29 successful guesses (instead of 20/29 with the criterion of Eq. (5)) for benzoic acid, to 15/19 successful guesses (as with the criterion of Eq. (5)) for salicylic acid, but for aspirin it leads to 16/23 successful guesses (instead of 18/23 with the criterion of Eq. (5)). Obviously, many more tests should be performed before securing that the terms of Eq. (6) present an improvement in the *Extended* criterion of Eq. (5) by replacing the last term of it.

In this work, we have tried to avoid involved equations and complex thermodynamic analysis, respecting the simplicity of the concept of solubility parameter as being the source of its strength. Such an analysis, however, would reveal in more realistic terms the potential of the present approach. In thermodynamic terms, the function or quantity that dictates solubility is the free energy change on mixing rather than the plain energy exchange. Thus, for a binary mixture of a solvent 1 with a polymeric solute 2, the classical Flory–Huggins theory (Hildebrand and Scott, 1962) gives the following expression for the molar Gibbs free energy of mixing:

$$\frac{\Delta G^M}{RT} = x_1 \ln \varphi_1 + x_2 \ln \varphi_2 + x_1 \varphi_2 \chi_{12} \quad (7)$$

where, x_i and φ_i are the mole fraction and volume fraction, respectively, of component i in the mixture. The Flory–Huggins χ_{12} parameter is the one that makes the difference from system to system and dictates its miscibility. So far, the relation of this parameter to the solubility parameters was given, almost invariably, by the following equation (Hansen, 2007)

$$\chi_{12} = \frac{V_1}{4RT}\left[4\left(\delta_{d1} - \delta_{d2}\right)^2 + \left(\delta_{p1} - \delta_{p2}\right)^2 + \left(\delta_{hb1} - \delta_{hb2}\right)^2\right] \quad (8)$$

It is clear from this equation that, regardless of the strength of intermolecular interactions, this equation is calculating χ_{12} as being always positive. With our approach, Eq. (8) changes andbecomes

$$\chi_{12} = \frac{V_1}{4RT}\left[4\left(\delta_{d1} - \delta_{d2}\right)^2 + \left(\delta_{p1} - \delta_{p2}\right)^2 + \left(\delta_{a1} - \delta_{a2}\right)\left(\delta_{b1} - \delta_{b2}\right) + \left(\delta_{a1} - \delta_{b1}\right)\left(\delta_{a2} - \delta_{b2}\right)\right] \quad (9)$$

where, V_1 is the molar volume of the solvent. Eq. (9) does allow for negative values (higher miscibility) of the χ_{12} parameter in cases of strong specific intermolecular interactions such as the strong acid–base or hydrogen-bonding cross-associations. In addition, the strong quantum chemical basis of our splitting process of the δ_{hb} parameter warrants an expansion of the approach to an equation of state framework, as we have shown recently (Stefanis et al., 2006; Stefanis and Panayiotou, 2008; Panayiotou, 2011). This, however, will be the subject of a forthcoming publication.

5. Conclusions

In this work we have presented a new and straightforward approach for the evaluation of the acidic and basic components of the hydrogen bonding solubility parameter. The application of the new method for splitting δ_{hb} into its acidic and basic components, δ_a and δ_b, requires two types of data. First, it requires data for the δ_{hb} itself. These data may be obtained either directly from available compilations (Abbott and Hansen, 2010; Hansen, 2007) or from calculations via robust and reliable approaches such as the updated group contribution method presented in Appendix A of this work. The second type of data requires data for the COSMO-RS third σ-moments for the hydrogen bonding acceptor and donor functions, *HB_acc3* and *HB_don3*, respectively. These data may be obtained directly from available compilations (COSMObase, 2006; VT Sigma Profile Databases, 2006) for thousands of compounds, or they may be derived from the sigma profiles of surface screening charge distributions obtained via widely available quantum chemical calculation software suites. In this sense, the calculation of the acidic and basic components, δ_a and δ_b is straightforward via the simple Eqs. (2) and (3). No additional experimental information is required and, thus, the proposed solubility scheme, as embodied in the *Extended* criterion of Eq. (5), is a purely predictive scheme. This is a most simple scheme and does not require any sophisticated calculations for its application. The extensive calculations presented in this work have shown that the new (4-parameter) approach of the extended solubility parameters is significantly improved over the classical (3-parameter) Hansen solubility parameter approach.

Appendix A.

An Update of the Stefanis–Panayiotou Group-Contribution Method (Stefanis and Panayiotou, 2008)

A.1. Introduction to group-contribution methods

Computer-aided molecular design is a very important tool for the prediction of properties of organic compounds, especially in the case of lack of experimental data, but most of all, for the selection of compounds with desired properties. In the last decades various characteristic group-contribution methods have been introduced. These methods have been widely used for the prediction of physicochemical properties of pure organic compounds. One of the first group-contribution methods was the UNIFAC method (Fredenslund et al., 1977), in which the values of each property were calculated from the sum of the contributions of simple first-order groups. Similar group-contribution methods, which were presented later, are the method by Joback and Reid (1987) and the method by Horvath (1992).

In an alternative category of group-contribution methods (Constantinou and Gani, 1994; Mavrovouniotis, 1990), second-order groups are defined to provide more structural information, to distinguish between isomers and to afford more accurate predictions. Second-order groups have a strong physicochemical meaning and can significantly improve the accuracy of property predictions. The definition of second-order groups is based on the theory of conjugation operators (Constantinou et al., 1993; Mavrovouniotis, 1990). Marrero and Gani introduced a higher level of approximation by defining third-order groups to provide more structural information about systems of fused aromatic and non-aromatic rings (Marrero and Gani, 2001).

In 2004, the Constantinou-Gani group-contribution method was extended by Stefanis et al. to new classes of compounds, which are of significant importance for the chemical, biochemical, pharmaceutical, and food industries, as well as for the environmental protection (Stefanis et al., 2004). The method included a large variety of first-order and second-order groups and was able to estimate properties of organic compounds with complex multi-ring, heterocyclic, and aromatic structures.

The Stefanis et al. method predicted a series of physicochemical properties: octanol–water partition coefficient (*logKow*), total (Hildebrand) solubility parameter at 25 °C, flash point (Stefanis et al., 2004), the three characteristic scaling constants (ε^*, v^*, $v_{sp}{}^*$) and the influence parameter, κ, of the NRHB model for estimating vapour pressure, liquid density, heat of vaporization and surface tension at various temperatures and pressures (Stefanis et al., 2005).

As already mentioned, the use of Hansen solubility parameters is much more appropriate than the total solubility parameter for the selection of solvents. For this reason, the Stefanis et al. group-contribution method was extended in 2008 to predict the Hansen solubility parameters (Stefanis and Panayiotou, 2008). The development of such a predictive method is of crucial importance not only for selecting the appropriate solvents for each given solute but also for the synthesis of new solvents with desired solubility properties.

A.2. The essentials of the new group-contribution method

In the Stefanis–Panayiotou group-contribution method, the molecular structure of each organic compound can be described using two kinds of functional groups: first-order groups (UNIFAC groups), which describe the basic molecular structure of compounds (Fredenslund et al., 1977) and second-order groups, which have the first-order groups as building blocks. The second-order groups improve the accuracy of predictions significantly and give a physical meaning to the method. This physical meaning is related to that their definition is based on the theory of conjugation operators, as formulated by the ABC framework (Constantinou and Gani, 1994). The principles concerning the determination of second-order groups and the methodology that is followed for their identification is thoroughly described in literature (Constantinou et al., 1993; Stefanis et al., 2004).

The basic equation of the model, which gives the values of solubility parameters according to the molecular structure of compounds, is Eq. (A.1):

$$\delta = \sum_i N_i C_i + W \sum_j M_j D_j \tag{A.1}$$

where C_i is the contribution of the first-order group of type i that appears N_i times in the compound and D_j is the contribution of the second-order group of type j that appears M_j times in the compound. In Eq. (A.1), δ is a single equation of a solubility parameter (δ_d, δ_p or δ_{hb}) and is selected after a thorough study. The constant W is equal to 0 for compounds without second-order groups and equal to 1 for compounds with second-order groups.

The calculation of group contributions is done by a two-step regression analysis. In the first step, the aim is to determine the first-order group contributions only (which are, the C_i's). In the second step, using the C_i's contributions, the second-order groups are activated and the second-order group contributions (D_j's) are calculated through regression. These contributions act as a correction to the first-order approximation.

The Levenberg–Marquardt approach is used to minimize the total sum of squared errors between experimental and predicted values of solubility parameters. This is the criterion for the selection of the most appropriate equation to fit the experimental data. It should be stressed that the model is applicable to organic compounds with three or more carbon atoms, excluding the atom of the characteristic group (e.g., –COOH or –CHO). Along these lines, Stefanis and Panayiotou developed in 2008 a simple, yet quite accurate method for the estimation of HSPs of pure organic compounds (Stefanis and Panayiotou, 2008).

A.3. Prediction of Hansen solubility parameters with the Stefanis–Panayiotou group-contribution method

This Appendix presents a recent update of the group contributions and equations which are useful for the estimation of Hansen solubility parameters, δ_d, δ_p and δ_{hb} with the Stefanis–Panayiotou method. Recently, it was concluded that a polynomial version of Eq. (A.1) fit the δ_d experimental data best. It gave a smaller sum of total squared errors between estimated and experimental values and, therefore, more accurate results compared to those obtained by applying the previous (Stefanis and Panayiotou, 2008), linear version of the equation. Slight improvements have been made to the group contributions and equations for the estimation of δ_p και δ_{hb}. These new updated group-contributions to Hansen solubility parameters are now shown in Tables A.1 and A.2. The updated equations for the estimation of Hansen solubility parameters are the following (Eqs. (A.2)–(A.4)):

$$\delta_d = \left(\sum_i N_i C_i + \sum_j M_j D_j + 959.11 \right)^{0.4126} \quad \mathrm{MPa}^{(1/2)} \tag{A.2}$$

$$\delta_p = \left(\sum_i N_i C_i + \sum_j M_j D_j + 7.6134 \right) \quad (\mathrm{MPa})^{(1/2)} \tag{A.3}$$

$$\delta_{hb} = \left(\sum_i N_i C_i + \sum_j M_j D_j + 7.7003 \right) \quad (\mathrm{MPa})^{(1/2)} \tag{A.4}$$

It is very important to stress that Eqs. (A.3) and (A.4) are valid only for Hansen solubility parameter values greater than 3 $\mathrm{MPa}^{(1/2)}$.

The statistical values of the first-order and second-order approximations for the prediction of Hansen solubility parameters are presented in Tables A.3 and A.4. One can see the improvement of the statistical values, compared to those of the previous article (Stefanis and Panayiotou, 2008). It should be mentioned that the values of correlation parameters reported in reference 24 are R values and not R^2 ones as erroneously indicated on the corresponding figures.

Figs. A.1–A.3 show the correlation between estimated and experimental values of Hansen solubility parameters.

The group contributions in the case of low δ_p or low δ_{hb} (less than 3 $\mathrm{MPa}^{(1/2)}$) are presented in Table A.5 (first-order groups) and

Table A.1
First-order group contributions to the dispersion partial solubility parameter, δ_d, the polar partial solubility parameter, δ_p, and the hydrogen-bonding partial solubility parameter, δ_{hb}.

First-order groups	δ_d	δ_p	δ_{hb}	Examples (occurrences)
$-CH_3$	−123.01	−1.6444	−0.7458	Propane (2)
$-CH_2-$	1.82	−0.3141	−0.3877	Butane (2)
–CH<	82.94	0.6051	−0.2064	Isobutane (1)
>C<	182.13	2.0249	−0.0113	Neopentane (1)
CH_2═CH–	−126.15	−2.0170	−1.1783	Propylene (1)
–CH═CH–	28.65	−0.5037	−0.1253	cis-2-Butene (1)
CH_2═C<	−31.62	−0.9052	−0.7191	Isobutene (1)
–CH═C<	62.48	−1.1018	−1.7171	2-Methyl-2-butene (1)
>C═C<	50.10	0.9957	−1.9773	2,3-Dimethyl-2-butene (1)
CH_2═C═CH–	−161.71	***	−0.7545	1,2-Butadiene (1)
CH≡C–	45.86	−1.5147	1.2582	Propyne (1)
C≡C	9.56	−0.9552	−1.0176	2-Butyne (1)
ACH	29.87	−0.5771	−0.3554	Benzene (6)
AC	98.84	0.7661	−0.1553	Naphthalene (2)
$ACCH_3$	27.67	−0.6212	−1.1409	Toluene (1)
$ACCH_2-$	89.07	0.8019	−0.2298	m-Ethyltoluene (1)
CH_3CO	−29.41	2.1567	−1.1683	Methyl ethyl ketone (1)
CH_2CO	114.74	3.6103	−0.3929	Cyclopentanone (1)
CHO (aldehydes)	−31.35	3.3159	0.2062	1-Butanal (1)
COOH	−38.16	0.7153	3.8422	Vinyl acid (1)
CH_3COO	−53.86	−0.6075	1.7051	Ethyl acetate (1)
CH_2COO	89.11	3.4942	1.3893	Methyl propionate (1)
HCOO	***	1.7056	2.3049	n-Propyl formate (1)
COO	27.57	3.3401	1.1999	Ethyl acrylate (1)
OH	−29.97	1.0587	7.3609	Isopropanol (1)
ACOH	58.52	1.0520	6.9757	Phenol (1)
CH_3O	−68.07	0.0089	0.2676	Methyl ethyl ether (1)
CH_2O	13.40	0.8132	−0.1196	Ethyl vinyl ether (1)
CHO (ethers)	111.46	1.6001	0.4873	Diisopropyl ether (1)
$C_2H_5O_2$	15.51	3.3880	8.5893	2-Methoxy-ethanol (1)
CH_2O (cyclic)	49.32	0.1227	0.1763	1,4-Dioxane (2)
CH_2NH_2	−49.96	−0.3449	2.7280	1-Amino-2-propanol (1)
$CHNH_2$	18.53	−1.4337	0.5647	Isopropylamine (1)
CH_3NH	***	0.5060	5.7321	n-Methylaniline (1)
CH_2NH	96.18	0.2616	1.4053	di-n-Propylamine (1)
CH_3N	170.59	1.0575	1.8500	Trimethylamine (1)
CH_2N	152.54	2.6766	1.5557	Triethylamine (1)
$ACNH_2$	253.66	1.6493	4.4945	Aniline (1)
$CONH_2$	−1.22	5.9361	5.3646	2-Methacrylamide (1)
$CON(CH_3)_2$	95.97	5.5309	3.2455	N,N-dimethylacetamide (1)
CH_2SH	214.84	−0.9940	4.5321	N-butyl mercaptan (1)
CH_3S	***	0.2451	−1.2669	Methyl ethyl sulfide (1)
CH_2S	168.57	0.5730	−0.0838	Diethyl sulfide (1)
I	197.67	0.1060	0.3321	Isopropyl iodide (1)
Br	109.79	0.5207	−0.9087	2-Bromopropane (1)
CH_2Cl	47.17	0.5013	−0.4498	n-Butyl chloride (1)
CHCl	73.01	2.6796	−1.3563	Isopropyl chloride (1)
CCl	385.39	1.8196	0.1473	t-Butyl chloride (1)
$CHCl_2$	197.67	1.6255	−3.0669	1,1-Dichloropropane (1)
CCl_2	72.60	0.1035	−1.3220	Pentachlorocyclopropane (2)
CCl_3	***	1.1060	−2.5679	Benzotrichloride (1)
ACCl	141.54	−0.0941	−0.7512	m-Dichlorobenzene (2)
ACF	27.74	0.1293	−0.6613	Fluorobenzene (1)
Cl–(C═C)	45.32	2.2673	−0.5258	2,3-Dichloropropene (1)
CF_3	−13.79	−2.1381	−1.2997	Perfluorohexane (2)
CH_2NO_2	***	6.6451	−1.0669	1-Nitropropane (1)
$CHNO_2$	***	7.7753	−2.1087	2-Nitropropane (1)
$ACNO_2$	219.22	4.4640	−0.7302	Nitrobenzene (1)
CH_2CN	−29.09	6.3586	−0.7297	n-Butyronitrile (1)
CF_2	−103.83	***	***	Perfluoromethylcyclohexane (5)
CF	20.32	***	***	Perfluoromethylcyclohexane (1)
F (except as above)	−80.11	***	***	2-Fluoropropane (1)
CH_2═C═C<	6.64	***	−1.7087	3-Methyl-1,2-butadiene (1)
O (except as above)	18.09	3.5248	0.0883	Divinyl ether (1)
Cl (except as above)	76.35	1.7491	−0.2917	Hexachlorocyclopentadiene (2)
>C═N–	−10.55	−0.1692	−5.3820	2,4,6-Trimethylpyridine (1)
–CH═N–	186.40	2.7015	0.5507	Isoquinoline (1)
NH (except as above)	***	−0.0746	2.0646	Dibenzopyrrole (1)
CN (except as above)	49.36	6.3705	−0.5239	cis-Crotonitrile (1)
O═C═N–	15.22	1.4695	4.1129	n-Butyl isocyanate (1)
SH (except as above)	190.87	1.8229	4.9279	2-Mercaptobenzothiazole (1)
S (except as above)	201.91	8.5982	−0.4013	Thiophene (1)
SO_2	182.83	11.0254	−0.3602	Sulfolene (1)
>C═S	−0.46	0.5216	3.0519	n-Methylthiopyrrolidone (1)
>C = 0 (except as above)	−127.16	0.7691	1.7033	Anthraquinone (2)
N (except as above)	267.06	2.2212	1.3655	Triphenylamine (1)

*** Not available.

Table A.2
Second-order group contributions to the dispersion partial solubility parameter, δ_d, the polar partial solubility parameter, δ_p, and the hydrogen-bonding partial solubility parameter, δ_{hb}.

Second-order groups	δ_d	δ_p	δ_{hb}	Examples (occurrences)
$(CH_3)_2$–CH–	7.63	0.0365	0.3019	Isobutane (1)
$(CH_3)_3$–C–	−0.03	1.1593	−0.1924	Neopentane (1)
Ring of 5 carbons	−81.93	−2.3673	0.2586	Cyclopentane (1)
Ring of 6 carbons	−26.15	−3.6661	***	Cyclohexane (1)
–C=C–C=C–	−5.69	−3.3100	−1.2207	1,3-Butadiene (1)
CH_3–C=	−4.45	0.3461	0.4418	Isobutene (2)
$-CH_2$–C=	−29.67	−2.3189	−0.5613	1-Butene (1)
>C{H or C}–C=	−5.38	***	−1.0241	3-Methyl-1-butene (1)
String in cyclic	−54.05	***	***	Ethylcyclohexane (1)
$CH_3(CO)CH_2$–	3.57	−0.4108	−0.3628	Methyl ethyl ketone (1)
C_{cyclic}=O	−46.57	0.1972	−0.4496	Cyclopentanone (1)
ACCOOH	−37.57	−0.6284	−0.8552	Benzoic acid (1)
>C{H or C}–COOH	***	−0.2450	1.2554	Isobutyric acid (1)
CH3(CO)OC{H or C}<	−40.56	−0.0652	0.3864	Isopropyl acetate (1)
(CO)C{H_2}COO	***	−2.3624	0.8545	Ethyl acetoacetate (1)
(CO)O(CO)	−92.46	−0.9818	1.5759	Acetic anhydride (1)
ACHO	46.84	−1.8120	−0.9192	Benzaldehyde (1)
>CHOH	16.54	0.2366	−0.2453	2-Propanol (1)
>C < OH	−5.97	−0.0069	1.3813	Tert-Butanol (1)
–C(OH)C(OH)–	***	0.6669	0.2493	1,2-Propanediol (1)
–C(OH)C(N)	−7.03	0.8750	−0.7322	1-Amino-2-propanol (1)
C_{cyclic}–OH	−6.40	−3.6065	0.5836	Cyclohexanol (1)
C–O–C=C	25.23	0.5480	1.1279	Ethyl vinyl ether (1)
AC–O–C	35.82	0.7781	0.6689	Methyl phenyl ether (1)
>N{H or C}(in cyclic)	53.29	−1.6876	−0.0132	Cyclopentimine (1)
–S–(in cyclic)	91.57	0.2513	0.2663	Tetrahydrothiophene (1)
ACBr	33.35	−0.4478	0.3149	Bromobenzene (1)
(C=C)–Br	−85.85	0.0686	−1.1154	2-Bromopropene (1)
Ring of 3 carbons	15.35	1.7962	−0.7224	Cyclopropane (1)
ACCOO	−38.39	0.3670	−0.2340	Methyl benzoate (1)
$AC(ACH_m)_2AC(ACH_n)_2$	−33.14	−1.4784	0.7468	Naphthalene (1)
O_{cyclic}–C_{cyclic}=O	13.89	2.7261	0.2185	Diketene (1)
AC–O–AC	−136.10	−3.4995	1.8763	Diphenyl ether (1)
$C_{cyclic}H_m$=N_{cyclic}–$C_{cyclic}H_n$=$C_{cyclic}H_p$	53.11	−0.5075	−2.1004	2,6-Dimethylpyridine (1)
$N_{cyclic}H_m$–C_{cyclic}=O	93.54	2.0813	1.2226	2-Pyrrolidone (1)
–O–CH_m–O–CH_n–	31.52	0.3293	0.2527	Methylal (1)
C(=O)–C–C(=O)	−61.38	−0.4126	1.2240	2,4-Pentanedione (1)

*** Not available.

Table A.6 (second-order groups). The equations for the estimation of δ_p and δ_{hb} in such cases are the following:

$$\delta_p = \left(\sum_i N_i C_i + \sum_j M_j D_j + 2.6560 \right) (\text{MPa})^{1/2} \quad \text{(A.5)}$$

$$\delta_{hb} = \left(\sum_i N_i C_i + \sum_j M_j D_j + 1.3720 \right) (\text{MPa})^{1/2} \quad \text{(A.6)}$$

A.4. Characteristic examples

Although we had presented application examples of the Stefanis–Panayiotou method in our previous article (Stefanis and

Table A.3
Comparison of total sum of squared errors for first and second-order approximations.

	First-order approximation	Second-order approximation	Change (%)
δ_d	85	73	−14.1
δ_p	649	425	−34.5
δ_{hb}	460	390	−15.2

Table A.4
Average absolute error (AAE)[a] for first- and second-order approximations.

	First-order approximation AAE ($MPa^{(1/2)}$)	Second-order approximation AAE ($MPa^{(1/2)}$)	Change (%)
δ_d	0.37	0.33	−10.8
δ_p	1.00	0.82	−18.0
δ_{hb}	0.85	0.77	−9.4

[a] sangeeta Absolute average error = AAE = $\frac{1}{N}\sum |X_{est} - X_{exp}|$
where N is the number of data points, X_{est} is the estimated values of solubility parameters and X_{exp} the experimental values.

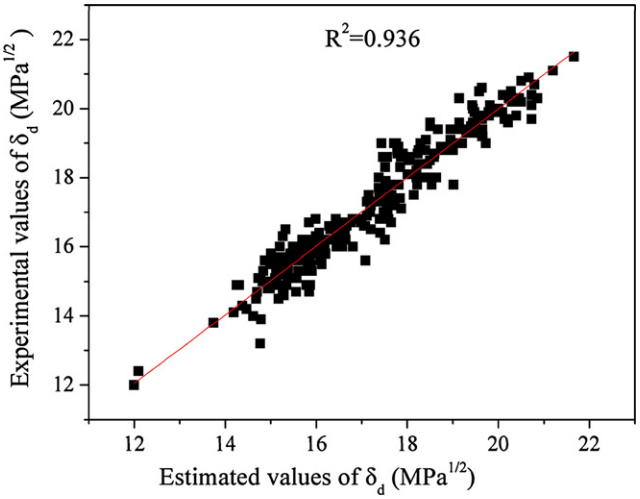

Fig. A.1. Correlation between estimated and experimental values of dispersion partial solubility parameter, δ_d (347 data points).

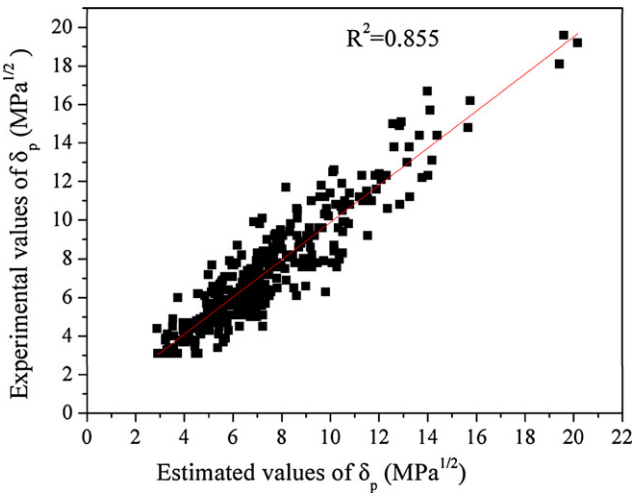

Fig. A.2. Correlation between estimated and experimental values of polar partial solubility parameter, δ_p (350 data points).

Table A.5
First-order group contributions to the polar partial solubility parameter, δ_p, and the hydrogen-bonding partial solubility parameter, δ_{hb}, when $\delta_p < 3$ MPa$^{(1/2)}$ or $\delta_{hb} < 3$ MPa$^{(1/2)}$.

First-order groups	$\delta_p < 3$ MPa$^{(1/2)}$	$\delta_{hb} < 3$ MPa$^{(1/2)}$
CH_3-	−0.7107	0.2990
$-CH_2-$	−0.1361	−0.1161
−CH<	0.6477	0.1386
CH_2=CH−	−0.2511	1.3552
−CH=CH−	−0.1503	0.4819
CH_2=C<	0.6956	0.1115
−CH=C<	1.2761	−0.0307
>C=C<	***	−0.1212
CH_2=C=CH−	−0.2453	***
CH≡C−	−0.7049	0.4385
C≡C	***	−0.3511
ACH	−0.1930	0.1353
AC	0.1745	−0.1740
$ACCH_3$	−0.4493	−0.2873
$ACCH_2-$	−0.2857	−0.8808
ACCH<	0.9303	−1.4467
−COOH	2.9098	***
CH_3COO	1.7711	***
CH_2COO	2.2096	0.3720
COO	1.4783	***
OH	***	***
CH_3O	−0.3600	***
CH_2O	***	−0.4067
CHO	***	***
CH_2O (cyclic)	−0.2919	***
CH_2NH_2	***	***
CH_2NH	0.8875	***
CHNH	1.2391	−0.1700
CH_3N	***	−1.0369
CH_2N	0.7055	0.1461
CH_2S	***	0.4895
CH_2Cl	***	0.1300
CHCl	***	0.5254
$CHCl_2$	***	0.4424
ACCl	−0.0927	***
CCl_2F	***	−0.3718
ACF	***	0.6606
Cl−(C=C)	***	−0.0887
CF_3	***	***
CH_2=C=C<	1.2654	***
O (except as above)	−0.5555	1.1251
Cl (except as above)	***	***
S (except as above)	0.0445	−0.0553
>C=O (except as above)	***	

*** Not available.

Table A.6
Second-order group contributions to the polar partial solubility parameter, δ_p, and the hydrogen-bonding partial solubility parameter, δ_{hb}, when $\delta_p < 3$ MPa$^{(1/2)}$ or $\delta_{hb} < 3$ MPa$^{(1/2)}$.

Second-order groups	$\delta_p < 3$ MPa$^{(1/2)}$	$\delta_{hb} < 3$ MPa$^{(1/2)}$
$(CH3)_2$−CH−	0.2246	0.0000
Ring of 5 carbons	−0.9657	0.1944
Ring of 6 carbons	−0.9615	0.0000
−C=C−C=C−	0.6463	***
CH_3−C=	−0.0063	−0.0614
$-CH_2$−C=	0.0192	0.0660
>C{H or C}−C=	−0.4460	0.3422
$AC(ACH_m)_2AC(ACH_n)_2$	0.0669	***
$-O-CH_m-O-CH_n-$	0.0000	***
String in cyclic	***	−0.2809
ACCOO	***	0.0000
$AC(ACH_m)_2AC(ACH_n)_2$	***	0.0864

*** Not available.

Panayiotou, 2008), we realized from the large number of messages we received since then, that the focus in the examples should be primarily on the use of *second-order groups*. With this in mind, we add below 20 carefully selected further examples which will help users apply much more easily the Stefanis–Panayiotou group-contribution method to the molecules and systems of their interest.

1. 2,4,4-trimethyl-1-pentene (Fig. A.4): $(CH_3)_3$−C− (1 occurrence), CH_3−C= (1 occurrence) and $-CH_2$−C= (1 occurrence).
2. 2-methyl-1-butene (Fig. A.5): CH_3−C= (1 occurrence) and $-CH_2$−C= (1 occurrence).
3. gamma-thiobutyrolactone (Fig. A.6): $C_{(cyclic)}$=O (1 occurrence) and −S− (in cyclic) (1 occurrence).
4. diisononyl phthalate (Fig. A.7): $(CH_3)_2$−CH− (2 occurrences) and ACCOO (2 occurrences).
5. o-bromoanisole (Fig. A.8): AC−O−C (1 occurrence) and ACBr (1 occurrence).
6. 1-bromo-propene (Fig. A.9): CH_3−C= (1 occurrence) and (C=C)−Br (1 occurrence).
7. vanillin (Fig. A.10): ACHO (1 occurrence) and AC−O−C (1 occurrence).
8. triisooctyl trimellitate (Fig. A.11): $(CH_3)_2$−CH− (3 occurrences) and ACCOO (3 occurrences).
9. alpha-terpinene (Fig. A.12): $(CH_3)_2$−CH− (1 occurrence), ring of 6 carbons (1 occurrence) and −C=C−C=C− (1 occurrence).

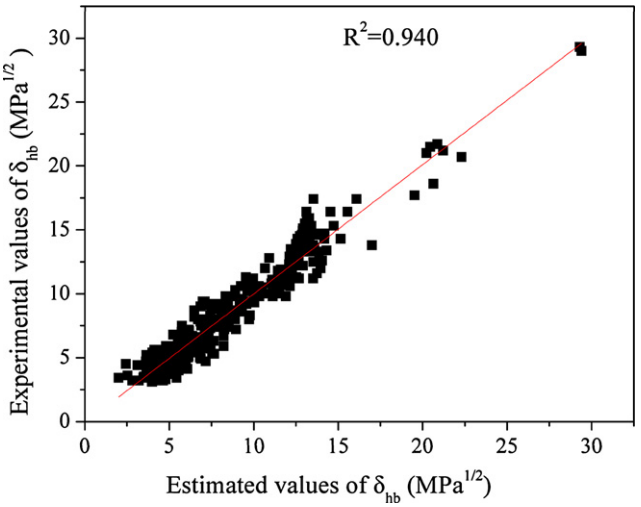

Fig. A.3. Correlation between estimated and experimental values of hydrogen-bonding partial solubility parameter, δ_{hb} (350 data points).

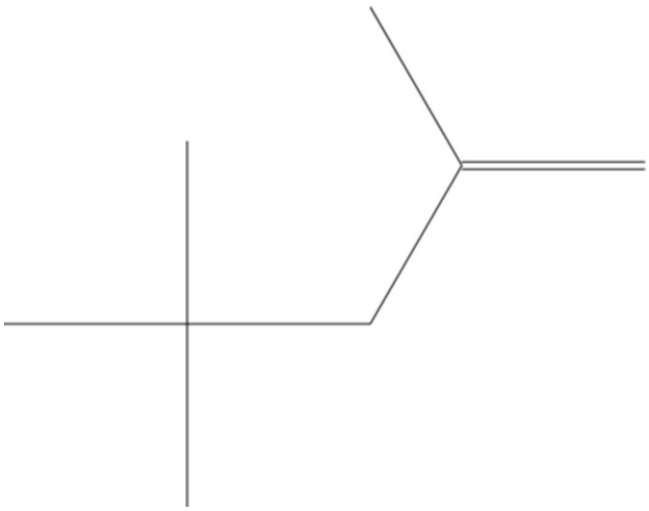

Fig. A.4. Structure of 2,4,4-trimethyl-1-pentene.

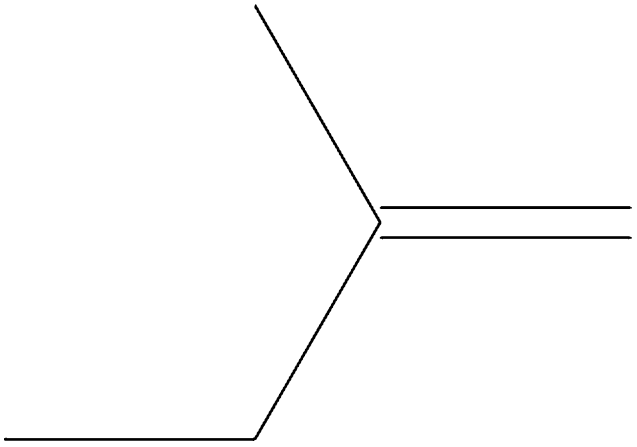

Fig. A.5. Structure of 2-methyl-1-butene.

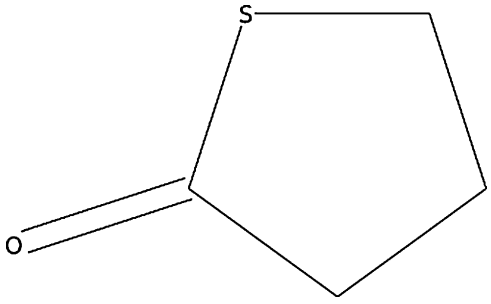

Fig. A.6. Structure of gamma-thiobutyrolactone.

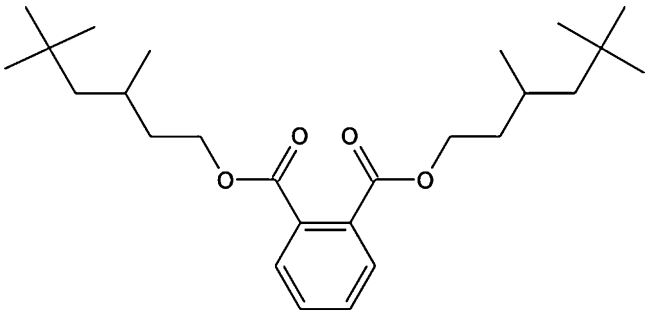

Fig. A.7. Structure of diisononyl phthalate.

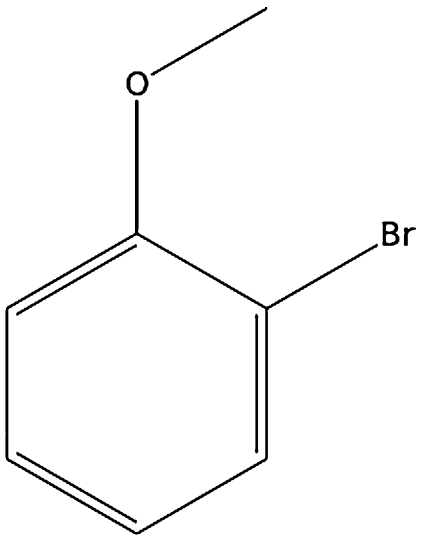

Fig. A.8. Structure of o-bromoanisole.

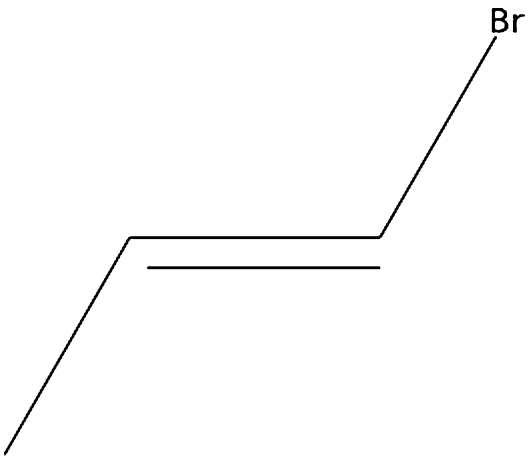

Fig. A.9. Structure of 1-bromo-propene.

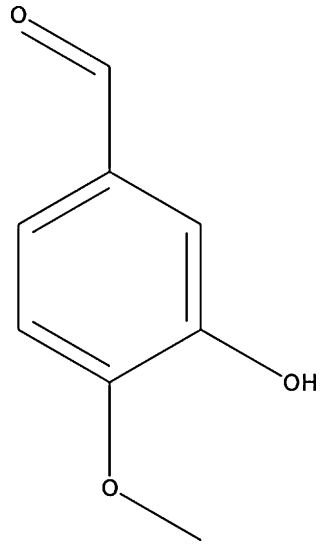

Fig. A.10. Structure of vanillin.

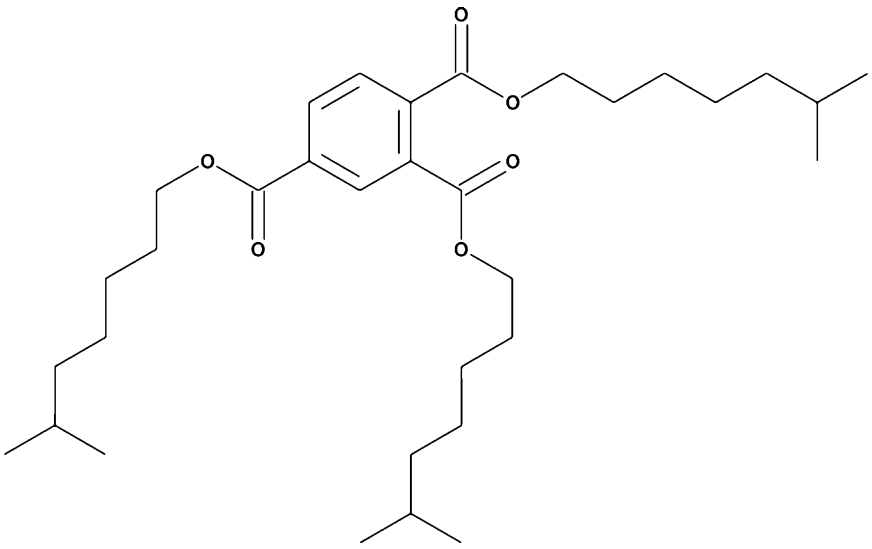

Fig. A.11. Structure of triisooctyl trimellitate.

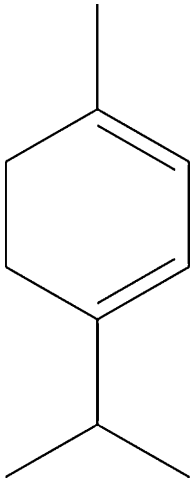

Fig. A.12. Structure of alpha-terpinene.

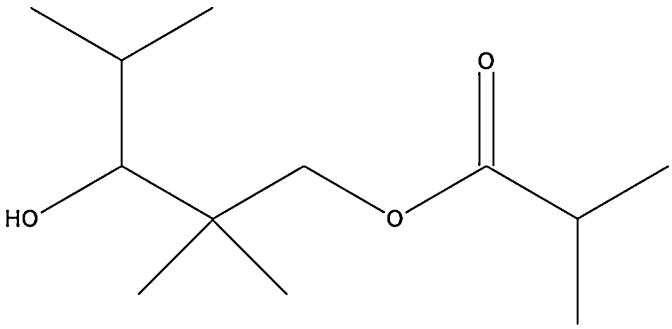

Fig. A.13. Structure of 2,2,4-trimethyl-1,3-pentanediol monoisobutyrate.

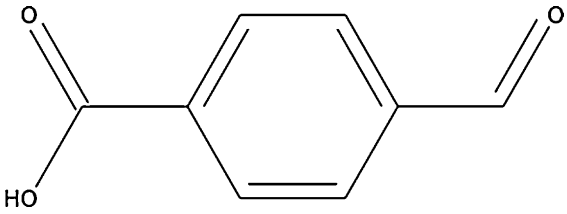

Fig. A.14. Structure of 4-carboxybenzaldehyde.

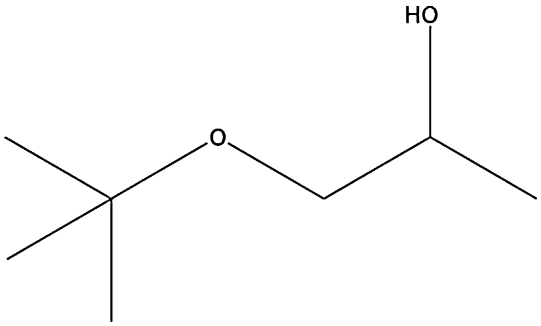

Fig. A.15. Structure of propylene glycol-tert-butyl ether.

10. 2,2,4-trimethyl-1,3-pentanediol monoisobutyrate (Fig. A.13): $(CH_3)_2$–CH– (2 occurrences) and >CHOH (1 occurrence).
11. 4-carboxybenzaldehyde (Fig. A.14): ACCOOH (1 occurrence) and ACHO (1 occurrence).
12. propylene glycol-tert-butyl ether (Fig. A.15): $(CH_3)_3$–C– (1 occurrence) and >CHOH (1 occurrence).
13. 1-bromonaphthalene (Fig. A.16): ACBr (1 occurrence) and $AC(ACH_m)_2AC(ACH_n)_2$ (1 occurrence).
14. p-diisopropylbenzene hydroperoxide (Fig. A.17): $(CH_3)_2$–CH– (1 occurrence) and CH_n–O–OH (1 occurrence).
15. di-t-butyl peroxide (Fig. A.18): $(CH_3)_3$–C– (2 occurrences) and CH_m–O–O–CH_n (1 occurrence).
16. anethole (Fig. A.19): CH_3–C= (1 occurrence) and AC–O–C (1 occurrence).
17. diisodecyl phthalate (Fig. A.20): $(CH_3)_2$–CH– (2 occurrences) and ACCOO (2 occurrences).
18. allyl methacrylate (Fig. A.21): CH_3–C= (1 occurrence) and $–CH_2–C=$ (1 occurrence).
19. 2,4,6-trimethylpyridine (Fig. A.22): $C_{cyclic}H_m{=}N_{cyclic}{-}C_{cyclic}H_n{=}C_{cyclic}H_p$ (1 occurrence).
20. methylcyclopentadiene (Fig. A.23): ring of 5 carbons (1 occurrence) and –C=C–C=C– (1 occurrence).

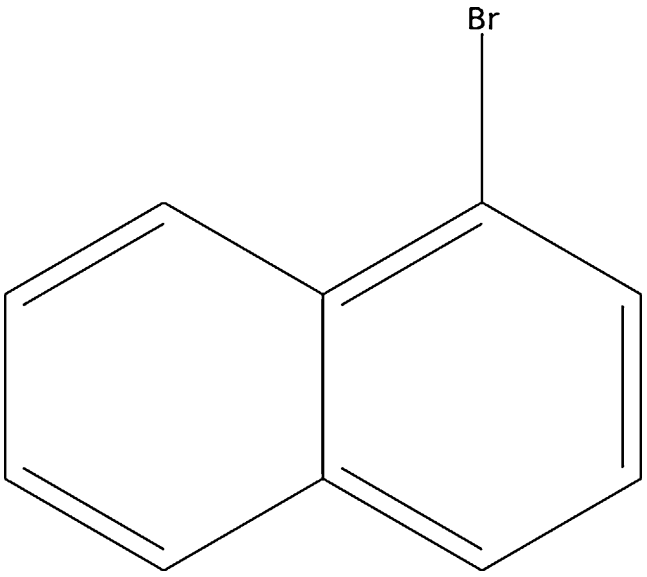

Fig. A.16. Structure of 1-bromonaphthalene.

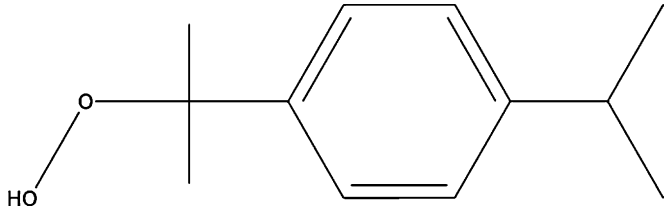

Fig. A.17. Structure of p-diisopropylbenzene hydroperoxide.

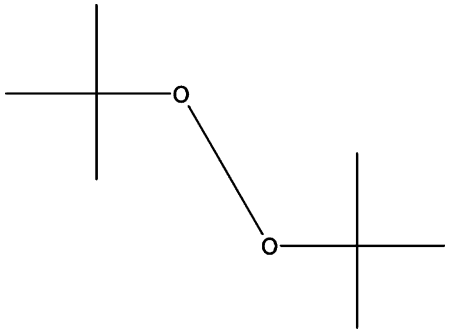

Fig. A.18. Structure of di-t-butyl peroxide.

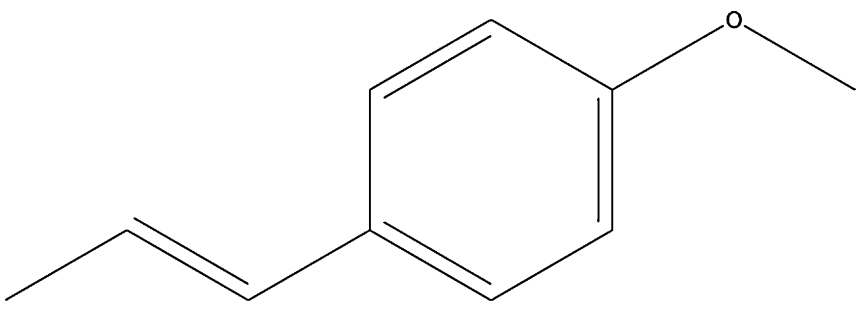

Fig. A.19. Structure of anethole.

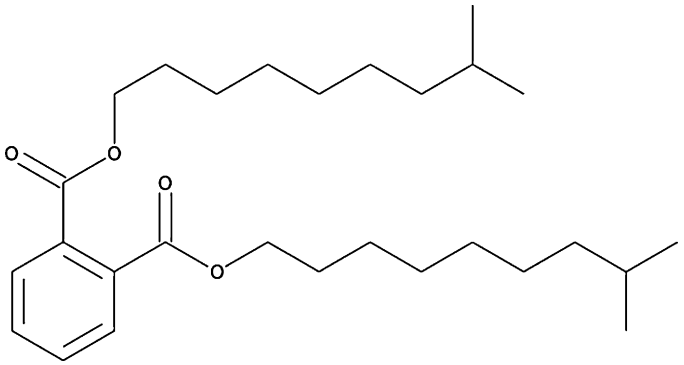

Fig. A.20. Structure of diisodecyl phthalate.

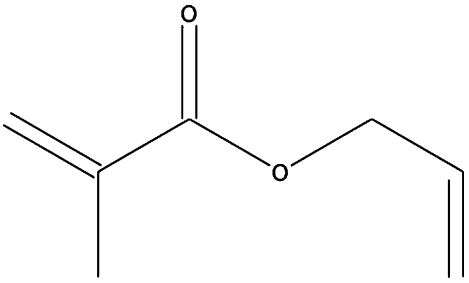

Fig. A.21. Structure of allyl methacrylate.

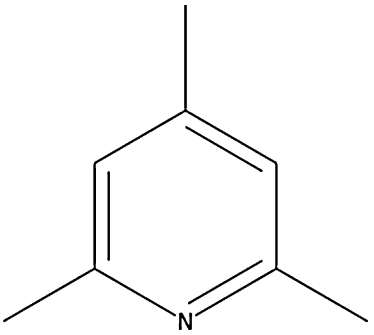

Fig. A.22. Structure of 2,4,6-trimethylpyridine.

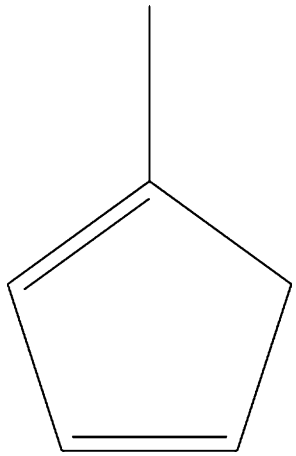

Fig. A.23. Structure of methylcyclopentadiene.

References

Abbott, S., Hansen, C.M., 2010. Hansen Solubility Parameters in Practice, Complete with Software, Data and Examples, 3rd Edition-version 3.1. Book and Software published by Hansen-Solubility.com.

Abraham, M., 1993. Scales of solute hydrogen-bonding: their construction and application to physicochemical and biochemical processes. Chem. Soc. Rev. 22, 73–83.

Ahlrichs, R., Bär, M., Häser, M., Horn, H., Kölmel, C., 1989. Electronic structure calculations on workstation computers: the program system TURBOMOLE. Chem. Phys. Lett. 162, 165–169.

Barra, J., Lescure, F., Doelker, E., Bustamante, P., 1997. The expanded Hansen approach to solubility parameters. paracetamol and citric acid in individual solvents. J. Pharm. Pharmacol. 49, 644–651.

Barton, A.F.M., 1983. Handbook of Solubility Parameters and Other Cohesion Parameters. CRC Press, Boca Raton, FL.

Barton, A.F.M., 1985. Applications of solubility parameters and other cohesion parameters in polymer science and technology. Pure Appl. Chem. 57, 905–912.

Beerbower, A., Martin, A., Wu, P.L., 1984. Expanded solubility parameter approach. I: naphthalene and benzoic acid in individual solvents. J. Pharm. Sci. 73, 179–188.

Bustamante, P., Peña, M.A., Barra, J., 1998. Partial solubility parameters of piroxicam and niflumic acid. Int. J. Pharm. 174, 141–150.

Constantinou, L., Prickett, S.E., Mavrovouniotis, M.L., 1993. Estimation of thermodynamic and physical properties of acyclic hydrocarbons using the ABC approach and conjugation operators. Ind. Eng. Chem. Res. 32, 1734–1746.

Constantinou, L., Gani, R., 1994. New group contribution method for estimating properties of pure compounds. AIChE J. 40, 1697–1710.

COSMObase Ver.C2.1 Rev.01.06, 2006. COSMOlogic GmbH & Co., K.G., Leverkusen, Germany.

Fredenslund, A., Gmehling, J., Rasmussen, P., 1977. Vapor Liquid Equilibria using UNIFAC. Elsevier Scientific, Amsterdam, Netherlands.

Hansen, C.M., 1967. The three dimensional solubility parameter—key to paint component affinities: I. Solvents, plasticizers, polymers, and resins. J. Paint Technol. 39, 104–117.

Hansen, C.M., 2004. Aspects of solubility, surfaces and diffusion in polymers. Prog. Org. Coat. 51, 55–66.

Hansen, C.M., 2007. Hansen Solubility Parameters: A User's Handbook. CRC Press, Boca Raton, FL.

Hildebrand, J., Scott, R.L., 1962. Regular Solutions. Prentice Hall, Englewood Cliffs, NJ.

Horvath, A.L., 1992. Molecular Design. Elsevier, Amsterdam, Netherlands.

Ikeda, H., Chiba, K., Kanou, A., Hirayama, N., 2005. Prediction of solubility of drug by conductor-like screening model for real solvents. Chem. Pharm. Bull. 53, 253–255.

Járvás, G., Quellet, C., Dallos, A., 2011. Estimation of Hansen solubility parameters using multivariate nonlinear QSPR modeling with COSMO screening charge density moments. Fluid Phase Equilib. 309, 8–14.

Jensen, W.B., 1987. The relevance of Lewis Acid–Base chemistry to surface interactions. In: Mittal, K.L. (Ed.), Surface and Colloid Science in Computer Technology. Plenum Press, New York, pp. 27–59.

Joback, K.G., Reid, R.C., 1987. Estimation of pure-component properties from group contributions. Chem. Eng. Commun. 57, 233–243.

Joris, L., Mitsky, J., Taft, R.W., 1972. Effects of polar aprotic solvents on linear free-energy relations in hydrogen-bonded complex formation. J. Am. Chem. Soc. 94, 3438–3442.

Jouyban, A., 2010. Handbook of Solubility Data for Pharmaceuticals. CRC Press, Boca Raton, Florida.

Kamlet, M.J., Taft, R.W., 1976. The solvatochromic comparison method. I. The.beta. -scale of solvent hydrogen-bond acceptor (HBA) basicities. J. Am. Chem. Soc. 98, 377–383.

Kamlet, M.J., Abboud, J.L.M., Taft, R.W., 1981. An examination of linear solvation energy relationships. In: Taft, R.W. (Ed.), Progress in Physical Organic Chemistry. Wiley, New York, pp. 485–630.

Karger, B.L., Snyder, L.R., Eon, C., 1976. An expanded solubility parameter treatment for classification and use of chromatographic solvents and adsorbents: parameters for dispersion, dipole and hydrogen bonding interactions. J. Chromatogr. 125, 71–88.

Karger, B.L., Snyder, L.R., Eon, C., 1978. Expanded solubility parameter treatment for classification and use of chromatographic solvents and adsorbents. Anal. Chem. 50, 2126–2136.

Klamt, A., 2005. COSMO-RS from Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design. Elsevier, Amsterdam, Netherlands.

Marrero, J., Gani, R., 2001. Group contribution based estimation of pure component properties. Fluid Phase Equilibr. 183–184, 183–208.

Mavrovouniotis, M.L., 1990. Estimation of properties from conjugate forms of molecular structures: the ABC approach. Ind. Eng. Chem. Res. 29, 1943–1953.

Panayiotou, C., 2011. Towards a COSMO equation-of-state model of fluids and their mixtures. Pure Appl. Chem. 83, 1221–1242.

Raevsky, O.A., 1987. QSAR description of molecular structure. In: Hadzi, D., Jerman-Blazic, B. (Eds.), QSAR in Drug Design and Toxicology. Elsevier, Amsterdam, Netherlands.

Sherry, A.D., Purcell, K.F., 1972. Linear enthalpy-spectral shift correlation for perfluoro-tert-butyl alcohol. J. Am. Chem. Soc. 94, 1853–1857.

Stefanis, E., Constantinou, L., Panayiotou, C., 2004. A group-contribution method for predicting pure component properties of biochemical and safety interest. Ind. Eng. Chem. Res. 43, 6253–6261.

Stefanis, E., Panayiotou, C., 2008. Prediction of Hansen solubility parameters with a New Group-Contribution Method. Int. J. Thermophys. 29, 568–585.

Stefanis, E., Tsivintzelis, I., Panayiotou, C., 2006. The partial solubility parameters: an equation-of-state approach. Fluid Phase Equilibr. 240, 144–154.

Stefanis, E., Constantinou, L., Tsivintzelis, I., Panayiotou, C., 2005. New group-contribution method for predicting temperature-dependent properties of pure organic compounds. Int. J. Thermophys. 26, 1369–1388.

Taft, R.W., Kamlet, M.J., 1976. The solvatochromic comparison method. 2. The. alpha. -scale of solvent hydrogen-bond donor (HBD) acidities. J. Am. Chem. Soc. 98, 2886–2894.

Tehrani, J., February 1993. Successful supercritical fluid extraction strategies. Am. Lab., 40hh–40mm.

Verheyen, S., Augustijns, P., Kinget, R., Van den Mooter, G., 2001. Determination of partial solubility parameters of five benzodiazepines in individual solvents. Int. J. Pharm. 228, 199–207.

VT Sigma Profile Databases, 2006. Virginia Polytechnic Institute and State University, Blacksburg, Virginia. Available at http://www.design.che.vt.edu/VTDatabases.html (accessed: 26 November 2011).