

Chapter 1

In Silico Search for Alternative Green Solvents

Laurianne Moity, Morgan Durand, Adrien Benazzouz, Valérie Molinier,
and Jean-Marie Aubry

Abstract The selection of the most appropriate alternative solvents requires efficient predictive tools that avoid resorting to time-consuming trial and error experiments. Several classifications of organic solvents exist but they most often require the knowledge of one or more experimental characteristics, which might be an obstacle in the case of emerging candidates. This chapter gives an overview of existing tools for the characterisation and classification of organic solvents and particular attention is given to purely predictive methods, such as the CONductor-like Screening MOdel for Real Solvents (COSMO-RS). A panorama of the currently available sustainable solvents is given, and these “green” alternatives are compared to the classical organic solvents, thanks to a completely in silico approach. Examples of substitutions are given to illustrate the methodology that can also be used to design new alternatives.

1.1 Tools for Solvent Selection

Solvents play an important role in a great number of unit operations in chemistry and chemical engineering. Resorting to solvents is usually required during a limited period of time during the process since they are most often expected to play a role of dissolvent, diluent, dispersant or extractant and should be removed afterwards. Nevertheless, the right choice of solvent is crucial, and through the ages, several methods for solvent selection have been developed. In former times, the choice of the most appropriate solvent was purely empirical and was often made through trial and error experiments and from empirical knowledge. This traditional approach to

L. Moity • M. Durand • A. Benazzouz • V. Molinier • J.-M. Aubry (✉)
EA 4478 Chimie Moléculaire et Formulation, University of Lille, USTL, ENSCL,
F59652 Villeneuve d’Ascq, France
e-mail: laurianne.moity@hotmail.fr; durand_morgan@yahoo.fr; adrienbenazzouz@aol.com;
Valerie.Molinier@univ-lille1.fr; Jean-Marie.Aubry@univ-lille1.fr

select a solvent usually followed the alchemist maxim *similia similibus solvuntur* that is still underlying popular contemporary approaches in use. Solvent effects were then related to the chemical structures, and several descriptors have been proposed to describe them.

1.1.1 Solvents Descriptors and Classifications

Organic chemists traditionally classify solvents as non-polar, aprotic polar and protic polar, according to their molecular structure and ability to establish hydrogen bonding. To refine this classification, solvent effects can be related to various kinds of descriptors that have evolved over time.

Initially, the only available quantitative descriptors were *physical*: enthalpy of vaporisation, dielectric constant, refractive index, boiling point, etc. However, quantifying solvent effects using physical descriptors demonstrated a moderate predictive power because these descriptors describe properly the bulk but neglect specific intermolecular interactions that are of the utmost importance whenever a second compound is added to the solvent. Therefore two types of solute-solvent interactions occur: the non-specific interactions (Van der Waals and the ion/dipole forces) and the specific interactions (hydrogen bond donor and/or hydrogen bond acceptor, electron pair donor/electron pair acceptor and solvophobic interactions) [1].

To assess these intermolecular forces, a solvent must be considered as a discontinuum, in which solvent molecules interact with each other or with the solute. For this purpose, well-chosen solutes, with a particular and quantifiable sensitivity to solvent effects, were used and allowed to access *empirical* descriptors that led to the emergence of numerous and useful empirical polarity scales, either uniparametric such as the ET(30) of Reichardt [1] or multi-parametric such as the solvatochromic parameters of Kamlet and Taft [2–4] or the Abraham parameters [5]. The last two approaches have been rationalised under the concept of linear solvation energy relationships (LSER) [6]. Recently, more than 180 polarity scales have been reviewed [7]. In the next section, the Hildebrand and Hansen solubility parameters will be emphasised since they are widely used in industry to compare and select solvents for various applications.

Over the past decade, purely *theoretical* descriptors have been introduced. They offer several advantages, the most important are being that they are easy to generate for any solvent and do not require any experiments. Different theoretical alternatives have been introduced as reviewed by Murray et al. [8]. Politzer and co-worker used electrostatic potentials computed on molecular surfaces to generate theoretical descriptors [9] that were found to be highly correlated to Kamlet and Taft solvatochromic parameters [10]. More recently, Katritzky et al. built QSPR (Quantitative Structure Property Relationship) models to predict 127 polarity scales based on theoretical descriptors. They carried out principal component analysis (PCA) of 100 solvent scales based on 703 solvents [11]. Relying on this extensive work, the authors emphasised that almost all theoretical descriptors can be related to one of the generally accepted types of intermolecular interactions.

1.1.2 Hansen Approach

The Hansen approach provides *empirical* descriptors, as presented above. At first, the solubility parameter δ_H was introduced by Hildebrand and Scott [12] and was defined as the square root of the cohesive energy density, correlated to the enthalpy of vaporisation ΔH_{vap} and to the molar volume, V (Eq. 1.1). As the difference between solute and solvent solubility parameters decreases, the tendency towards solubilisation increases.

$$\delta_H = \sqrt{\frac{\Delta H_{\text{vap}} - RT}{V}} \quad (1.1)$$

The Hildebrand parameter was extended by Hansen by splitting it into three components called the Hansen solubility parameters (Eq. 1.2). They correspond to the three main molecular interactions, namely, dispersive (δ_d), polar (δ_p) and hydrogen-bonding contributions (δ_h).

$$\delta_H = \sqrt{\delta_d^2 + \delta_p^2 + \delta_h^2} \quad (1.2)$$

The partial parameter for dispersive interactions, δ_d , is obtained from corresponding state principles, by considering the so-called homomorph of the molecule, while δ_p is derived from the ratio of the dipolar moment and the square root of the molar volume. The hydrogen-bonding contribution is calculated as the subtraction of δ_d and δ_p to the Hildebrand parameter. The three Hansen parameters can also be totally predicted by various group contribution methods such as the thermodynamically consistent model of Stefanis and Panayiotou [13]. Alternatively, the Hansen solubility parameters can be experimentally determined by individually mixing the solute in a ratio 1:10 to a proper set of solvents having a wide range of solubility parameters [14]. After 24 h of stirring at room temperature, the solubility is visually evaluated by a score ranging from 1 (soluble) to 6 (non-soluble). These scores are computed with a quality-to-fit function in order to build a solubility domain [15].

The three Hansen solubility parameters define a three-dimensional space, known as the Hansen space, in which all solvents and solutes can be located. A solute can be visualised as a point surrounded by its solubility sphere. All solvents and mixtures located inside this volume are likely to solubilise the solute. The closer the solute and solvent parameters are, the better the solubility is [16]. The solute-solvent distance, D , is defined according to Eq. 1.3:

$$D = \sqrt{4(\delta_{d,\text{solvent}} - \delta_{d,\text{solute}})^2 + (\delta_{p,\text{solvent}} - \delta_{p,\text{solute}})^2 + (\delta_{h,\text{solvent}} - \delta_{h,\text{solute}})^2} \quad (1.3)$$

The ratio between the distance D and the radius R of the solubility sphere is called the ‘‘Relative Energy Difference’’ (*RED*) – see Eq. 1.4 – and allows a fast

screening of molecules. $RED < 1$ indicates that a molecule is inside the sphere and is likely to have a high affinity with the solute while higher values of RED indicate a poor affinity:

$$RED = \frac{D}{R} \quad (1.4)$$

The semi-empirical Hansen approach demonstrated its ability to correlate and predict the behaviour of solvents. It provides reasonable results for the description of molecular and macromolecular solubility and is thus a useful tool for various industrial applications ranging from polymer processing [14] to coatings [17] and cosmetics [16].

1.1.3 COSMO-RS Approach

The significant improvement in computational power and the sophistication of recent algorithms led to the possibility to use extensively *quantum* descriptors of the solvent effect. Cartier et al. showed that quantum chemistry provides a more accurate and more detailed description of electronic effects than empirical methods [18]. Thanks to a combination of a dielectric continuum solvation model and a thermodynamic treatment of the molecular interactions, Klamt developed a general approach in which a solvent can be treated in the liquid state. In the first step, the CONductor-like Screening MODEL (COSMO) [19], the solute molecule is considered to be embedded in a cavity that is surrounded by a virtual conductor. The CONductor-like Screening MODEL for Real Solvent (COSMO-RS) then allows the transfer from the state of the molecule embedded in a virtual conductor to a real solvent [20]. COSMO-RS has already been successfully used for the prediction or the modelisation of various properties in solution, as partition coefficients (for instance, octanol-water [21] and blood brain [22]), pKa [23] or solubilisation of cosmetic ingredients [16].

In a recent work [24], we have evaluated the potentialities of the COSMO-RS approach to generate quantum descriptors for an a priori classification of solvents. The descriptors obtained from COSMO-RS were treated by principal component analysis coupled with a clustering procedure to provide a classification of solvents. This a priori classification was compared to the one of Chastrette [25], who first proposed a classification of solvents by resorting to a multi-parametric statistical approach based on a selection of six physical descriptors – boiling point, molecular dipole moment, molecular refraction, index of refraction, Hildebrand solubility parameter and Kirkwood function – in conjunction with two microscopic quantum descriptors (HOMO and LUMO energies).

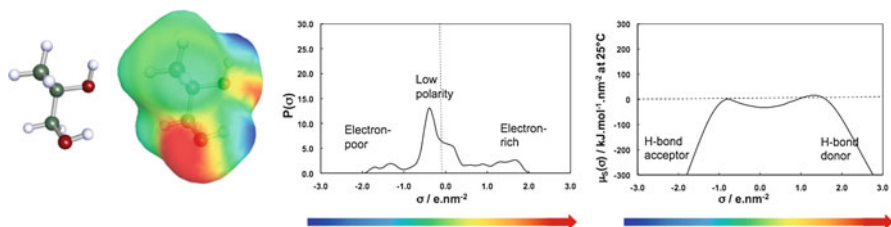


Fig. 1.1 σ -surface, σ -profile and σ -potential of 1,2-propanediol

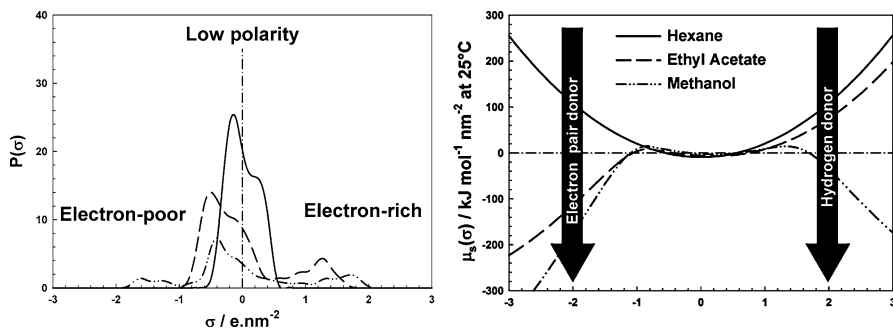


Fig. 1.2 σ -profile ($P(\sigma)$, left) and σ -potential ($\mu_S(\sigma)$, right) of three typical solvents: *n*-hexane (apolar), ethyl acetate (hydrogen bond acceptor) and methanol (amphiprotic) (Adapted from Ref. [24])

After DFT/COSMO geometry optimisations, COSMO surfaces can be generated via COSMOtherm. An example of COSMO surface is given in Fig. 1.1 for 1,2-propanediol. In the σ -surface representation, green to yellow codes the weakly polar surfaces, blue represents electron-deficient regions (δ^+) and red codes electron-rich regions (δ^-). This 3D information on the repartition of charge density on the molecular surface can be reduced to a histogram $P(\sigma)$ that expresses the redundancy of a surface density in a polarity interval. Such histograms have been defined as σ -profiles [21]. In the framework of COSMO-RS, it is also possible to generate the so-called σ -potential plots. This plot represents the chemical potential $\mu_S(\sigma)$ of a molecular surface fragment in a solvent S as a function of the polarisation charge density of this surface fragment (ranging from -3 to 3 e.nm $^{-2}$). This representation is of particular interest as it underlines the affinity of solvent S for a polarity of kind σ . The σ -surface, σ -profile and σ -potential of 1,2-propanediol are presented in Fig. 1.1.

Figure 1.2 shows the σ -profile and σ -potential of typical solvents. In the case of apolar solvents, exemplified by *n*-hexane, the σ -profile exhibits only one large shouldered peak centred close to 0, with a maximum at -0.1 e.nm $^{-2}$ corresponding

to the protons. Since no hydrogen bond interaction can occur with a solute surface, the corresponding σ -potential curve $\mu_S(\sigma)$ exhibits a U shape that is typical of apolar solvents: the contact between *n*-hexane and the molecular surface of a solute with a positive or negative charge density distant from 0 e.nm^{-2} will be energetically unfavourable ($\mu_S(\sigma) > 0 \text{ kJ.nm}^{-2}$). Methanol is a typical example of an amphiprotic solvent. Its sp^3 -oxygen induces at the same time a hydrogen acceptor and a hydrogen donor character. Consequently, on the σ -profile $P(\sigma)$, two secondary maxima can be observed beside the central peak corresponding to the carbon and hydrogens of the methyl group. One maximum, in the negative σ , corresponds to the hydrogen bond donor character, and the other one, in the positive σ , corresponds to the hydrogen bond acceptor character. Therefore, the σ -potential curve $\mu_S(\sigma)$ will be the opposite of that observed for apolar solvents in the regions distant from 0 e.nm^{-2} : negative, i.e. energetically favourable, $\mu_S(\sigma)$ are obtained for both negative and positive regions of a solute molecule, leading to a \cap -shaped curve. Ethyl acetate possesses only one hydrogen acceptor group due to its sp^2 -oxygen. With the same analysis, it is easy to understand why its σ -potential curve is S-shaped, i.e. negative for $\sigma < 0$ and positive for $\sigma > 0$.

This is a purely qualitative interpretation of σ -profiles and σ -potentials. In actual fact, much more quantitative information is enclosed in these curves, and therefore we have used them to extract theoretical molecular descriptors of the solubilising properties further used as input parameters for solvent classification.

To attain a quantitative comparison of the σ -potential plots, 61 discrete values of $\mu_S(\sigma)$ given by COSMOtherm can be extracted for every 0.1 e.nm^{-2} increment within the interval -3 to 3 e.nm^{-2} . It was performed for the 153 solvents of the chosen dataset (the one of Chastrette [25]), and these 61 points were used to give a description of the solubilising properties of each solvent and thus make up a set of 61 descriptors. This set could be reduced by PCA to a smaller number of relevant descriptors since most of them contain redundant information. In our case, the vector space could be reduced to only four eigenvectors, still accounting for 96.4 % of the variance. By neglecting the fourth eigenvector, more than 85 % of the variance is still expressed and all solvents can then be positioned in a pseudo-3D space (F1, F2, F3) (see next paragraph).

A clustering procedure allowed gathering the 153 traditional solvents into ten classes for which the sigma profiles and sigma potentials are presented in Fig. 1.3. The description of these clusters will be discussed in the next paragraph. They are in good agreement with the ten classes defined by Chastrette, and they even allow a more accurate positioning of solvents that were mispositioned in this original work [24].

The σ -potentials derived from the COSMO-RS theory can thus be successfully employed to describe and classify solvents in a purely predictive manner, with a good consideration of hydrogen bond donor/acceptor interactions. This approach is of particular interest in the context of solvent design and will be addressed in Sect. 1.4.

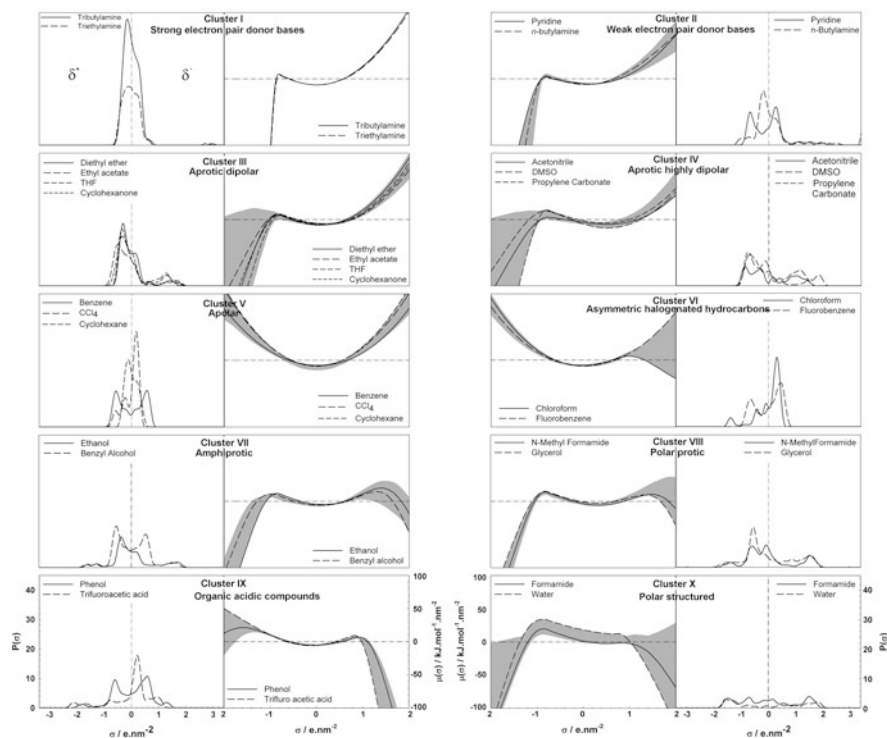


Fig. 1.3 σ -profile and σ -potential of typical solvents of the ten clusters. The grey regions show the dispersion of the σ -potential curves within each cluster (Adapted from Ref. [24])

1.2 Panorama of Current “Green Solvents”

1.2.1 Classes of “Green Solvents”

The adjective “sustainable” or “green” is used to describe different types of solvents including the ones that are produced from biomass feedstock and eco-friendly petrochemical-based solvents that are non-toxic and/or biodegradable. Figure 1.4 shows the different families of solvents that are generally considered as “green”. It is worth noticing that the greenness of some solvents is questionable with regard to toxicity (e.g. ionic liquids) or biodegradability (fluorinated solvents and silicones).

The family of “eco-friendly” solvents is the most heteroclite one, since it gathers all kinds of solvents with a good EHS (Environment, Human, Safety) profile. These solvents may also be obtained by the valorisation of industrial by-products, as is the case for the dimethyl, diethyl and dibutyl esters of glutaric, succinic and adipic acid, and by-products of the nylon 6,6 manufacture (the so-called dibasic esters). Another example of an “eco-friendly” solvent is 3-methoxy-3-methyl-butan-1-ol (MMB)

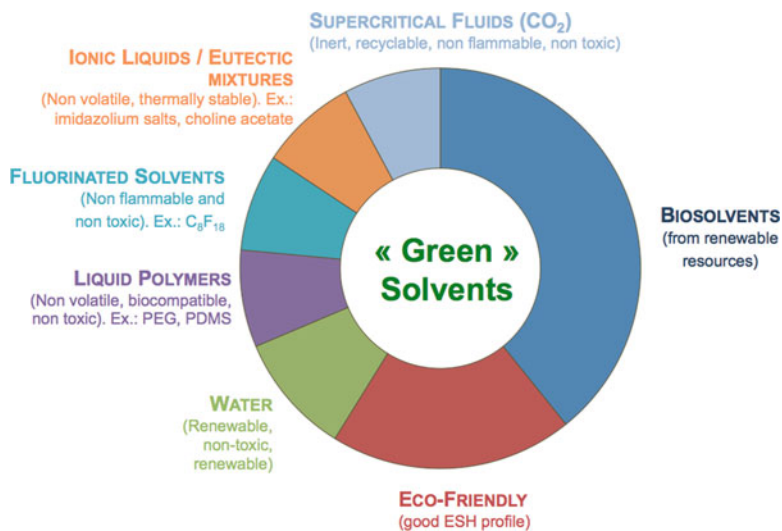


Fig. 1.4 The seven classes of solvents generally claimed as “green” solvents [26] (EHS Environment, Human, Safety)

which is a non-toxic and non-VOC solvent used in air freshener, household and industrial cleaner formulations. For the same reasons, the alkyl alcanolamides are considered as “eco-friendly” solvents, as well as some alkanes or dimethylsulfoxide.

Biosolvents mainly belong to three chemical families, namely, esters, alcohols and terpenes. These bio-based solvents are obtained by chemical or biochemical transformations of agro-synthons, i.e. defined molecules obtained from the biomass feedstock. Grains and oleaginous plants contain vegetable oils, which are converted to glycerol and fatty acids. They themselves are a source of solvents, giving rise to glycerol carbonate, glycerol triacetate, or vegetable oil methyl esters. The “sugar” platform (cellulose, hemicellulose, starch, sucrose) is the source of simple sugars and polyols that can be further transformed, chemically or enzymatically, to solvents. It should be stressed that, contrary to the “eco-friendly” family, all bio-based solvents do not have a good EHS profile. For instance, furfuraldehyde is a solvent readily obtained from various plant wastes, as corn stalk or sawdust, by acidic hydrolysis of hemicellulose into pentosidic units that are themselves dehydrated. This biosolvent is both toxic and carcinogenic. A first hydrogenation gives furfuryl alcohol that is also classified among the CMR (Carcinogenic Mutagenic Reprotoxic, in the European Union) substances, and a subsequent hydrogenation yields tetrahydrofurfuryl alcohol that has a good EHS profile.

Liquid polymers, such as polyethylene glycol, and silicone oils can also be considered as green solvents because of their non-volatility.

Finally, there is a growing interest for ionic liquids, i.e. salts with a melting point below 100 °C [27]. Typical ionic liquids have cations such as imidazolium, pyridinium or pyrrolidinium and anions such as hexafluorophosphate, tetrafluoroborate

or triflate. They are frequently considered as green solvents because of their non-volatility. Nevertheless, the toxicity and biodegradability of such compounds are currently questioned [28], and most common ionic liquids such as [Bmim][BF₄] (3-butyl-1-methylimidazolium tetrafluoroborate) do not show good biodegradation. Because the toxicity of ionic liquids is often associated to the cation part, choline has been investigated as a benign quaternary ammonium ion derived from renewable resources [29, 30].

In a recent work [31], we have listed 138 “green” solvents through the review of technical, commercial and scientific literature to provide a “panorama” of sustainable solvents through the COSMO-RS approach. Supercritical fluids have not been considered since they cannot be straightly modelled by COSMO-RS. This list is recalled in Table 1.1.

1.2.2 Positioning of Alternative Solvents

The list of “green” solvents presented in Table 1.1 has been studied via the COSMO-RS approach, as presented schematically in Fig. 1.5. One hundred fifty-three traditional solvents were also included in the analysis, as presented in Sect. 1.1, and PCA analysis of the sigma potentials followed by a clustering procedure further provided the ten classes presented in Table 1.2 and in Figs. 1.6 and 1.7.

Ionic liquids cannot be positioned in any of the ten families, and therefore, they should be considered as a full-fledged cluster. A1 and A3 coordinates of choline acetate, for example, are much higher than the ones of classical organic solvents.

The positioning of “green” solvents shows some distinctive features from the one of classical organic solvents. Cluster III (Aprotic dipolar) is the most populated family for green solvents, while cluster V (Apolar) is the largest one for classical solvents. Cluster III is made up of esters (dibasic esters or fatty acid esters) and ethers that are highly represented among sustainable solvents. If we take a closer look at fatty acid esters, we observe that they either belong to cluster III or to cluster V because of their particular chemical structure between alkane and aprotic dipolar molecules. Because of this duality, such compounds could be considered as a fully independent cluster. Cluster V contains few solvents and is mainly composed of terpenes, the main representative apolar solvents among “green” solvents. Clusters VII and VIII are also much populated by “green” solvents that are mainly alcohols coming from renewable resources. All other clusters are much less populated. Cluster II (weak electron pair donor bases) is only composed of 4 solvents, mainly amides. Since no amines were encountered among the “green” solvents listed, there is no solvent in cluster I. The same observation can be made for cluster VI since this cluster is mainly made up of halogenated compounds in the classical organic solvents. The lack of amines, aromatics and halogenated compounds among “green” solvents has already been noted using a similar approach based on Kamlet and Taft parameters [32].

Table 1.1 The 138 “green” solvents with their CAS registry numbers, positioned in the ten clusters as evidenced by the COSMO-RS analysis

Cluster I: Strong electron pair donor bases	
<i>No “green” solvents</i>	
Cluster II: Weak electron pair donor bases	
<i>Acetone</i>	67-64-1
N,N-Dimethyloctanamide	1118-92-9
Methyl 5-(dimethylamino) –2-methyl-oxopentanoate	1174627-68-9
<i>2-Pyrrolidone</i>	616-45-5
Cluster III: Aprotic dipolar	
<i>Acetyltributyl citrate</i>	77-90-7
<i>Benzyl benzoate</i>	120-51-4
Butyl acetate	123-86-4
Butyl laurate	106-18-3
1,4-Cineol	470-67-7
1,8-Cineol	470-82-6
Cyclopentyl methyl ether	5614-37-9
<i>Dibutyl sebacate</i>	109-43-3
Diethyl adipate	141-28-6
Diethyl glutarate	818-38-2
<i>Diethyl phthalate</i>	84-66-2
Diethyl succinate	123-25-1
Diisoamyl succinate	818-04-2
Diisobutyl adipate	141-04-8
Diisobutyl glutarate	71195-64-7
Diisobutyl succinate	925-06-4
Diisooctyl succinate	2915-57-3
Dimethyl adipate	627-93-0
Dimethyl glutarate	1119-40-0
<i>Dimethyl phthalate</i>	131-11-3
Dimethyl succinate	106-65-0
N,N-Dimethyldecanamide	14433-76-2
<i>Dimethyl isosorbide</i>	5306-85-4
Dioctyl succinate	14491-66-8
1,3-Dioxolane	646-06-0
<i>Ethyl acetate</i>	141-78-6
Ethyl laurate	106-33-2
Ethyl linoleate	544-35-4
Ethyl linolenate	1191-41-9
Ethyl myristate	124-06-1
Geranyl acetate	105-87-3
<i>Glycerol triacetate</i>	102-76-1
Glycerol-1,2,3-tributyl ether	131570-29-1
Glycerol-1,2,3-triethyl ether	162614-45-1
Glycerol-1,2,3-trimethyl ether	20637-49-4
Glycerol-1,3-dibutyl ether	2216-77-5

(continued)

Table 1.1 (continued)

Isoamyl acetate	123-92-2
Isobutyl acetate	110-19-0
Isopropylacetate	108-21-4
<i>Isopropyl myristate</i>	<i>110-27-0</i>
Isosorbide dioctanoate	64896-70-4
Methyl abietate	127-25-3
Methyl acetate	79-20-9
Methyl laurate	111-82-0
Methyl linoleate	112-63-0
Methyl linolenate	301-00-8
Methyl myristate	124-10-7
Methyl oleate	112-62-9
Methyl palmitate	112-39-0
Dimethyl 2-methyl glutarate	14035-94-0
2-Methyltetrahydrofuran	96-47-9
Menthanyl acetate	58985-18-5
n-Propyl acetate	109-60-4
Terpineol acetate	8007-35-0
<i>Tributyl citrate</i>	<i>77-94-1</i>
<i>Triethyl citrate</i>	<i>77-93-0</i>
Cluster IV: Aprotic highly dipolar	
<i>Dimethylsulfoxide</i>	<i>67-68-5</i>
2-Furfuraldehyde ^a	98-01-1
<i>Propylene carbonate</i>	<i>108-32-7</i>
γ -Valerolactone	108-29-2
Cluster V: Apolar	
Butyl myristate	110-36-1
Butyl palmitate	111-06-8
Butyl stearate	123-95-5
Cyclohexane	110-82-7
p-Cymene	99-87-6
β -Myrcene	123-35-3
<i>Decamethylcyclopentasiloxane</i>	<i>541-02-6</i>
β -Farnesene	18794-84-8
<i>Ethyl oleate</i>	<i>111-62-6</i>
Ethyl palmitate	628-97-7
<i>Isopropyl palmitate</i>	<i>142-91-6</i>
D-Limonene	5989-27-5
Methyl stearate	112-61-8
<i>Isododecane</i>	<i>31807-55-3</i>
Perfluorooctane	307-34-6
α -Pinene	80-56-8
β -Pinene	127-91-3
Terpinolene	586-62-9
Cluster VI: Asymmetric halogenated hydrocarbons (aprotic slightly dipolar)	
<i>No "green" solvents</i>	

(continued)

Table 1.1 (continued)

Cluster VII: Amphiprotic	
<i>Benzyl alcohol</i>	100-51-6
1-Butanol	71-36-3
Cyclademol	25225-09-6
1-Decanol	112-30-1
Dihydromyrcenol	18479-58-8
1,3-Dioxolane-4-methanol	5660-53-7
<i>Ethanol</i>	64-17-5
Ethylhexyl lactate	6283-86-9
<i>Ethyl lactate</i>	97-64-3
Geraniol	106-24-1
Glycerol-1,3-diethyl ether	4043-59-8
Glycerol-1,2-dibutyl ether	91337-36-9
Glycerol-1,2-diethyl ether	4756-20-1
Glycerol-1,2-dimethyl ether	40453-77-8
Glycerol-1,3-dimethyl ether	623-69-8
Glycerol-1-butyl monoether	624-52-2
Glycerol-1-ethyl monoether	1874-62-0
Glycerol-2-butyl monoether	100078-36-2
Glycerol-2-ethyl monoether	22598-16-9
Glycofurool ($n = 2$)	52814-38-7
N,N-Diethylolcapramide	136-26-5
Caprylic acid diethanolamide	3077-30-3
Isoamyl alcohol	123-51-3
<i>Isopropyl alcohol</i>	67-63-0
Methyl ricinoleate	141-24-2
Methanol	498-81-7
Nopol	128-50-7
1-Octanol	111-87-5
<i>Oleic acid</i>	112-80-1
1-Octanol	111-87-5
<i>Oleic acid</i>	112-80-1
<i>Oleyl alcohol</i>	143-28-2
Polyethylene glycol 600	25322-68-3
Solketal	100-79-8
<i>Ricinoleic acid</i>	141-22-0
α -Terpineol	98-55-5
β -Terpineol	138-87-4
Tetrahydrofurfurylic alcohol	97-99-4
Cluster VIII: Polar protic	
1,3-Dioxan-5-ol	4740-78-7
<i>1,3-Dioxolane-4-methanol</i>	5464-28-8
Ethylene glycol	107-21-1
<i>Dipropylene glycol</i>	110-98-5
Furfurylic alcohol ^a	98-00-0

(continued)

Table 1.1 (continued)

<i>Glycerol</i>	<i>56-81-5</i>
Glycerol carbonate	931-40-8
Glycerol-1-methyl monoether	623-39-2
Glycerol-2-methyl monoether	761-06-8
5-(Hydroxymethyl)furfural	67-47-0
3-Hydroxypropionic acid	503-66-2
3-Methoxy-3-methyl-1-butanol	56539-66-3
<i>Polyethylene glycol 200</i>	<i>112-60-7</i>
1,3-Propanediol	504-63-2
<i>Propylene glycol</i>	<i>57-55-6</i>
Cluster IX: Organic acidic compounds	
Acetic acid	64-19-7
Propionic acid	79-09-4
Cluster X: Polar structured	
Water	7732-18-5
Ionic liquids	
Choline acetate	14586-35-7
3-Butyl-1-methylimidazolium	174501-65-6
Tetrafluoroborate	

Adapted from Ref. [31]

Solvents coming from renewable resources (biosolvents) are indicated in *italics*. The solvents acceptable for pharmaceutical or cosmetic applications are in *italics*

^aCMR compounds

Current sustainable solvents thus mostly belong to aprotic dipolar, amphiprotic and polar protic compounds, while strong electron pair donor bases, weak electron pair donor bases and aprotic slightly dipolar (asymmetric halogenated hydrocarbons) are scarcely or not represented at all among them.

1.3 Selection of Alternative Solvents for Extraction

Because of renewed toxicology standards and exposure guidelines, ever extending lists of volatile organic compounds (VOCs), ozone depleting substances, hazardous air pollutants (HAPs, in the USA) or CMR compounds (Carcinogenic Mutagenic Reprotoxic, in the European Union), solvent substitution is not a novel concern. When the use of a solvent becomes forbidden by new regulations, effective and quick substitution solutions have to be found. In the case of extraction, chlorinated hydrocarbons and *n*-hexane are two archetypal examples of such problematic solvents. Their replacements by biosolvents have been analysed below in light of the COSMO-RS and Hansen approaches.

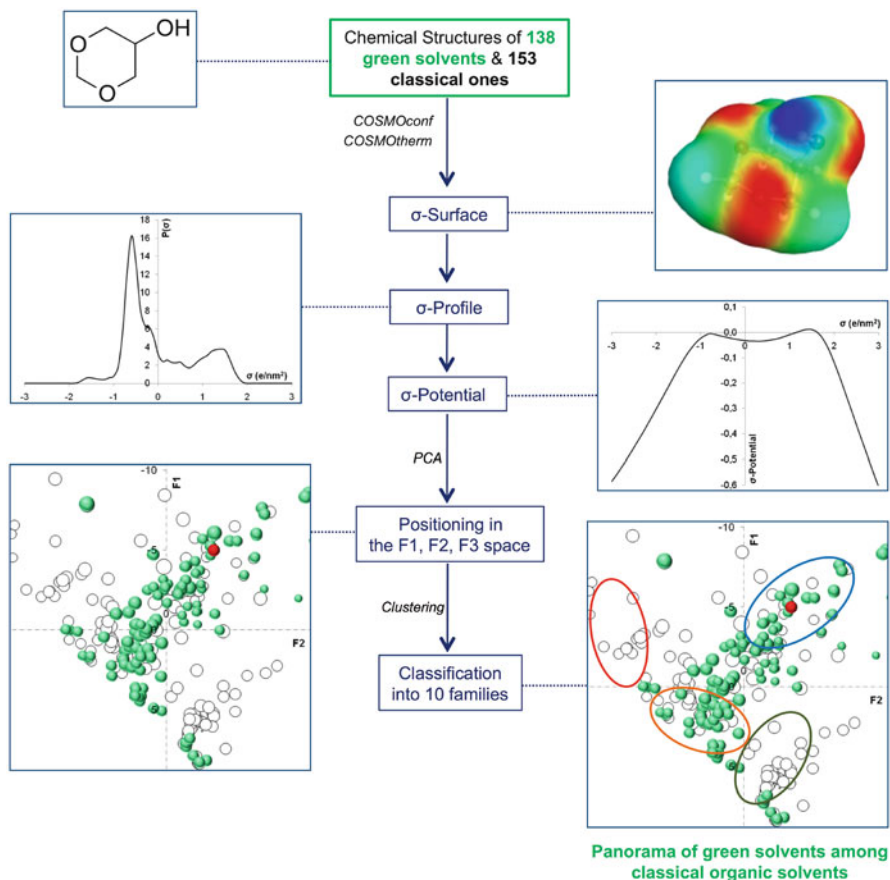


Fig. 1.5 Strategy used to position 138 green solvents (*green dots*) in the predefined 3D-space, thanks to 153 classical organic solvents (*empty dots*) using the COSMO-RS approach. The procedure is exemplified in the case of glycerol formal (1,3-Dioxolane-4-methanol) (*red dot*) (Reproduced from Ref. [31] with permission from The Royal Society of Chemistry)

1.3.1 Replacement of Chlorinated Solvents

For many years, chlorinated solvents have been in widespread use in numerous sectors, as degreasing of metallic surfaces, dry cleaning, paints (as thinner or stripper), organic synthesis and extraction among others [33]. Dichloromethane was previously used for the decaffeination of coffee which is now performed using supercritical carbon dioxide. This keen interest in chlorinated solvents is due to their outstanding physico-chemical properties, particularly their excellent solvent power, their low inflammability and high volatility. However, from a EHS point of view, chlorinated solvents exhibit a particularly bad footprint: most of them are classified among VOCs and some are blamed for stratospheric ozone depletion. Their low

Table 1.2 Clustering of “classical” and “green” solvents with typical examples in each group

Cluster	Name	Classical solvents	“Green” solvents
I	Strong electron pair donor bases	Tributylamine	–
II	Weak electron pair donor bases	Pyridine	2-Pyrrolidone
III	Aprotic dipolar	Diethyl ether	Glycerol triacetate
		Cyclohexanone	Dioxolane
IV	Aprotic highly dipolar	Sulfolane	γ -Valerolactone
		Acetonitrile	Propylene carbonate
V	Apolar	Benzene	Methyl stearate
		CCl_4	D-limonene
VI	Aprotic slightly dipolar (asymmetric halogenated hydrocarbons)	Dichloromethane	–
		Nitrobenzene	
VII	Amphiprotic	Ethanol	Isoamyl alcohol
		Benzyl alcohol	α -Terpineol
VIII	Polar protic	2-Aminoethanol	Glycerol carbonate
		Methanol	Ethylene glycol
IX	Organic acidic compounds	Phenol	Acetic acid
X	Polar structured	Water	Water
		Formamide	

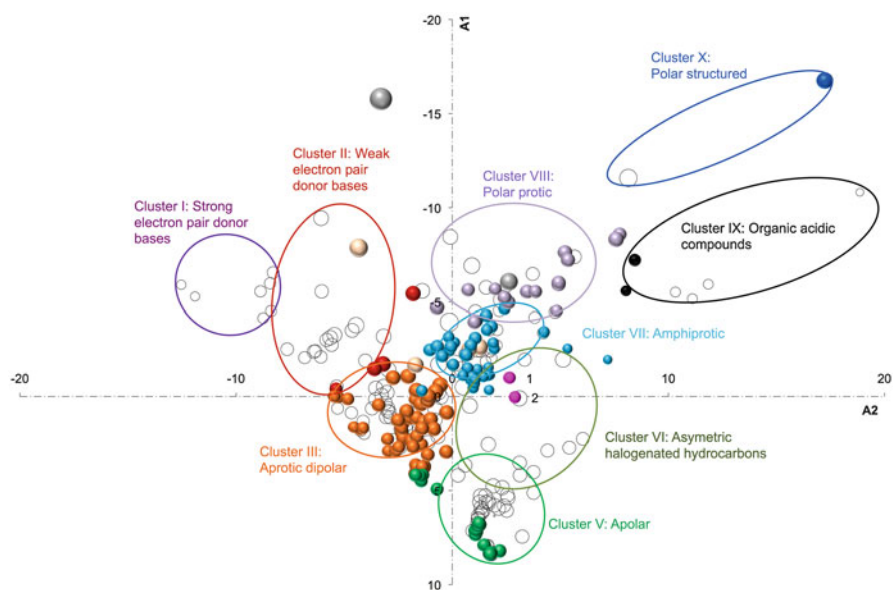


Fig. 1.6 2D – representation (A_1 vs. A_2) of green solvents (coloured circles) positioned within the clusters previously defined with classical solvents (empty circles) (Reproduced from Ref. [31] with permission from The Royal Society of Chemistry)

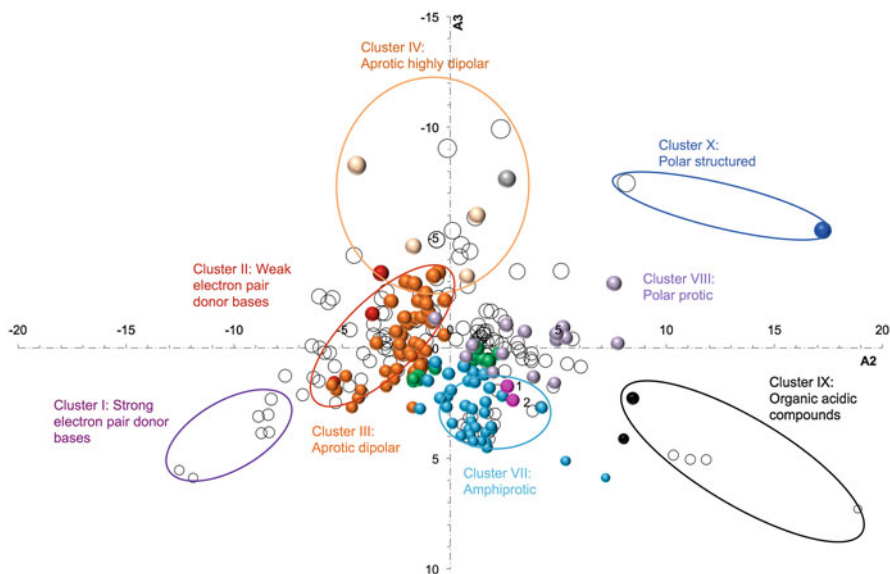
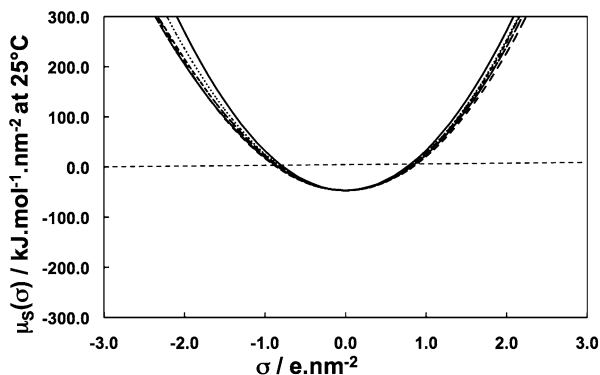


Fig. 1.7 2D – representation (A3 vs. A2) of green solvents (coloured circles) positioned within the clusters previously defined with classical solvents (empty circles) (Reproduced from Ref. [31] with permission from The Royal Society of Chemistry)

solubility in water and poor biodegradability induce long-term pollution of soil and groundwater [33]. Some of them are listed among HAP or CMR substances according to the US and/or European legislations. Therefore, for several years, substitution of chlorinated solvents has been encouraged. Among all proposed solutions, terpenes have been presented as alternatives in cleaning applications [34].

The closest neighbours of dichloromethane, tetrachloroethylene, carbon tetrachloride, 1,1,1-trichloroethane, trichloroethylene and chloroform have been looked for in the COSMO-RS classification presented before. Terpenic solvents emerge as possible substitutes for tetrachloroethylene, carbon tetrachloride and trichloroethylene (α -pinene for the first two and β -myrcene for the last one). Actually, terpenes, such as *p*-cymene, terpinolene, α -pinene or D-limonene, belong to cluster V, i.e. the cluster that contains, *inter alia*, some chlorinated solvents. The closest neighbour of 1,1,1-trichloroethane is perfluorooctane that also belongs to cluster V. This solvent has been chosen as a representative example of fluorinated solvents that are presented as “green solvents” for organic synthesis or in the electronics industry [26]. More surprisingly, the closest neighbours encountered for dichloromethane and chloroform are respectively benzyl benzoate that belongs to cluster III (aprotic dipolar) and benzyl alcohol, belonging to cluster VII (amphiprotic). Actually, chloroform and dichloromethane belong to a cluster that is not populated at all by the existing green solvents (cluster VI), which justify the current interest in the design of new bio-based solvents (see paragraph 1.4).

Fig. 1.8 σ -potential of *n*-hexane (plain line) compared to the ones of *D*-limonene, *p*-cymene, α -pinene and β -pinene (dotted lines)



1.3.2 Replacement of *n*-Hexane

n-hexane is a major solvent for the extraction of natural products and particularly vegetable oils in the food industry. It has many advantages, in particular its high solubilising capacity of oily constituents and its low boiling point, which facilitates the recovery of solutes and solvent recycling. However, it is listed among VOCs and HAPs, and in Europe, it belongs to the CMR list for its reprotoxicity and its neurotoxic metabolite, 2,5-hexanedione [14]. Substitutes to *n*-hexane for extraction are thus wanted, and terpenes are often put forward for this application. In particular, Tanzi et al. [35] have shown that terpenes could be efficiently used for the recovery of triglycerides from the algae *Chlorella vulgaris*. This substitution solution can be investigated using the COSMO-RS and Hansen approaches.

1.3.2.1 Positioning of *n*-Hexane in the COSMO-RS Panorama

As already mentioned previously for chlorinated solvents, looking for solvents having the closest σ -potentials is a way to identify potential substitutes. *n*-hexane belongs to cluster V (apolar compounds), with a typical U-shaped σ -potential showing the lack of H-bond donor and H-bond acceptor character. *D*-limonene, *p*-cymene, α -pinene and β -pinene are common terpene solvents that also belong to this cluster, as presented in Fig. 1.8, which indicates that their solubilising properties should be close.

1.3.2.2 Combined Hansen and COSMO-RS Approaches for the Substitution of *n*-Hexane

The solvents highlighted by the COSMO-RS panorama can be positioned in the Hansen space. The Hansen solubility parameters of *n*-hexane, *D*-limonene, *p*-cymene, α -pinene and β -pinene listed in references [14, 36] are positioned in

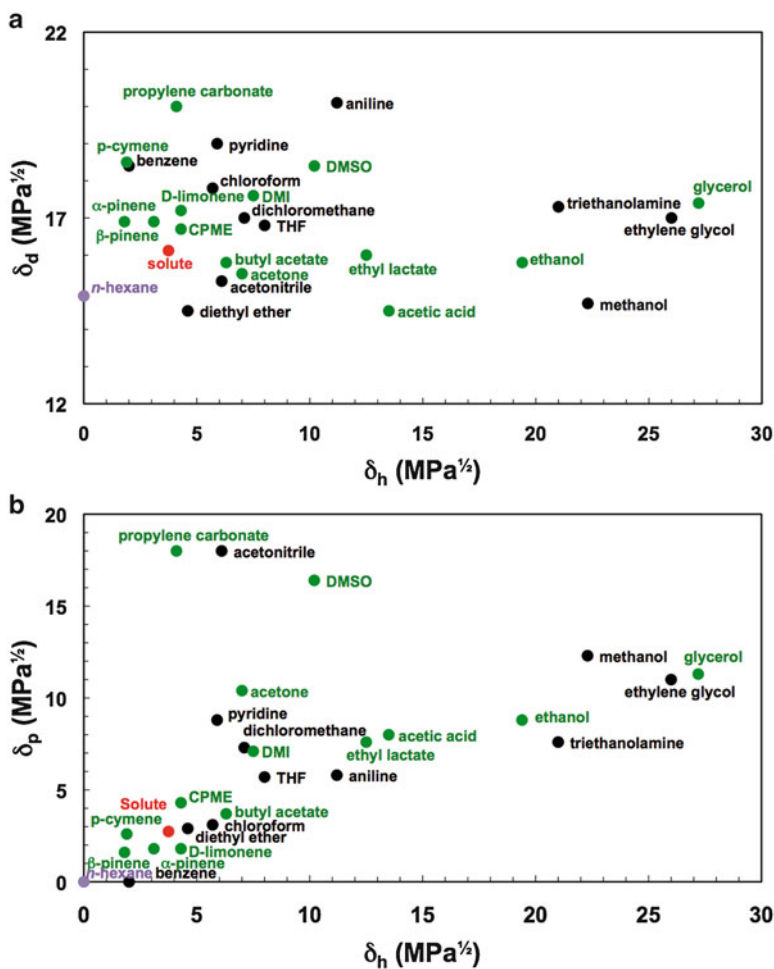


Fig. 1.9 Location of *n*-hexane, potential substitutes and fatty acid methyl esters (solute) within the (δ_p, δ_h) (top) and (δ_d, δ_h) (bottom) maps

the δ_p/δ_h and δ_d/δ_h 2-D maps in Fig. 1.9. The “solute” to extract (lipidic fraction) has been modelled by the methyl esters of palmitic, linolenic and oleic acids that are the three main fatty chains encountered in the oily fraction extracted from *Chlorella vulgaris* [35]. The position of this “solute” takes into account the relative proportions of each type of fatty acid as described in reference [35].

The close location of *n*-hexane and terpenes in the Hansen space is in good agreement with the close positioning in the COSMO-RS panorama. Other “green” solvents are found in the vicinity of *n*-hexane in the Hansen space, namely, cyclopentyl methyl ether (CPME) and *n*-butyl acetate. They both belong to cluster III (aprotic dipolar) in the COSMO-RS classification, in which solvents have an electron-donor ability (H-bond acceptor).

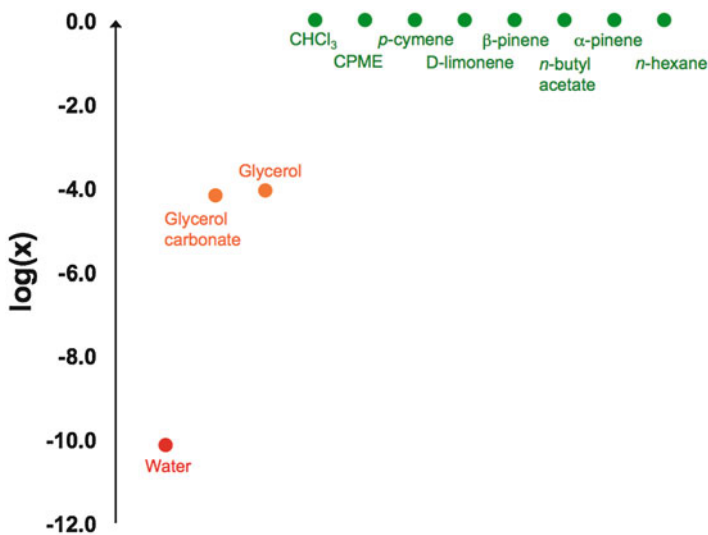


Fig. 1.10 Relative solubility of the mixture of fatty acids methyl esters considered in good (*green*), poor (*orange*) and very bad (*red*) solvents, as determined through the “solvent screening” tool of COSMOtherm

To compare the solubilising abilities of these “green” alternatives towards fatty acid methyl esters, the “solvent screening” tool implemented in COSMOtherm was used. It provides a ranking of relative solubilities from the prediction of μ^i_{solv} , the chemical potential of the compound of interest in a list of selected solvents. This solvent ranking is presented in Fig. 1.10 for the solubilisation of the mixtures of fatty acid methyl esters considered. The logarithm of solubility in mole fractions is calculated, and the logarithm of best solubility is set to 0, all other solvents being given relative to the best solvent. In the present case, *n*-hexane is the best solvent, and the “green” substitutes cannot be distinguished, they are all set at $\log(x) = 0$. For comparison purposes, the relative solubilities in water (worse solvent) and in glycerol and glycerol carbonate (poor solvents) are also given.

Finally, it is also interesting to show that the Hansen approach can be used in a very simplistic manner to identify possible efficient solvent mixtures. In the present example, Fig. 1.9 shows that α -pinene can be brought closer to the position of the target solute by addition of a solvent with higher δ_p and δ_h and lower δ_d .

Figure 1.11 shows such a procedure in the case of the addition of ethanol to α -pinene. The distance D of the solvent mixture to the solute in the Hansen space is computed from Eq. 1.3 (see Sect. 1.1.2). The addition of *ca.* 6 % of ethanol to α -pinene allows reducing the distance to the solute and is expected to enhance the solubilising capacities. This effect was observed by Tanzi et al. [35] during the recovery of triglycerides from *Chlorella vulgaris* for which the addition of a small amount of methanol to chloroform proved to increase the extraction yield.

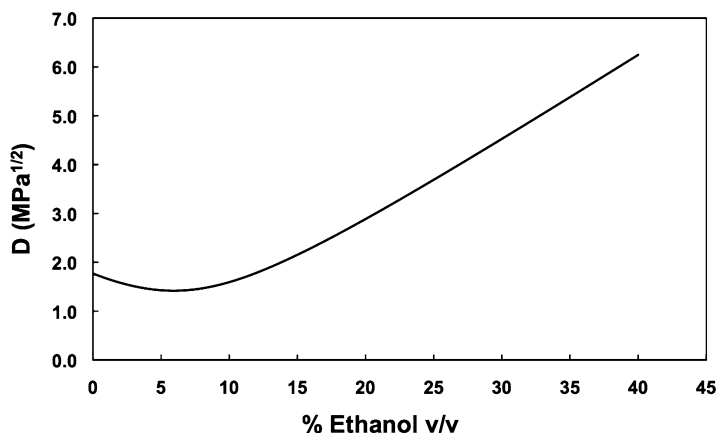


Fig. 1.11 “Synergetic effect” according to Hansen for the solubilisation of fatty acid methyl esters using an α -pinene/ethanol mixture

1.4 Design of New Solvents with Tailored Properties

1.4.1 Lack of Structures with Specific Properties

The panorama of green solvents according to the COSMO-RS approach evidences that some solvent families are little populated or even not populated at all by “green” solvents (clusters I, II, IV and VI, see Figs. 1.3, 1.6 and 1.7). In particular, the quasi-emptiness of cluster II (weak electron pair donor bases) shows that “green” amide-containing solvents should be developed. Some industrial solutions are starting to emerge, such as amide-homologues of the dibasic esters.

To enlarge the scope of substitution solutions, it is thus of utmost interest to develop new solvent structures with tailored properties for a given application. Reverse engineering is a powerful tool for such an approach, which uses a “top-bottom” strategy [37]. In a recent paper, we have presented a different approach that starts from a chosen bio-sourced building block and generates new molecules by applying chosen chemical transformations [38]. This approach has been exemplified on the generation of itaconic acid-derived solvents and is presented in the last section.

1.4.2 Automatic Generation of New Solvent Structures

The interest for using renewable resources (biomass) instead of fossil resources (coal, oil) has grown exponentially in the last years. The “biorefinery” concept,

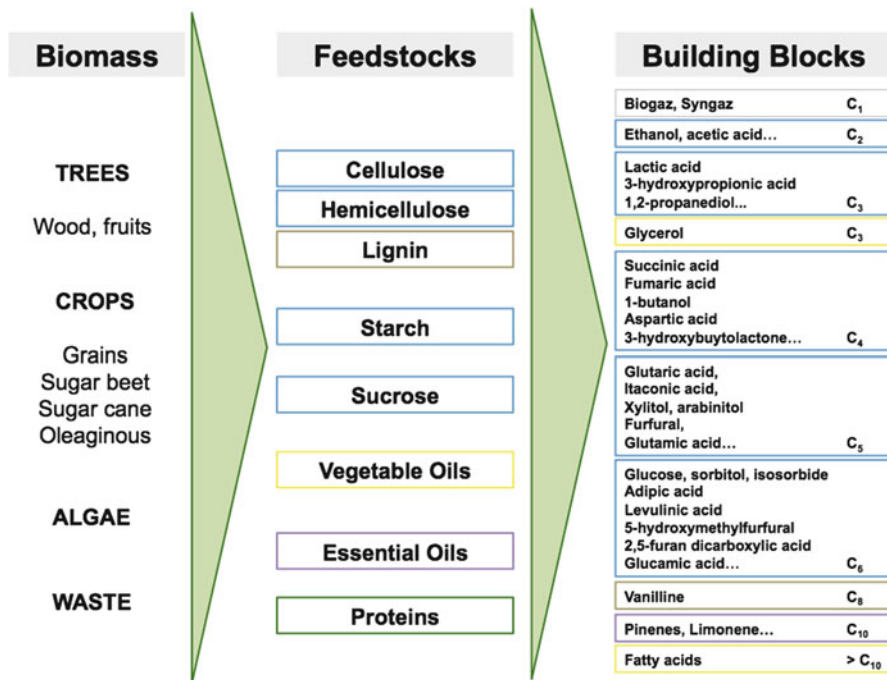


Fig. 1.12 Schematic representation of the building blocks that are currently or will soon be obtained from biorefineries, ordered by increasing number of carbon atoms. The frame colours refer to the chemical feedstocks they come from

i.e. the transposition of the petrorefinery scheme to the processing of biomass, is gaining importance, and a spectrum of products is expected to be obtained from the biomass feedstock in the forthcoming years.

Figure 1.12 shows a schematic route from biomass to defined chemicals. The main biomass sources providing incomes for biorefineries are forestry, dedicated crops and vegetable residues (cobs, straw, sugarcane bagasse, etc.). The exploitation of aquatic biomass (algae) is also a promising source of renewable carbon. The chemical or biochemical transformation of this biomass feedstock may lead to a spectrum of bio-based building blocks available for chemistry. As the main sources are polysaccharides, the available building blocks are currently sugars and sugar derivatives (polyols, organic acids obtained by fermentation).

The methodology used to generate virtual solvents is presented schematically in Fig. 1.13 [38]. The software developed has been named GRASS as the acronym of GeneratoR of Agro-based Sustainable Solvents. It requires three inputs: a bio-based building block, readily available co-reactants and a list of selected transformations that can be applied to the substrate and co-reactants. Virtual products are then automatically generated using the architecture developed by Barone et al. in the previous versions of the programme [39, 40]. This set can, in turn, be an input to GRASS and transformed again.

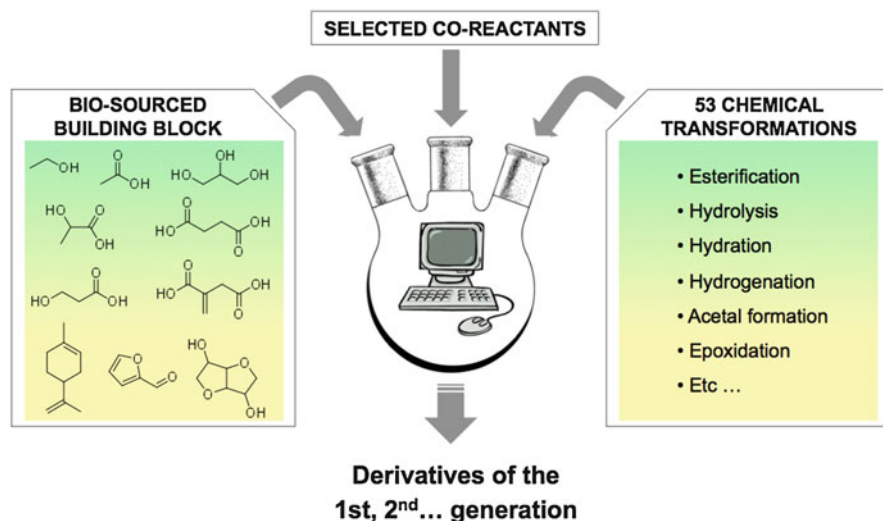


Fig. 1.13 Schematic representation of the “GRASS” programme used to generate automatically virtual solvents

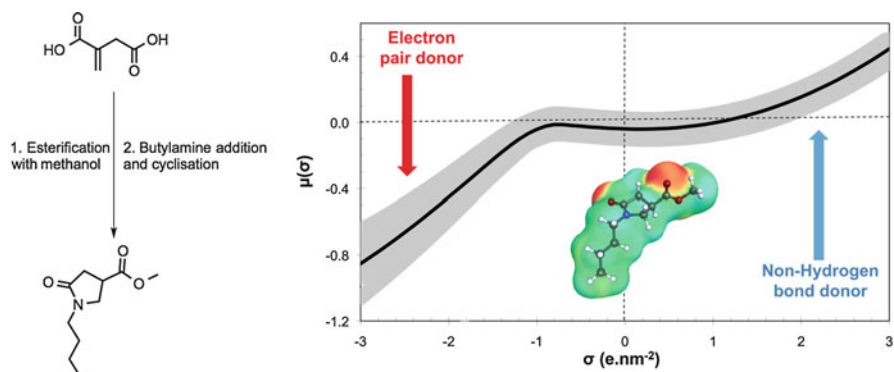


Fig. 1.14 Preparation of *N*-butyl-4-carboxypyrrolidinone ester starting from itaconic acid (left). σ -potential of the product $\mu(\sigma)$ compared to those of classical weak electron pair donor bases belonging to cluster II (in grey) (Adapted from Ref. [38])

The methodology was applied to itaconic acid, an organic acid obtained by sugar fermentation, and it highlighted a family of solvents that are readily obtained in two steps, the *N*-butyl-4-carboxypyrrolidinone esters (Fig. 1.14). This family containing a lactam function is of great interest since it is positioned in cluster II according to the COSMO-RS classification, a class of solvents that is not much populated by current “green” solvents (Fig. 1.14).

1.5 Conclusion

The search for alternative solvents is a hot topic in many industrial fields, including the extraction of natural products. A panel of so-called green solvents is already available, and several tools exist to guide the selection of the most appropriate alternative for a given application. In particular, the traditional Hansen approach coupled with more modern modelling tools such as COSMO-RS provides accurate predictions of solubilising abilities. These predictive tools can also be used to design new solvent structures with tailored properties, in order to enlarge the scope of “green” alternative solvents.

References

1. Reichardt C (1988) *Solvents and solvent effects in organic chemistry*, 2nd edn. Wiley, New York
2. Kamlet MJ, Taft RW (1976) The solvatochromic comparison method. 1. The β -scale of solvent hydrogen-bond acceptor (HBA) basicities. *J Am Chem Soc* 98:377–383
3. Taft RW, Kamlet MJ (1976) The solvatochromic comparison method. 2. The α -scale of solvent hydrogen-bond donor (HBD) acidities. *J Am Chem Soc* 98:2886–2894
4. Kamlet MJ, Abboud JLM, Taft RW (1977) The solvatochromic comparison method. 6. The π^* scale of solvent polarities. *J Am Chem Soc* 99:6027–6038
5. Abraham MH (1993) Scales of solute hydrogen-bonding: their construction and application to physico-chemical and biochemical processes. *Chem Soc Rev* 22:73–83
6. Taft RW, Abboud JLM, Kamlet MJ, Abraham MH (1985) Linear solvation energy relations. *J Solution Chem* 14:153–186
7. Katritzky AR, Fara Dan C, Yang H, Tamm K, Tamm T, Karelson M (2004) Quantitative measures of solvent polarity. *Chem Rev* 104:175–198
8. Murray JS, Politzer P, Famini GR (1998) Theoretical alternatives to linear solvation energy relationships. *Theochem-J Mol Struc* 454:299–306
9. Brinck T, Murray JS, Politzer P (1993) Octanol/water partition coefficients expressed in terms of solute molecular surface areas and electrostatic potentials. *J Org Chem* 58:7070–7073
10. Lowrey AH, Cramer CJ, Urban JJ, Famini GR (1995) Quantum chemical descriptors for linear solvation energy relationships. *Comput Chem* 19:209–215
11. Katritzky AR, Fara DC, Kuanar M, Hur E, Karelson M (2005) The classification of solvents by combining classical QSPR methodology with principal component analysis. *J Phys Chem A* 109:10323–10341
12. Hildebrand J, Scott R (1950) *The solubility of nonelectrolytes*, 3rd edn. Reinhold, New York
13. Stefanis E, Panayiotou C (2008) Prediction of Hansen solubility parameters with a new group-contribution method. *Int J Thermophys* 29:568–585
14. Benazzouz A, Moity L, Pierlot C, Sergent M, Molinier V, Aubry JM (2013) Selection of a greener set of solvents evenly spread in the Hansen space by space-filling design. *Ind Eng Chem Res* 52:16585–16597
15. Gharagheizi F, Sattari M, Angaji MT (2006) Effect of calculation method on values of Hansen solubility parameters of polymers. *Polym Bull* 57:377–384
16. Benazzouz A, Moity L, Pierlot C, Molinier V, Aubry JM (2014) Hansen approach versus COSMO-RS for predicting the solubility of an organic UV filter in cosmetic solvents. *Colloids Surf A Physicochem Eng Asp* (in press). doi: 10.1016/j.colsurfa.2014.03.065
17. Hansen CM (2004) 50 years with solubility parameters – past and future. *Prog Org Coat* 51:77–84
18. Cartier A, Rivail JL (1987) Electronic descriptors in quantitative structure-activity relationships. *Chemometr Intell Lab* 1:335–347

19. Klamt A, Schueuermann G (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J Chem Soc Perk T* 2:799–805
20. Klamt A (1995) Conductor-like screening model for real solvents: a new approach to the quantitative calculation of solvation phenomena. *J Phys Chem* 99:2224–2235
21. Klamt A (2005) COSMO-RS: from quantum chemistry to fluid phase thermodynamics and drug design. Elsevier, Amsterdam
22. Wichmann K, Diedenhofen M, Klamt A (2007) Prediction of blood–brain partitioning and human serum albumin binding based on COSMO-RS σ -moments. *J Chem Inf Model* 47:228–233
23. Klamt A, Eckert F, Diedenhofen M, Beck ME (2003) First principles calculations of aqueous pKa values for organic and inorganic acids using COSMO–RS reveal an inconsistency in the slope of the pKa scale. *J Phys Chem A* 107:9380–9386
24. Durand M, Molinier V, Kunz W, Aubry JM (2011) Classification of organic solvents revisited by using the COSMO-RS approach. *Chem-Eur J* 17:5155–5164
25. Chastrette M (1979) Etude statistique des effets de solvant—I: principes et applications à l'évaluation des paramètres de solvant et à la classification. *Tetrahedron* 35:1441–1448
26. Kerton FM (2009) Alternative solvents for green chemistry. RSC Publishing, Cambridge
27. Plechkova NV, Seddon KR (2008) Applications of ionic liquids in the chemical industry. *Chem Soc Rev* 37:123–150
28. Ranke J, Stolte S, Störmann R, Arning J, Jastorff B (2007) Design of sustainable chemical products. *Chem Rev* 107:2183–2206
29. Klein R, Zech O, Maurer E, Kellermeier M, Kunz W (2011) Oligoether carboxylates: task-specific room-temperature ionic liquids. *J Phys Chem B* 115:8961–8969
30. Imperato G, König B, Chiappe C (2007) Ionic green solvents from renewable resources. *Eur J Org Chem* 7:1049–1058
31. Moity L, Durand M, Benazzouz A, Pierlot C, Molinier V, Aubry JM (2012) Panorama of sustainable solvents using the COSMO-RS approach. *Green Chem* 14:1132–1145
32. Jessop PG (2011) Searching for green solvents. *Green Chem* 13:1391–1398
33. Danaché B, Févotte J, Work team of Matgéné (2009) Éléments techniques sur l'exposition professionnelle à cinq solvants chlorés (trichloroéthylène, perchloroéthylène, chlorure de méthylène, tétrachlorure de carbone, chloroforme) – matrices emplois – expositions à cinq solvants chlorés. Institut de veille sanitaire, Umrestte Lyon, Saint-Maurice
34. Abel S (1990) Fate and exposure assessment of aqueous and terpene cleaning substitutes for chlorofluorocarbons and chlorinated solvents. U.S. Environmental Protection Agency Office of Toxic Substances Exposure Assessment Branch, Washington, DC
35. Tanzi C, Vian M, Ginies C, Elmaataoui M, Chemat F (2012) Terpenes as green solvents for extraction of oil from microalgae. *Molecules* 17:8196–8205
36. Hansen CM (2007) Hansen solubility parameters. CRC Press, Taylor & Francis Group, Boca Raton
37. Heintz J, Touche I, Teles dos Santos M, Gerbaud V (2012) An integrated framework for product formulation by computer aided mixture design. *Comput Aided Chem Eng* 30:702–706
38. Moity L, Molinier V, Benazzouz A, Barone R, Marion P, Aubry JM (2014) In silico design of bio-based commodity chemicals: application to itaconic acid based solvents. *Green Chem* 16:146–160
39. Barone R, Chanon M, Vernin G, Parkanyi C (2005) Generation of potentially new flavoring structures from thiamine by a new combinatorial chemistry program. In: Mussinan CJ, Ho CT, Tatra Contis E, Parliment TH (eds) *Food flavor and chemistry: explorations into the 21st century*. RSC, Cambridge, pp 175–212
40. Barone R, Chanon M, Vernin G, Parkanyi C (2010) Computer-aided organic synthesis as a tool for generation of potentially new flavoring compounds from ascorbic acid. In: Ho CT, Mussinan CJ, Shahidi F, Tatra Contis E (eds) *Recent advances in food and flavor chemistry: food flavors and encapsulation, health benefits, analytical methods and molecular biology of functional foods*. RSC, Cambridge, pp 81–126