From chemical platform molecules to new biosolvents: Design engineering as a substitution methodology

Manon Bergez-Lacoste, Sophie Thiebaud-Roux, Pascale De Caro, and Jean-François Fabre, Université de Toulouse and INRA, LCA (Laboratoire de Chimie Agro-Industrielle) Toulouse, France
Vincent Gerbaud, Université de Toulouse and CNRS LGC (Laboratoire de Génie Chimique), Toulouse, France
Zéphirin Mouloungui, Université de Toulouse and INRA, Toulouse, France

Abstract: The substitution of conventional solvents, in line with regulation changes, requires the use of appropriate methodologies able to generate candidate molecules. Starting from the widely used trial and error approach, we developed two improved, time- and cost-saving methodologies, involving the prediction of molecule properties and reverse design. Reverse design is an innovative methodology to design biosolvents through a virtual laboratory: stages of generation of molecular structures and properties prediction are integrated into a computer-aided molecular design tool providing solutions that meet targeted specifications. These two substitution methodologies were applied in a case study aiming at replacing acetone and methyl ethyl ketone for the solubilization of epoxy resin prepolymers. The generation of performing biosolvents was carried out from furfural as a bio-based platform molecule, thanks to the prediction of different relevant properties (physico-chemical, safety, and environmental characteristics). The reverse design succeeded in ranking these solvent candidates according to their capacity to match the required specifi cations.

Keywords: biosolvents; substitution methodologies; reverse design; furfural

Introduction

Solvents play a major role in the chemical industry. They are essential for many applications, including cleaning, pesticide delivery, coatings, and synthetic chemistry.[1] The market for conventional solvents totals 18 million tonnes per year worldwide, and 4 million tonnes per year in Europe. The most commonly used solvents to date have been of petrochemical origin. These conventional solvents include aliphatics, aromatics, halogenated hydrocarbons, terpenes, alcohols, esters, ketones, and glycol ethers. They account for about 20% of the volatile organic compound (VOC) emissions and are, therefore, a major environmental concern, due to their contribution to the accumulation of tropospheric ozone through photochemical reactions in the atmosphere. They therefore play an indirect role in the greenhouse effect. Moreover, conventional solvents are often highly flammable and

toxic, and their manufacture is also dependent on fossil resources. This context has tended to modify the array of solvents used in industry, and regulations have evolved with advances in our understanding of the potential effects of solvents on the environment and human health. Solvent use is governed by several European directives: directive 1999/13/CE limits VOC emissions from industrial equipment; directive 2004/42/CE concerns the VOC content of solvents used to dilute architectural paints and varnishes and car refinishing paints; and directives 64/548/CEE and 1999/45/CE govern the use of dangerous chemicals. In addition, petrochemical solvents have been subject to REACH regulations since 2007. This regulatory context resulted, for example, in the banning of dichloromethane in paint strippers in 2012 (REACH 276/2010 regulation). Thus, many studies are currently focusing on the replacement of hazardous solvents with more environment-friendly alternatives, according to the ecodesign approach.[2] In recent years, solvent-free processes, ionic liquids, eutectic solvents, fluorous solvents, supercritical fluids, and water as a reaction medium have emerged as interesting alternatives.[3,4] For other applications, however, such as industrial cleaning, these alternatives are not applicable and the replacement of conventional solvents remains difficult. Biomass-derived chemicals offer promising opportunities in the search for eco-friendly, 'sustainable', or 'green' solvents.[5] The market for bio-based solvents in Europe is currently 60 000 tonnes, but is expected to expand over the next few years, in a context of changes to the regulations in force. Biosolvents are not universal solvents like most of the chlorinated or hydrocarbon solvents for which substitutions are required. They are therefore often suitable only for specific applications. For instance, fatty-acid methyl esters have proved effective solvents for cleaning purposes. Dimethyl carbonate is used as a toluene substitute in the pharmaceutical industry. Methyl THF is an interesting solvent for organic synthesis. This lower level of versatility means that a larger number of biosolvents are required to cover the principal applications.

New tools and methodologies are being developed in the face of the need for alternative solvents. Selecting new solvents is a laborious, time-consuming task, as the new solvents must have technical performances at least as good as those of the conventional solvents they replace, together with better health, safety, and environmental profiles. Three different types of methodology can be applied in the substitution approach. We will introduce these methodologies here, comparing their efficiency through an example of substitution. In the case study presented, the aim is to replace acetone or methyl ethyl ketone, both of which have

high vapor pressures, for the solubilization of epoxy resin prepolymers: DGEBA (bisphenol A diglycidyl ether) and TGPA (triglycidyl p-aminophenol ether). These polymers are widely used in industry as starting monomers for the production of epoxy resins and polycarbonates, polymers widespread in several industrial sectors, including the petrochemical industry, packaging, restoration, adhesives, panels, and composites.

Substitution of conventional solvents

Criteria for the selection of new alternatives

The goal of substitution approaches is to move progressively to the use of safer products and processes, in accordance with legislation, public concern, and customer needs. However, clear criteria must be defined for the identification of a safer alternative with a similar technical characteristics profile and production price. Thus, for the design of eco-friendly solvents, the following criteria must be considered:[1]

– **Technical specifications**: linked to the desired performance for the targeted application and the required safety properties.
– **Environmental and health properties**: to meet particular requirements in terms of regulations, standards, and ecolabel specifications.
– **Ecocompatibility** of the process for biosolvent production: it must respect most of the 12 principles of green chemistry proposed by Anastas and Warner.[6]
– The **cost** of the biosolvent, depending on the raw materials and the processes used for their transformation.

These criteria can be checked by evaluating several properties, by experimental or predictive methods (Table 1).

In this context, the selection of biosolvents is based on the determination of several parameters described below.

Solubility power may be assessed by various experimental and theoretical methods:

– **The Kauri-butanol index,** which indicates the maximum amount of solvent that can be added to a solution of Kauri gum in *n*-butanol without causing cloudiness. This measurement is described by ASTM D1133 and provides a scaleless index.
– **The Kamlet Taft solvatochromic scale** is a widely used multiparameter scale for investigating and predicting solvent behavior.[7-9] The Kamlet-Taft solvent parameters

Table 1. Criteria for the selection of new biosolvents.

Criteria	Properties	Associated measurements	Experimental methods	Predictive tools*
Technical	Solubility power	- Kauri-butanol index	ASTM[e] D1133	-
		- Kamlet Taft parameters	Solvatochromic method	-
		- Hildebrand solubility parameter	ASTM D3132	HSPiP
		- Hansen solubility parameters	Calculation	HSPiP
		-Sigma profile, chemical potential	-	COSMO-RS
	Physical state	- Melting point	ISO[f] 1392	IBSS, HSPiP
Technical/Environmental	Volatility	- Vapor pressure	ASTM D2879	IBSS, HSPiP
		- Boiling point	ISO 918	IBSS, HSPiP
		- Evaporation rate	ASTM D3539	HSPiP
Technical/Safety	Flammability	- Flash point	ASTM D92	IBSS, HSPiP
Sanitary	Toxicity	- IC50[a]	OECD[g] section 4	-
Environmental	Biodegradability	- Biodegradation rate	OECD 301	Epi suite, PBT
	Bioaccumulation	- log Kow[b]	OECD 107	IBSS, PBT
		- BCF[c]	OECD section 3	IBSS, Epi suite
	Ecotoxicity	- LC50[d]	OECD section 2	Tox Predict
	Raw materials	- Origin	-	-
Economic/Environmental		- Availability	-	-
		- Cost	-	-
	Chemical process	- Green chemistry indicators -Cost	Calculation from flowsheet	EATOS[h]

*Tools described in the predictive methodology section
[a]IC50: half inhibitory concentration; [b]log kow: octanol-water partition coefficient,
[c]BCF: BioConcentration Factor; [d]LC50: median lethal concentration.
[e]ASTM (American Society for Testing and Materials); [f]ISO (International Organization for Standardization); [g]OECD (Organisation for Economic Co-operation and Development), guidelines for the testing of chemicals; [h]EATOS: Environmental Assessment Tool for Organic Synthesis.

are hydrogen bond donation ability α, hydrogen bond acceptor ability β, dipolarity-polarizability π^*, and a correction term, δ. These parameters were recently measured by Jessop *et al* for solvents of interest in green chemistry.[10]

– **The Hildebrand solubility parameter** δ indicates the relative solvency behavior of a specific solvent.[11] It is derived from the cohesive energy density of the solvent, which is linked to the heat of vaporization.

– **Hansen solubility parameters** are derived from Hansen theory, which describes the cohesion of a substance as the result of the combination of three intermolecular interactions. These interactions correspond to three parameters dividing the Hildebrand parameter into three parts: δD reflects the non-polar interactions derived from London dispersion forces, δP represents the polar interactions linked to Keesom forces and

δH reflects the ability to form hydrogen bonds. These parameters are used as coordinates, to place solvents or solutes in a three-dimensional space. The affinity of a solvent and solute increases with the closeness of their Hansen parameters. Hansen solubility parameters can be determined experimentally by solubility tests in a set of solvents with known Hansen parameters. The solubility sphere includes the maximum number of miscible solvents and no non miscible solvents. This sphere is characterized by a radius. The RED number, reflecting the relative energy difference, is defined as the ratio D/R, where D is the distance between a solute and a solvent (or two solvents), and R is the radius. This number is, thus, a powerful indicator of predicted solubility (or miscibility). If RED < 1, the studied solvent is within the solubility sphere and should dissolve the target. If RED > 1, its position outside the sphere indicates that it is not
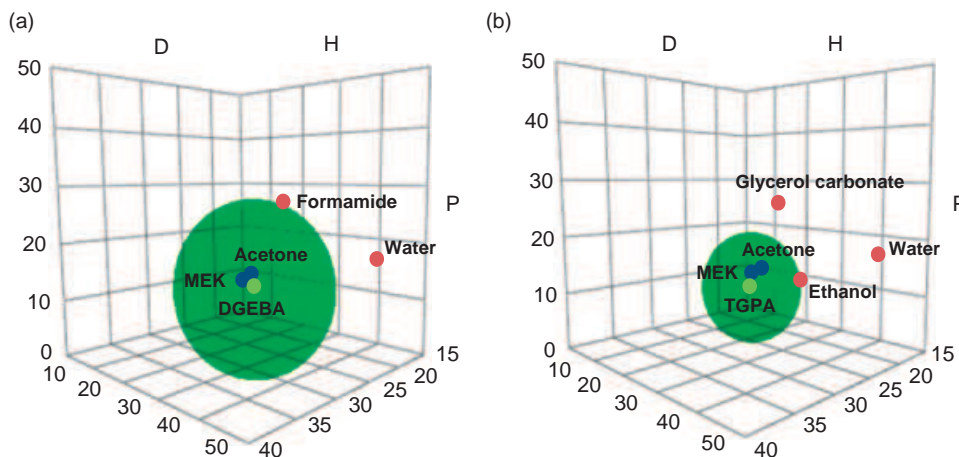
Figure 1. Hansen solubility spheres for (a) DGEBA and (b) TGPA: ● epoxy resin prepolymers, ● conventional solvents and ● non solvents.

predicted to be a good candidate. This notion is illustrated in Fig. 1. The solubility spheres of epoxy resin prepolymers (studied in the next section) are represented. Their conventional solvents, methyl ethyl ketone (MEK) and acetone, have been also located with non-solvents in the Hansen space.

– **COSMO-RS (**Conductor-like Screening Model for Real Solvents) is a new approach for the classification of organic solvents.[12] It is based purely on solvent structure and the solvents are located in a pseudo 3D-space on the basis of principal component analysis and clustering procedures. This representation of solvents as a function of their electron density, sigma profiles and chemical potential may facilitate the identification of potential alternatives to an undesirable solvent.[13]

VOC emissions are strictly limited by regulations and the VOC emissions of industrial equipment must thus be carefully controlled. Vapor pressure and boiling point are good indicators of solvent **volatility**. A high evaporation rate is useful for limiting drying time (for applications such as coatings, nail polish, and paints), but a compromise must be found to decrease VOC emissions.

Flash point is related to **flammability** and must be high enough for safety purposes during the production, conditioning, and transportation steps.

The **impact on the environment** is correlated with solvent biodegradability, the ability of the compound to be broken down by natural micro-organisms into water, carbon dioxide, and biomass. Products with a biodegradability rate of more than 60% in 28 days are generally described as easily biodegradable. Bioaccumulation and ecotoxicity can also be used as indicators of the

environmental impact of a chemical substance: these indicators are assessed through their action on sample populations of living organisms (measurements of median lethal concentration and of the concentrations of substances accumulating in specific organisms).

Economic aspects can be evaluated in terms of the cost of the raw materials and processes involved in the production chain. Raw materials should be renewable, easily available, affordable and should not also be used in the food industry (to avoid competition between the two sectors, pushing up food prices).

The **processes** should be evaluated by green metrics, which can be calculated from the mass balance by EATOS software.[14]

All of these critical factors are taken into account in solvent substitution. In all the various alternative approaches described since the 1990s by Fishkov *et al.*,[15,16] the British Health and Safety Executive,[17] or Collaton and Green,[18] similar criteria were taken into account in the analysis of needs: technical efficiency, health aspects, safety, protection of the environment, process ecoefficiency, economic aspects, and human and organizational factors.

Solvent substitution methodologies

Computing resources, such as the CAGE (coating alternatives guide) or SAGE (solvent alternative guide) software developed by the Research Triangle Institute, can be used to find substitutes for cleaning and paint solvents.[19,20] However, these tools are not suitable for the identification of appropriate alternatives for all applications.

Various different methodologies have been developed for the substitution of conventional solvents.

Trial and error methodology

'Trial and error' is an experimental problem-solving method that can be seen as a basic approach built on practical experiments. For a conventional solvent to be replaced, one or several hypothetical candidates are identified and tested (trials). If these molecules do not match requirements (errors), they are eliminated and others are proposed and tested, with a view to identifying a suitable substitution solvent. Bégin and Gérin have described this methodology for solvent substitution in greater detail, focusing on the following major steps: [21]

– **Identification of the solvent to be replaced** for reasons of toxicity, a poor environmental footprint, high costs, etc.
– **Proposition of alternative solvents** by a chemist able to select bio-based building blocks and to propose derived compounds. The choice of the proposed alternative solvents is generally guided by existing solvent classifications, which are based on dipolar moment, dielectric constant and protic character, but also on the Hansen solubility parameters or Kamlet Taft parameters or the COSMO-RS approach.[13]
– **Tests** of candidates for the target application, such as solubilization of the active principle or industrial cleaning.
– Finally, **candidate selection** tends to be based on a compromise between several criteria.

This methodology has been depicted and used in several case studies aiming to replace trichloroethylene or xylene, for instance.[21] Ethyl lactate was identified by the same authors as a good biosolvent for paint stripper, degreasing agents and precision cleaning.[22]

This method is time-consuming; every solution has to be tested. Moreover, the selected solvent may not necessarily be the best alternative.

Predictive methodology

This methodology was developed as an improvement of the trial and error method, based on prediction of the properties of the molecules proposed by the chemist. As for the trial and error method, the substitution problem and the associated specifications are first identified and alternative solvents are then proposed by the chemist. In this case, the chemist, in charge of the generation of molecules derived from bio-based building blocks, selects the transformation routes respecting green chemistry principles. He may be then helped by a computer-assisted

organic synthesis program able to suggest virtual compounds from one bio-based platform molecule. [23]

Before testing these molecules in the laboratory, they are subjected to a screening process: their properties are predicted with the various tools listed in Table 2 and compared with the specifications. This makes it possible to eliminate any molecules with predicted properties not meeting the defined specifications. The selected molecules are synthesized and tested. This methodology is clearly more productive, because only potentially interesting molecules (predicted properties consistent with the specifications) are synthesized, saving time, energy and raw materials.

The models implemented in the IBSS tool were selected based on their fiability regarding the comparison between predicted properties and experimental properties for a series of known compounds.

All these computational systems were developed to overcome the lack of experimental data for many chemicals. Indeed, physico-chemical properties have been experimentally determined for less than 1% of the approximately 100 000 chemicals registered in the REACH program.[58] They are based on property modeling and must be handled with care, as their domain of applicability determines the molecules to which they can be applied.

IBSS is a CAPD (computer-aided product design) tool developed by the Laboratoire de Génie Chimique (Toulouse, France) for the InBioSynSolv project, a research program supported by the French National Research Agency.[59] The CAPD approach is an extension to multicomponent mixtures of the CAMD (computer-aided molecular design) approach, which principle is described in the next section. Within this tool, a library of property prediction methods is available including physico-chemical properties, as well as toxicity and ecotoxicity models. Once target property values are defined, they can be aggregated into a performance function enabling to check the match of a molecule or mixture with property specifications.

Some of the other tools are dedicated to the prediction of chemical and physical properties, whereas others specifically predict the toxicity and ecotoxicity properties. HSPiP was developed by Professor Abbot and its main application is the prediction of Hansen solubility parameters, although it can also be used to predict various other properties.[60] Sci Finder is a well known database of chemical abstracts with integrated ACD/Labs software for the prediction of some properties. SPARC (Sparc Performs Automated Reasoning in Chemistry) was developed by EPA (Environment Protection Agency) researchers and is

Table 2. Properties determined with predictive tools and predictive models.

Properties	Predictive tools and associated models					
Physico-chemistry	IBSS[a]	PBT[b]	Epi suite[c]	HSPIP[d]	Sci Finder	Spare[e]
Melting point	Marrero, 2001[24]	Reid, 1987[25]	Reid 1987[25]; Gold, 1969[26]	Joback, 1987[27]		
Boiling point	Marrero, 2001[24]		Stein, 1994[28]	Joback, 1987[27]	ACD/Labs[29]	Hilal, 2003[30]
Flash point	Catoire, 2006[31]			Yamamoto[32]	ACD/Labs[29]	
Vapor pressure	Riedel, 1954[33]	Lyman, 1990[34]	Lyman, 1990[34]			
			Lyman, 1985[35]	Riedel, 1954[33]		Hilal, 2003[30]
Molecular volume	Yamada, 1973[36]			Yen, 1966[37]		Hilal, 2007[38]
Density	Hukkerikar, 2012[39]			Yen, 1966[37]		Hilal, 2007[38]
Water solubility	Marrero, 2002[40]	Meylan, 1996[41]	Meylan, 1994[42]			Hilal, 2004[43]
Hansen parameters	Hukkerikar, 2012[39]			Yamamoto[32]		
Viscosity	Joback, 1987[27]			Joback, 1987[27]		
Superficial tension	Conte, 2008[44]					
Toxicity/ecotoxicity	IBSS	PBT	Epi suite	Caesar[f]	Toxpredict[g]	
Persistence		Meylan, 1993[45]				
		Boethling, 1994[46]				
Bioaccumulation		Meylan, 1999[47]				
Toxicity		ECOSAR program[48]				
Kow	Marrero, 2002[40]	Meylan, 2005[49]	Meylan, 2005[49]			
LC$_{50}$	Martin, 2001[50]				Patlewicz, 2008[51]	
BCF	Veith, 1975[52]		Meylan, 1999[47]	Dimitrov, 2005[53]		
Developmental toxicity				Cassano, 2010[54]		
Skin sensitivity				Gerberick, 2006[55]		
Mutagenicity				Ferrari, 2010[56]		
Carcinogenicity				Fjodorova, 2010[57]		

[a]IBSS: InBioSynSolv; [b]PBT: Persistence, Bioaccumulation, Toxicity; [c]Epi suite: Estimation program interface; [d]HSPiP: Hansen Solubility Parameters in Practive; [e]Sparc: Sparc Performs Automated Reasoning in Chemistry; [f]Caesar: Computer- Assisted Evaluation of industrial chemical Substances According to Regulations, [g]Tox Predict: Estimate toxicological hazard.

available online.[61] It uses computational algorithms based on fundamental chemical structure theory to estimate a large array of physical/chemical properties.

PBT (Persistance, Bioaccumulation Toxicity), Caesar (Computer-Assisted Evaluation of industrial chemical Substances According to Regulations), Epi suite (Estimation Program Interface), and Toxpredict were developed more specifically for predicting the impact of products on human health and the environment.[62,63] The PBT and Epi suite were designed under contract for the Office of Chemical Safety and Pollution Prevention of the US Environmental Protection Agency. Caesar and Toxpredict were developed with funding from the European Community. These predictive tools are available either online or as free downloads.

SMILES notation or molecule-drawing software is required as an input for these tools, with the predicted properties as the output.

In the absence of experimental information for a candidate molecule, prediction tools are useful for the primary selection of molecules.

Reverse design

This methodology, which combines the chemist's knowledge and CAMD, is suitable for use in the substitution of conventional solvents and has already proven successful for the identification of alternative products in various fields, such as processes, solvent design, refrigerant fluids, and polymer design.[64–68] Indeed, computer-aided

molecular design has emerged as a powerful tool, given the progress made with group contribution methods in the estimation of values for particular properties. The CAMD method consists in finding molecules satisfying a set of requirements defined *a priori*. Alternative structures are generated by an optimization technique; their properties are predicted and compared to the requirements transformed into property value specifications; and molecules are ranked according to their match with the specifications. The chemist can then select candidate molecules, based on the feasibility of their synthesis. The IBSS tool developed for the InBioSynSolv project is an innovative CAMD tool dedicated to the design of bio-based solvents.[59,69] Indeed, in addition to the abovementioned characteristics of CAMD implemented with a genetic algorithm, a specific functionality has been added for molecular architecture: the ability to impose one or more bio-based building blocks as fragment in the molecule structure, to ensure the development of bio-based solvent.

The **specifications** are provided as target property values to build a performance objective function in the CAMD tool. Thus, the targeted properties are defined as constraints of the optimization problem to be met by the generated structures. The other input of IBSS software is a **platform molecule**. A bio-based building-block (like furfural) is thus given as the starting point for molecular architecture. Common **functional groups** (simple ones like -CH_3, >C=O, -OH or complex ones like -COOH) are also listed in the CAMD tool as elementary blocks. By connecting these building blocks together, IBSS generates molecules containing the bio-based core, predicts their properties and compares them to the specifications. A large set of property models can be used in the specifications, including technical performance (melting point, boiling point), Hansen solubility parameters, safety properties (flash point), and environmental impact (vapor pressure, bioaccumulation factor, and LC_{50} acute toxicity). After a given number of algorithm generations, a population of 100 molecules is obtained. These molecules are ranked according to their performances describing how well they match the specifications. Different weighting factors can be applied to the properties in the performance function, to weight the selection criteria as a function of the targeted application. The choice of weighting factors is usually done based on the user perception of each property importance for his problem to be solved. As the CAPD approach relies upon property prediction methods, a lower weight may be attributed for properties with methods having poor known prediction accuracy, so as to enable the search to find good compromise solutions. Alternatively,

less strict target property value specifications can be set for such properties. Once candidate molecules have been selected, syntheses and tests are performed.

This methodology is highly innovative and the most exhaustive, because specifications and renewable materials are used for the generation of alternative solvents. Although it does not take into account the ecocompatibility of the synthesis process (greenness of the reaction) according to the 24 principles of green chemistry and green engineering, the multicriterion search of the CAPD makes it possible to achieve 'the greenest solvent' in terms of health, safety, and environmental considerations for its use while at the same time, considering technical.[70] The ecocompatibility of the process can be evaluated further, by calculating green metrics.[71,72]

The three methodologies are depicted in Fig. 2.

The results provided by the second and third methodologies will be compared through a case study.

Case study: substitution of common epoxy prepolymer solvents

The substitution methodologies described above are illustrated here for the replacement of conventional solvents for the solubilization of epoxy resins.

Targeted application: cleaning of resins by solubilization

The aim is to design an efficient biosolvent for the solubilization of two epoxy resin prepolymers presented in Fig. 3: bisphenol A diglycidyl ether (DGEBA) and triglycidyl p-aminophenol ether (TGPA).

These prepolymers are widely used in industry as basic monomers for the production of two major polymers: epoxy resins and polycarbonates. The cleaning of manufacturing and conditioning materials requires conventional volatile solvents, such as acetone and methyl ethyl ketone (MEK). Alternative solvents are now required, to limit the emission of VOCs.[73]

The specifications for this context are listed in Table 3, for definition of the required technical and environmental properties of the substitution solvent.

The Hansen parameters of the prepolymers were determined experimentally in a previous study,[74] from the solubility spheres of radii, R = 23.7 MPa$^{1/2}$ for DGEBA and R = 12.3 MPa$^{1/2}$ for TGPA (Fig. 1).

Both the predictive and reverse design methods were used, with a selected bio-based building block, furfural, as the starting material.
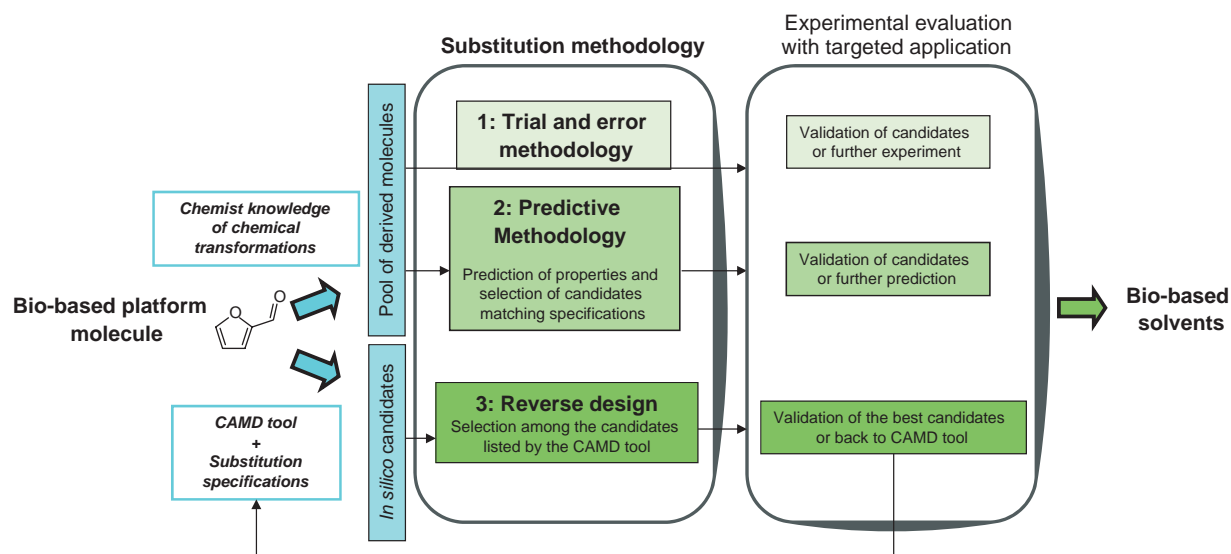
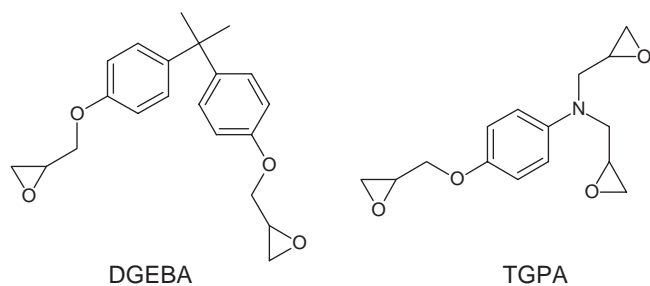Figure 2. Substitution methodologies: Trial and error, predictive method, and reverse design.

Figure 3. Epoxy resin prepolymers.

Furfural as a platform molecule

Furfural is a widely used platform molecule with several advantages for the preparation of new eco-friendly solvents.[75] It is readily available, exclusively from plant feedstock, as no synthetic route exists. Furfural is thus prepared by dehydration of the xylose present in the hemicellulose of plant waste.[76]

The annual production of furfural worldwide is 250 000 tonnes, 60% of which is converted into furfuryl alcohol. The rest is used as a selective solvent for the refinement

Properties	Specifications	Acetone	MEK
Table 3. Specifications for the dissolution of DGEBA and TGPA and properties of conventional solvents to be substituted.			
Technical performance			
Melting point	< 0ºC	–95ºC	–86ºC
Boiling point	100ºC <...< 250ºC	56ºC	80ºC
Solubility			
Hansen parameters (MPa$^{1/2}$)	DGEBA: $\delta D = 17.2$, $\delta P = 6.9$, $\delta H = 8.2$, R = 23.7 TGPA: $\delta D = 17.4$, $\delta P = 6.1$, $\delta H = 6.4$, R = 12.3	$\delta D = 15.5$, $\delta P = 10.4$, $\delta H = 7.0$	$\delta D = 16.0$, $\delta P = 9.0$, $\delta H = 5.1$
Safety			
Flash point	> 61ºC	–18ºC	–9ºC
Environment			
Vapor pressure	< 10 Pa, 20ºC	22 800 Pa, 20ºC	10500 Pa, 20ºC
Bioaccumulation			
BCF	BCF < 500	0.69	-
log kow	< 3	0.24	0.29
R: radius of the Hansen sphere			

of lubricating oils, or as a synthesis intermediate for the production of furan and tetrahydrofuran (Fig. 4).[75] This upgrading scheme is consistent with the condition that bio-based chemistry should not compete with the food industry. Different types of agricultural waste are used, depending on the country: corn cobs in China, sugarcane bagasse in India, the black liquor from wood pulping in Austria, and apricot and olive seeds in Spain.

Alternative structures derived from furfural were screened by the predictive methodology. For the reverse design, furfural was used as the initial building block for CAMD.

Results for the predictive methodology

Several chemical transformations consistent with the principles of green chemistry (e.g. atom economy, small number of steps, limited waste and energy) were considered and applied to the platform molecule. This led to the identification of several candidate compounds, presented in Fig. 5.

The predictive tools were applied only to molecules that could be synthesized easily. The properties predicted by

IBSS (Table 4) were correlated and validated with the results obtained with other tools (e.g. Sparc, HSPiP), for the selection of promising candidates.

The most promising molecule (A) is selected by comparing the predicted properties with the specifications. The actual properties of the candidate molecule can be cheked after synthesis.

Results for the reverse design

The innovative IBSS tool was also used to find alternative solvents for the solubilization of epoxy resins. The specifications defined in Table 3 were taken as the functional objectives and furfural was used as the platform molecule. IBSS can build furfural-derived molecules and predict their properties, for comparison with the specifications. A selection of molecules ranked by IBSS is depicted in Fig. 6, and their predicted properties are gathered in Table 5 as well as the weights attributed to the considered properties in the performance function. The molecule (A) selected by the predictive methodology is among the molecules selected by CAMD, but this tool also proposes new structures. The chemist must consider the feasibility
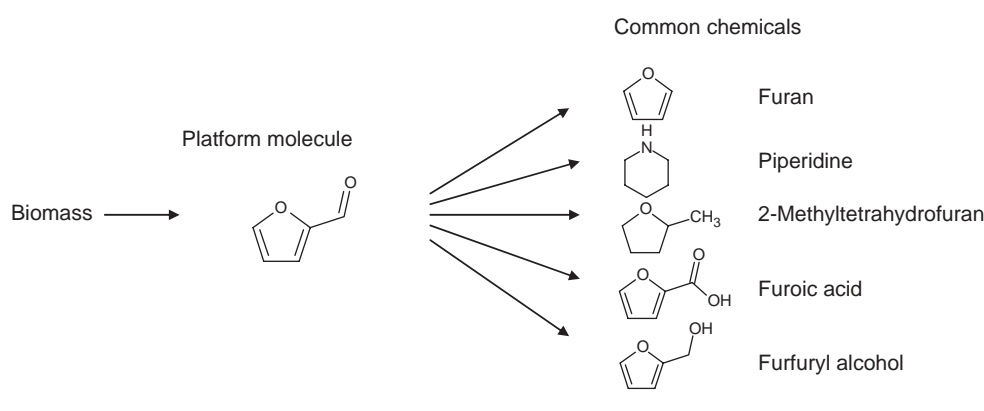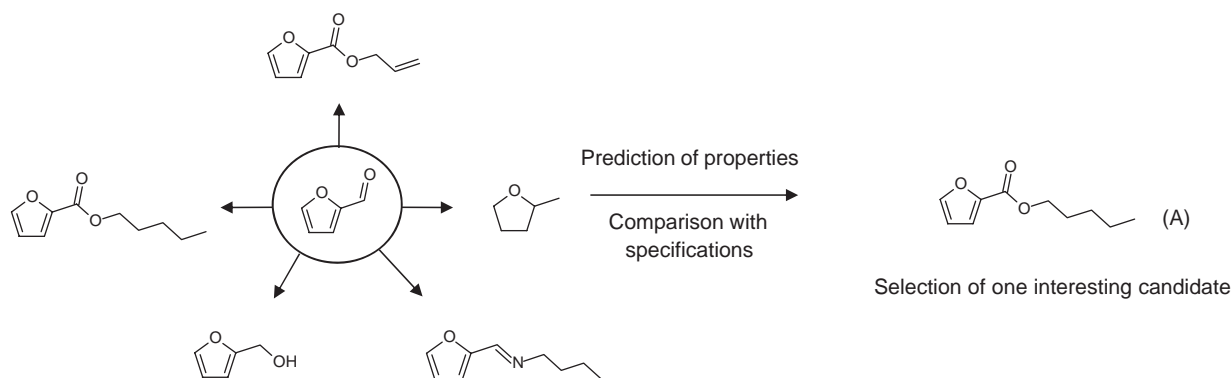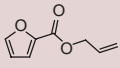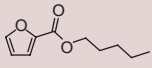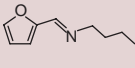

Figure 4. Furfural as a platform molecule.

Figure 5. Selection of a candidate molecule after property prediction.

Table 4. Properties predicted by IBSS for comparison with specifications.

						Specifications
Melting Point [°C]	3	**0**	**–81**	**–15**	109	**≤0**
Boiling Point [°C]	**200**	**236**	77	**163**	**201**	**100 <…< 250**
Flash Point [°C]	59	**88**	–12	**69**	35	**>61**
Vapor Pressure(@293.15K) [Pa]	26	**4**	12839	59	**3**	**<10**
BCF	**4**	**39**	**2**	**0**	**2**	**<500**
Log(Kow)	**1.6**	**2.7**	**1.2**	**0.4**	**2.9**	**<3**
Hansen solubility parameters						
δD [MPa$^{1/2}$]	17.3	17.0	16.8	18.0	17.5	
δP [MPa$^{1/2}$]	7.3	6.5	5.0	9.5	10.9	**close to resins**
δH [MPa$^{1/2}$]	8.5	6.6	4.0	18.5	8.8	
RED*$_{DGEBA}$	**0.02**	**0.07**	**0.20**	**0.45**	**0.34**	**<1**
RED*$_{TGPA}$	**0.19**	**0.07**	**0.23**	1.02	**0.50**	**<1**

*RED = distance between epoxy prepolymer resin and solvent/radius of the resin solubility sphere.
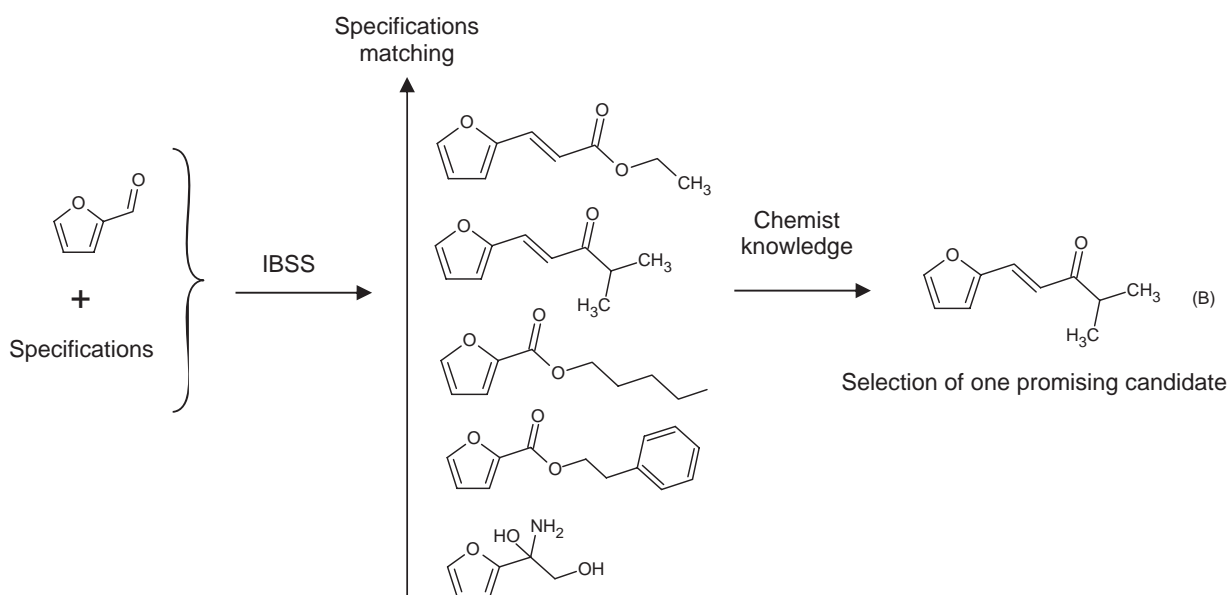Bold values are in agreement with the specifications.

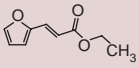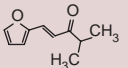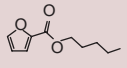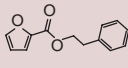Figure 6. Selection from the results generated by the reverse design.

and relative costs of the various syntheses when choosing a candidate from the results obtained with IBSS.

This innovative reverse design methodology has proved efficient for the generation of alternative solvents in accordance with defined specifications. IBSS seems to be a valuable tool for the rapid design of bio-based molecular architectures (a few minutes of calculation).

The performances of the two alternative solvents selected (A and B) by the predictive methodology and reverse design were experimentally validated for the solubilization of DGEBA and TGPA. A solubility of 100 g/L was achieved, confirming the satisfactory performances of these two solvents for resin dissolution in industrial cleaning, for example. The corresponding boiling points are 205 °C for (A) and 225 °C for (B). And the Hansen solubility parameters were determined for (B): δD = 17.1, δP = 8.6, δH = 9.7 with RED$_{DGEBA}$ = 0.09, RED$_{TGPA}$ = 0.33. The experimental values were then compared with the predicted properties, and were found in agreement.

Table 5. Properties predicted by IBSS and comparison with specifications.

	(furan acrylate ethyl ester)	(furan ketone)	(furan butyl ester)	(furan phenethyl ester)	(furan amino diol)	Specifications	Weight
Melting Point [ºC]	4	6	**0**	29	169	≤0	3
Boiling Point [ºC]	**221**	**223**	**236**	289	278	**100 <...< 250**	1
Flash Point [ºC]	**79**	**91**	**88**	**155**	**181**	>61	1
Vapor Pressure(@293.15K) [P_a]	**4.8**	12	**4.0**	**0.04**	**0.003**	<10	1
BCF	**1**	**2**	39	**2**	**2**	<500	1
Log(Kow)	**1.7**	**2.3**	**2.7**	**2.8**	**1.8**	<3	1
Hansen solubility parameters							
δD	17.2	17.1	17.0	19.0	18.6		-
δP	6.3	7.0	6.5	6.9	7.8	**close to resins**	-
δH	8.8	6.9	6.6	6.5	20.9		-
RED*$_{DGEBA}$	**0.03**	**0.06**	**0.07**	**0.17**	0.6	<1	4
RED*$_{TGPA}$	**0.19**	**0.09**	**0.07**	**0.27**	1.2	<1	4

*RED = distance between epoxy prepolymer resin and solvent/radius of the resin solubility sphere.
Bold values are in agreement with the specifications.

Conclusion

Research into green solvents is often a complex issue for the chemical industry. This study highlights the importance of using a rational methodology to achieve high-performance substitutions.

Both predictive methodology and reverse design proved effective for the development of new functional solvents meeting specifications. We checked the performances of the furfural derivatives selected with these methodologies for solubilization of the targeted solutes, DGEBA and TGPA.

Several tools for the prediction of solvent properties have been described. In combination with the predictive methodology, they facilitate the transition from conventional solvents to eco-friendly solvents. This approach saves time and reduces costs, by ensuring that only the best candidates are synthesized. The predictive methodology is easy to apply when alternative structures have already been identified by the chemist.

However, when starting from a bio-based building block, the reverse design is the most complete and rapid methodology, as all steps are managed by the CAMD tool, including the generation of derivative molecules. The reverse design approach goes further than other methodologies: it has the advantage of integrating constraints relating to both properties and molecular structures.

Moreover, the CAMD tool IBSS is being extended for prediction of the properties of mixtures, opening up new possibilities for increasing the substitution potential of biosolvents.[77]

Finally, the substitution process deserves to be complemented by a 'green chemistry' approach to the chemical conversion step. Green metrics, such as atom economy and environmental factors, can then be taken into account in the selection of ecofriendly processes for the preparation of biosolvents.

Acknowledgement

References

1. De Caro P and Thiebaud-Roux S, Biosolvants. *Techniques de l'Ingénieur* **IN 102**:1–9 (2008).
2. Jessop PG, Searching for green solvents. *Green Chem* **13**:1391–1398 (2011).
3. Sheldon RA, Green solvents for sustainable organic synthesis: State of the art. *Green Chem* **7**:267–278 (2005).
4. DeSimone JM, Practical approaches to green solvents. *Science* **297**:799–803 (2002).
5. Bandres M, de Caro P, Thiebaud-Roux S and Borredon ME, Green syntheses of biobased solvents. *C R Chimie* **14**:636–646 (2011).
6. Anastas PT and Warner JC, *Green Chemistry: Theory and Practice*. Oxford University Press, New York, USA (2000).
7. Kamlet MJ and Taft RW, The solvatochromic comparison method. I. The β-scale of solvent hydrogen-bond acceptor (HBA) basicities. *J Am Chem Soc* **98**:377–383 (1976).
8. Kamlet MJ, Abboud JL and Taft RW, The solvatochromic comparison method. 6. The π* scale of solvent polarities. *J Am Chem Soc* **99**:6027–6038 (1977).
9. Kamlet MJ, Abboud JLM, Abraham MH and Taft RW, Linear solvation energy relationships. 23. A comprehensive collection of the solvatochromic parameters, π*, α, and β, and some

methods for simplifying the generalized solvatochromic equation. *J Org Chem* **48**:2877–2887 (1983).

10. Jessop PG, Jessop DA, Fu D and Phan L, Solvatochromic parameters for solvents of interest in green chemistry. *Green Chem* **14**:1245–1259 (2012).

11. Hildebrand JH and Scott RL, *The solubility of Non-Electrolytes*, 3rd ed., Reinhold Publishing Co., New York (1950).

12. Durand M, Molinier V, Kunz W and Aubry J-M, Classification of organic solvents revisited by using the COSMO-RS approach. *Chemistry* **17**:5155–5164 (2011).

13. Moity L, Durand M, Benazzouz A, Pierlot C, Molinier V and Aubry J-M, Panorama of sustainable solvents using the COSMO-RS approach. *Green Chem* **14**:1132–1145 (2012).

14. Eissen M and Metzger JO, Environmental performance metrics for daily use in synthetic chemistry. *Chemistry* **8**:3580–3585 (2002).

15. Filskov P, Goldschmidt G, Hansen MK, Höglund JL, Johansen T, Pedersen CL and Wibroe L, *Substitutes for Hazardous Chemicals in the Workplace*. CRC Press Inc., Lewis Publishers, Boca Raton, FL (1996):

16. Goldschmidt G, An analytical approach for reducing workplace hazards through substitution. *Am Ind Hyg Assoc J* **54**:36–43 (1993).

17. HSE, *7 Steps to Succesful Substitution of Hazardous Substances*. United Kingdom Health and Safety Executive, HSE Books, Sudbury, Suffolk, UK (1994).

18. Callahan MS and Green B, *Hazardous Solvent Source Reduction*. McGraw-Hill, NY (1995).

19. Cornstubble DR, *CAGE 1.0 - Coating Alternatives Guide - User's Guide*. Research Triangle Institute and united States Environmental Protection Agency, National Risk Management Research Laboratory, Air Pollution Prevention and Control Division, Research Triangle Park, NC (1996).

20. Monroe KR and Hill EA, Solvent Alternatives Guides (SAGE), Computer assisted guidance for solvent replacement. *CFC and Halon Alternatives Conference - Stratospheric Ozone Protection for the 90s*, Washington, October 20–22 (1993).

21. Bégin D and Gérin M, *Substitution des Solvants. Etude de cas d'implantation*. Institut de Recherche Robert-Sauvé en Santé et Sécurité du Travail, Montréal (2001).

22. Bégin D and Gérin M, *La substitution des solvants par le lactate d'éthyle, in Bilans de connaissances, Université de Montréal (IRSST)*, Montréal (2005).

23. Moity L, Molinier V, Benazzouz A, Barone R, Marion P and Aubry J-M, In silico design of bio-based commodity chemicals: Application to itaconic acid based solvents. *Green Chem* **16**:146–160 (2014).

24. Marrero J and Gani R, Group-contribution based estimation of pure component properties. *Fluid Phase Equilibr* **183/184**:183–208 (2001).

25. Reid RC, Prausnitz JM and Poling BE, *The Properties of Gases and Liquids*, *Chapter 2*. McGraw-Hill Inc., NY (1987).

26. Gold PI and Ogle GJ, Estimating thermophysical properties of liquids. Part 4 - Boiling, freezing and triple-point temperatures. *Chem Eng* **76**:119–122 (1969).

27. Joback KG and Reid RC, Estimation of pure-component properties from group-contributions. *Chem Eng Commun* **57**:233–243 (1987).

28. Stein SE and Brown RL, Estimation of normal boiling points from group contributions. *J Chem Inf Comp Sci* **34**:581–587 (1994).

29. Advanced Chemistry Development (ACD/Labs), *Software V11.02 (© 1994–2013 ACD/Labs)*. [Online]. ACD/Labs Percepta Predictors. Available at: http://www.acdlabs.com/products/percepta/predictors.php [September 26, 2013].

30. Hilal SH, Karickhoff SW and Carreira LA, Prediction of the vapor pressure boiling point, heat of vaporization and diffusion coefficient of organic compounds. *QSAR CombSci* **22**:565–574 (2003).

31. Catoire L, Paulmier S and Naudet V, Estimation of closed cup flash points of combustible solvent blends. *J Phys Chem* **35**:9–14 (2006).

32. Yamamoto H. [Online]. About the Pirika.com. Available at: http://pirika.com/ [September 26, 2013].

33. Riedel L, Eine neue universelle admpfdruckformel untersuchungen über eine Erweiterung des Theorems der übereinstimmenden Zustände. Teil I. *Chem-Ing-Tech* **26**:83–89 (1954).

34. Lyman WJ, Reehl WF and Rosenblatt DH, *Handbook of Chemical Property Estimation Methods*. American Chemical Society, Washington DC (1990),

35. Lyman WJ, *Environmental Exposure From Chemicals, Volume I*, ed by NeelyWB and Blau GE. CRC Press, Boca Raton, FL (1985).

36. Yamada T and Gunn RD, Saturated liquid molar volumes. Rackett equation. *J Chem Eng Data* **18**:234–236 (1973).

37. Yen LC and Woods SS, A generalized equation for computer calculation of liquid densities. *AIChE J* **12**:95–99 (1966).

38. Hilal SH, Saravanaraj AN, Whiteside T and Carreira LA, Calculating physical properties of organic compounds for environmental modeling from molecular structure. *J Comput-Aid Mol Des* **21**:693–708 (2007).

39. Hukkerikar AS, Sarup B, Ten Kate A, Abildskov J, Sin GR and Gani R, Group-contribution+ (GC+) based estimation of properties of pure components: Improved property estimation and uncertainty analysis. *Fluid Phase Equilibr* **321**:25–43 (2012).

40. Marrero J and Gani R, Group-contribution-based estimation of octanol/water partition coefficient and aqueous solubility. *Ind Eng Chem Res* **41**:6623–6633 (2002).

41. Meylan WM, Howard PH and Boethling RS, Improved method for estimating water solubility from octanol/water partition coefficient. *Environ Toxicol Chem* **15**:100–106 (1996).

42. Meylan WM and Howard PH, *Validation of Water Solubility Estimation Methods Using Log Kow for Application in PCGEMS & EPI*. Office of Pollution Prevention and Toxics, U.S. Environmental Protection Agency, Environmental Science Center, Washington DC (1994).

43. Hilal SH, Karickhoff SW and Carreira LA, Prediction of the solubility, activity coefficient and liquid/liquid partition coefficient of organic compounds. *QSAR Comb Sci* **23**:709–720 (2004).

44. Conte E, Martinho A, Matos HA and Gani R, Combined group-contribution and atom connectivity index-based methods for estimation of surface tension and viscosity. *Ind Eng Chem Res* **47**:7940–7954 (2008).

45. Meylan WM and Howard PH, Computer estimation of the atmospheric gas-phase reaction rate of organic compounds with hydroxyl radicals and ozone. *Chemosphere* **26**:2293–2299 (1993).

46. Boethling RS, Howard PH, Meylan W, Stiteler W, Beauman J and Tirado N, Group contribution method for predicting probability and rate of aerobic biodegradation. *Environ Sci Technol* **28**:459–465 (1994).

47. Meylan WM, Howard PH, Boethling RS, Aronson D, Printup H and Gouchie S, Improved method for estimating bioconcentration/bioaccumulation factor from octanol/water partition coefficient. *Environ Toxicol Chem* **18**:664–672 (1999).

48. United States Environmental Protection Agency (EPA). [Online]. Ecological Structure Activity Relationships (ECOSAR). Available at: http://www.epa.gov/oppt/newchems/tools/21ecosar.htm [September 27, 2013].

49. Meylan WM and Howard PH, Estimating octanol-air partition coefficients with octanol-water partition coefficients and Henry's law constants. *Chemosphere* **61**:640–644 (2005).

50. Martin TM and Young DM, Prediction of the acute toxicity (96-h LC50) of organic compounds to the fathead minnow (Pimephales promelas) using a group contribution method. *Chem Res Toxicol* **14**:1378–1385 (2001).

51. Patlewicz G, Jeliazkova N, Safford RJ, Worth AP and Aleksiev B, An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR QSAR Environ Res* **19**:495–524 (2008).

52. Veith GD and Konasewich DE, *Structure-activity correlations in studies of toxicity and bioconcentration with aquatic organisms*. Great Lakes research advisory Board, Windsor, Ontario (1975)

53. Dimitrov S, Dimitrova N, Parkerton T, Comber M, Bonnell M and Mekenyan O, Base-line model for identifying the bioaccumulation potential of chemicals. *SAR QSAR Environ Res* **16**:531–554 (2005).

54. Cassano A, Manganaro A, Martin T, Young D, Piclin N, Pintore M *et al.,* CAESAR models for developmental toxicity. *Chem Cent J* **4**:S4 (2010).

55. Gerberick GF, Ryan CA and Kern PS, Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. *Dermatitis* **16**:157–202 (2005).

56. Ferrari T and Gini G, An open source multistep model to predict mutagenicity from statistical analysis and relevant structural alerts. *Chem Cent J* **4**:S2 (2010).

57. Fjodorova N, Vracko M, Novic M, Roncaglioni A and Benfenati E, New public QSAR model for carcinogenicity. *Chem Cent J* **4**:S3 (2010).

58. United States Environmental Protection Agency (EPA). [Online]. SPARC Performs Automated Reasoning in Chemistry. Available at: http://www.epa.gov/athens/research/sparc.html [September 26, 2013].

59. Heintz J, Belaud J-P and Gerbaud V, Chemical enterprise model and decision-making framework for sustainable chemical product design. *Computers in Industry* **65**:505–520 (2014).

60. Abbott S, Hansen CM and Yamamoto H. [Online]. The Official Hansen Solubility Parameter Site. Available at: http://www.hansen-solubility.com [September 26, 2013].

61. SPARC. [Online]. SPARC Performs Automated Reasoning in Chemistry. Available at: https://archemcalc.com/sparc/ [September 26, 2013].

62. PBT profiler, a component of OPPT's P2 Framework, assessing chemical in the absence of data. [Online]. Persistent, Bioaccumulative, and Toxic Profiles Estimated for Organic Chemicals. Available at: http://www.pbtprofiler.net/default.asp [September 26, 2013].

63. Computer Assisted Evaluation of Industrial Chemical Substances According to Regulations (CAESAR). [Online]. Available at: http://www.caesar-project.eu/ [September 26, 2013].

64. Song J and Song HH, Computer-aided molecular design of environmentally friendly solvents for separation processes. *Chem Eng Technol* **31**:177–187 (2008).

65. Sherman J, Chin B, Huibers PDT, Garcia-Valls R and Hatton TA, Solvent replacement for green processing. *Environ Health Persp* **106**:253–271 (1998).

66. Sinha M, Achenie LEK and Ostrovsky GM, Environmentally begnin solvent design by global optimization. *Comput Chem Eng* **23**:1381–1394 (1999).

67. Constantinou L, Bagherpour K, Gani R, Klein JA and Wu DT, Computer aided product design: Problem formulations, methodology and applications. *Comput Chem Eng* **20**:685–702 (1996).

68. Maranas CD, Optimization accounting for property prediction uncertainty in polymer design. *Comput Chem Eng* **21**:S1019–S1024 (1997).

69. Heintz J, *Systemic approach and decision process for sustainability in chemical engineering: Application to computer aided product design*, PhD thesis. Laboratoire de Génie Chimique, Toulouse. (2012)

70. Capello C, Fischer U and Hungerbuhler K, What is a green solvent? A comprehensive framework for the environmental assessment of solvents. *Green Chem* **9**:927–934 (2007).

71. Andraos J, Unification of reaction metrics for green chemistry II: Evaluation of named organic reactions and application to reaction discovery. *Org Process Res Dev* **9**:404–431 (2005).

72. Constable DJC, Curzons AD and Cunningham VL, Metrics to 'green' chemistry-which are the best? *Green Chem* **4**:521–527 (2002).

73. Medina-Gonzalez Y, Thiebaud-Roux S, De-Caro P, Hernandez-Ochoa L and Mouloungui Z, The role of co-additive on epoxy resin pre-polymers solubilization in supercritical CO_2. *J Supercrit Fluid* **38**:13–17 (2006).

74. Medina Gonzalez Y, de Caro P, Thiebaud-Roux S and Lacaze-Dufaure C, Fatty acid methyl esters as biosolvents of epoxy resins: A physicochemical study. *J Solution Chem* **36**:437–446 (2007).

75. Hoydonckx HE, Van Rhijn WM, Van Rhijn W, De Vos DE and Jacobs PA, Furfural and Derivatives, in *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim (2007).

76. Mamman AS, Lee JM, Kim YC, Hwang IT, Park NJ, Hwang YK et al., Furfural: Hemicellulose/xylosederived biochemical. *Biofuel Bioprod Bioref* **2**:438–454 (2008).

77. Heintz J, Touche I., Teles dos Santos M and Gerbaud V, An integrated framework for product formulation by computer aided mixture design. *Comput-Aided Chem Eng* **30**:702–706 (2012).

Manon Bergez-Lacoste

Manon Bergez-Lacoste is a PhD student in Bio-based Chemistry in Laboratoire de Chimie Agro-industrielle, at the University of Toulouse, France. Her research interests include the *in silico* and *in vivo* design of chemicals, particularly biosolvents, in the framework of an eco-design of bio-based products.

Sophie Thiebaud-Roux

Dr Sophie Thiebaud-Roux is Associate Professor and head of the Department of Chemistry at INP-ENSIACET in Toulouse (France). She holds a PhD degree in organic chemistry applied to agro-molecules. During the last ten years, her research work focuses on the development of bio-solvents under reaction conditions close to the principles of Green chemistry.

Pascale de Caro

Dr Pascale de Caro is an associate Professor at INPT-ENSIACET in Toulouse. She is graduated from the Chemical Engineer School of Toulouse. Her research work in Laboratoire de Chimie Agro-Industrielle is focused on the development of bio-based functional products from vegetable raw materials, according to a green chemistry approach.

Jean-François Fabre

Dr Jean-François Fabre is a Physico-Chemistry Engineer at Laboratoire de Chimie Agro-Industrielle, Toulouse. His research interests include extraction and valorization of agricultural derived products with particular emphasis on colloids chemistry and molecular modeling.

Vincent Gerbaud

Dr Vincent Gerbaud is a research director at the French National Scientific Research Center CNRS at Laboratoire de Génie Chimique in Toulouse. His research field concerns modeling and simulation in process system engineering, with a strong interest in small scale modeling: molecular simulation, thermodynamics, computer aided molecular and mixture design. At the process scale, He also focuses with the design and synthesis of distillation processes, like extractive and azeotropic distillation.

Zéphirin Mouloungui

Dr Zephirin Mouloungui is a research director at INRA (Institut National pour la Recherche Agronomique), at the head of the research team for the chemical reactivity of agromolecules and oleochemicals, in Laboratoire de Chimie Agro-Industrielle. His expertise deals with the development of the renewable resources in the fields of chemistry, bioproducts and neo-materials, according to innovating integrated schemes of production.