Integrated Solvent and Process Design Exemplified for a Diels–Alder Reaction

Teng Zhou and Kevin McBride
Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, D-39106 Magdeburg, Germany

Xiang Zhang
Process Systems Engineering, Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany

Zhiwen Qi
Max Planck Partner Group at the State Key Laboratory of Chemical Engineering, East China University of Science and Technology, 130 Meilong Road, 200237 Shanghai, China

Kai Sundmacher
Process Systems Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, D-39106 Magdeburg, Germany, Process Systems Engineering, Otto-von-Guericke-University Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany

DOI 10.1002/aic.14630
Published online October 1, 2014 in Wiley Online Library (wileyonlinelibrary.com)

A new kind of solvent descriptor obtained from quantum chemical calculations is introduced. Group contributions to each solvent descriptor are regressed for 71 UNIFAC groups. A reaction kinetic model is built by correlating a set of experimentally determined reaction rate constants in various solvents with the corresponding theoretical solvent descriptors. Based on the kinetic model and the developed group contribution method, a computer-aided molecular design problem is formulated and optimal solvents to achieve highest reaction rates are identified. For considering the multiple and complicated effects of solvents on a chemical process, an integrated solvent and process design is performed. Solvent molecular structures and process operations are simultaneously optimized by the formulation and solution of a mixed-integer nonlinear program. The proposed design methodology is exemplified for a selected Diels–Alder reaction.
 AIChE J, 61: 147–158, 2015
Keywords: computer-aided molecular design, solvent design, process design, process optimization, Diels–Alder reaction

Introduction

Solvents are widely used as reaction media in liquid phase chemical processes. It has been proven that the variation of the type of solvent can dramatically change the reaction rate and even the reaction mechanism.[1] Therefore, the selection of solvents is regarded as one of the most important decisions in the early stages of chemical process design. Unfortunately, it is unrealistic to perform kinetic experiments for each solvent due to the huge number of potential candidates. To allow for an efficient prescreening of existing solvents or the design of new solvents, one has to make use of reliable theoretical methods.

In the last three decades, computer-aided molecular design (CAMD) approaches have been proposed and successfully applied to solvent design. In CAMD methods, each chemical species is usually represented as a collection of functional groups. Their physical and chemical properties are calculated from various structure-based property predictive methods. Promising solvents were identified through either generate-and-test methods[2–9] or optimization-based mathematical programming methods.[10–21] In generate-and-test methods, all chemically feasible combinations of the preselected groups are generated and then filtered by a series of boundaries on the target solvent and process properties. In optimization-based molecular design methods, CAMD problems are formulated as optimization programs where discrete variables denote the number of groups that are present in the molecule, and chemical feasibilities and property requirements are reformulated as either constraints or objectives. Optimal molecular structures are identified by solving this optimization program without testing all possible solvents in the design space. Traditionally, for most solvent-involved processes, solvent design and process optimization are treated as two separated problems[22–24] or reconstructed into interlinked substages[25–30] due to the many complex combinations of integer and continuous design variables. Recently,

Additional Supporting Information may be found in the online version of this article.

Correspondence concerning this article should be addressed to K. Sundmacher at sundmacher@mpi-magdeburg.mpg.de.

researchers explored the possibility of combining solvent and process design by replacing solvents with certain continuous parameters that can be optimized simultaneously with process variables and then mapped from a database onto an existing optimal solvent.[31–33]

Despite the wide applicability of CAMD methods in solvent design, it should be noted that until now most CAMD methods focus on solvents for separation processes. Rarely are such methods implemented for solvent selection in chemical reactions. Gani et al.[34] developed a systematic method for the selection of solvents for organic reactions. Their method shows excellent applicability for certain reaction systems, but it requires a large amount of *a priori* knowledge about reactions and solvents to construct the solvent score table. Importantly, the method cannot give quantitative predictions on solvent effects. As is well known, the quantification of reaction rates in solvents relies on the prediction of Gibbs energies in solution of the reactants and the transition state. For estimating these energy levels, quantum chemical (QC) calculations (e.g., *ab initio* or DFT computations) are currently state-of-the-art. But, since the transition state configuration is often difficult to determine, QC calculations are often time-consuming and the resulting reaction rates are, in many cases, quantitatively inaccurate. Nowadays, the most common and practical way to study solvent effects on chemical reaction rates is to correlate a set of experimental reaction rate constants in different solvents with some physical properties of these solvents. The linear solvation energy relationships method is one of the most successful approaches where empirical or experimentally determined solvent parameters, such as the solvent solvatochromic parameters, are used to quantify the effect of solvents on chemical reactions.[35,36] This method can generally provide satisfactory predictions of the rates of a specific reaction in some prespecified solvents. However, it can neither be directly applied for extensive solvent screening nor for solvent design, because the required parameters are available for a limited number of solvents only.

Over the past decade, chemical engineers have gained better access to theoretical chemistry tools and computer power has reached a very high level. This opens the possibility to make use of theoretical descriptors for correlating solvent effects. Theoretical descriptors can be derived from molecular electronic or structural information, which can be correlated with experimentally determined solvent parameters.[37] Because theoretical descriptors are easy-to-generate and do not require experimental data, they allow for the screening and design of solvents in a very large molecular structure space. Recently, our group has developed a fast solvent screening method by which experimental reaction properties (rates or selectivities) in various solvents were correlated with theoretical solvent descriptors obtained from solvent conductor-like screening model (COSMO)-RS calculations.[38] This *a priori* solvent profiling method allows for an efficient prescreening of a great set of solvents. However, as COSMO-RS treats solvents with their real molecular structures, such an approach cannot be used for solvent molecular design. In this work, we introduce a new kind of solvent theoretical descriptor and determine the contributions of common molecular subgroups to these descriptors. Based on a prebuilt reaction kinetic model and the developed group contribution (GC) method, optimal reaction solvents are identified through the solution of a CAMD optimization problem.

Conversely, we should be aware that the identification of optimal reaction solvents is not the ultimate goal of chemical engineers. It is clear that solvents not only change the reaction rate but also have significant impact on many other process properties. Considering the strong interdependencies between solvent characteristics and the process performance, the integration of solvent design and process optimization-based on a process-wide approach is a prerequisite for achieving an optimal chemical process. Folić et al.[39,40] developed a method for designing solvents for chemical reactions, where solvent solvatochromic parameters are used to quantify their kinetic effects. Despite the wide applicability of their method, it is still not clear how this technique can be incorporated into integrated solvent-process optimization problems. This motivates the development of a new method for the selection of reaction solvents under the consideration of process constraints.

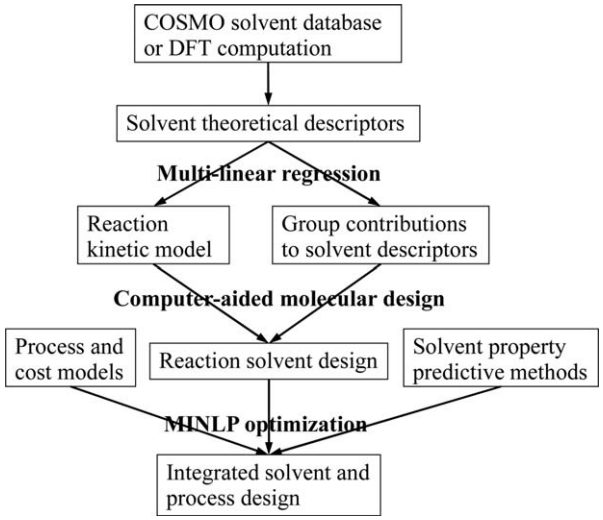

Figure 1. Structural outline of the integrated solvent and process design methodology.

The overall framework of this work is illustrated in Figure 1. First, we introduce a new kind of solvent theoretical descriptor. A kinetic model is then regressed by correlating the reaction rate constants of the here selected Diels–Alder (DA) reaction example in different solvents with the corresponding solvent theoretical descriptors. Afterward, a GC method, that can be used to estimate the solvent descriptors, is developed. Based on the reaction kinetic model and the GC method, solvents are designed for maximizing the reaction rates. Finally, the reaction solvent and process operation conditions are simultaneously optimized through the solution of a mixed-integer nonlinear program (MINLP)-formulated integrated solvent and process design problem.

Solvent Theoretical Descriptors and Reaction Rate Regression

In quantum chemistry, solvents are normally simulated by continuum solvation models (CSMs), where the real solvent is approximated by a dielectric continuum of permittivity, ε.[41] Klamt and Schüürmann[42] developed a technical modification of the dielectric CSMs, known as the COSMO, which replaces the dielectric boundary conditions with a much simpler scaled-conductor boundary condition. COSMO is now a widely used CSM due to its considerable reduction of the computational complexity and increased numerical robustness. Figure 2 briefly depicts the COSMO calculation procedure. The molecule of interest is embedded into a virtual

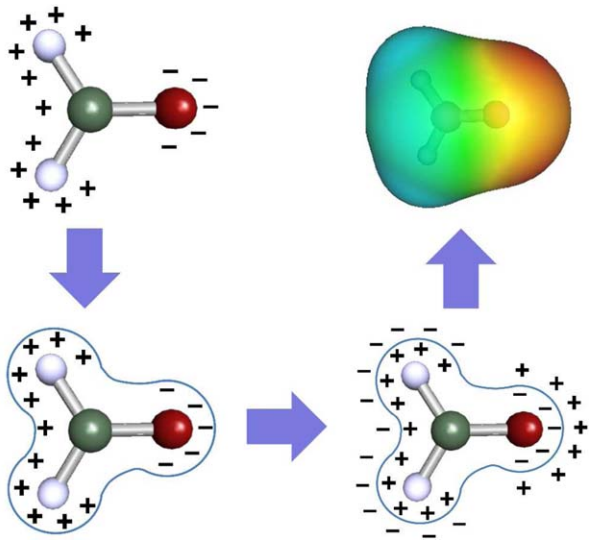

Figure 2. Schematic diagram of the COSMO model.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

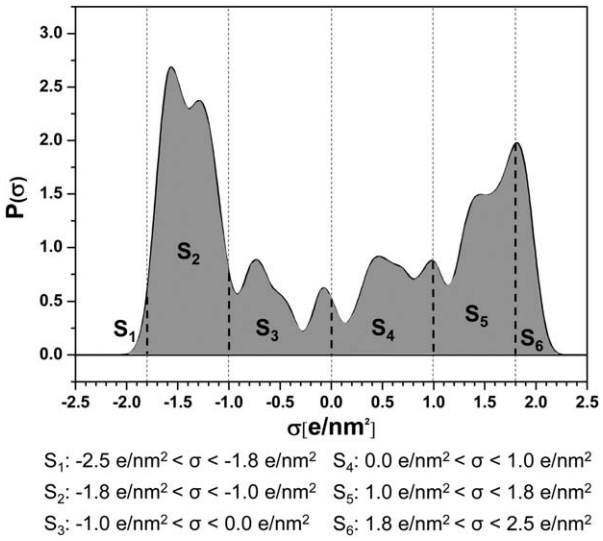

Figure 4. Solvent theoretical descriptors defined by COSMO σ-profile areas (σ-profile of water is used for illustration).

conductor, and the conductor screening charge density (SCD) σ is ultimately induced. Through QC self-consistency computations (usually the DFT method), the molecule converges to its optimal COSMO state with respect to its energy, electronic density, conductor SCD, and the geometry in the conductor. After the DFT calculation, the three-dimensional distribution of the SCD on the surface of the molecule is converted into a composition function, the so called "σ-profile", $P(\sigma)$, which quantifies the number of surface segments one can find within a certain SCD σ-interval $[\sigma - d\sigma/2, \sigma + d\sigma/2]$. It should be noted that the σ-profile of a compound needs to be calculated only once and afterward it is stored in a COSMO file. The available COSMO database already contains COSMO files of several hundred common solvents as well as thousands of other chemicals. A detailed description of the COSMO model and its extensions can be found in the literature.[43]

Figure 3 shows the σ-profiles of four representative solvents, namely water, hexane, chloroform, and acetone. In general, the whole SCD range can be divided into three different regions, that is, the nonpolar region (-1.0 e/nm$^2 \le \sigma \le 1.0$ e/nm^2), the electronic acidity region ($\sigma < -1.0$ e/nm^2), and the electronic basicity region ($\sigma > 1.0$ e/nm^2). As depicted, water has a very broad σ-profile distribution with two pronounced peaks around -1.6 and $+1.8$ e/nm^2, resulting from the two strongly polar hydrogen atoms and lone-pairs of the oxygen atom, respectively. Hexane is almost a nonpolar compound, which is reflected by the narrow distribution of the charge densities around zero. The slightly positive σ can be assigned to the hexane carbon atoms and the slightly negative σ denotes the hydrogen atoms (Please note that negative partial charges of atoms cause positive screening charges and vice versa). The σ-profile of acetone is asymmetric. The peak around $+1.3$ e/nm^2 arises from its carbonyl group. As for chloroform, the three peaks in the nonpolar region are resulted from the chlorine atoms and the peak at -1.4 e/nm^2 corresponds to the acidic hydrogen atom.

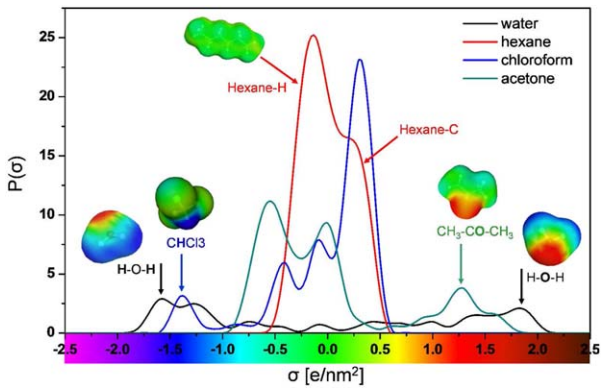

Figure 3. σ-profiles of four representative solvents: nonpolar (hexane), highly polar HB acceptor and HB donor (water), HB acceptor (acetone), and HB donor (chloroform).

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Derived from unimolecular QC calculations, σ-profiles can well characterize the electrostatic polarity and charge distribution of a solvent molecule. With the main advantage of providing fundamental information at the molecular level, σ-profiles have been successfully applied to a variety of areas, such as the interpretation of complex molecular interactions,[44] stereoisomer differentiation,[45] estimation of physical properties of ionic liquids,[46] prediction of drug similarity,[47] and so forth. More importantly, the molecule-specific characteristics of σ-profile make it possible to get not only qualitative but also quantitative information on solvent properties and behaviors. In this work, solvent σ-profile curves were divided into six segments in the entire [–2.5 e/nm^2, +2.5 e/nm^2] σ-region. Six areas (S_1 to S_6) were obtained for each solvent by integrating those segments over the SCD, as clearly depicted in Figure 4. Afterward, the six area parameters were used as solvent theoretical descriptors to quantify their effects on chemical reactions.

The DA reaction is one of the most important carbon–carbon bond forming reactions. The reaction facilitates the rapid development of molecular complexity and thus has been widely used in organic synthesis to prepare cyclic structures.

Table 1. Six COSMO σ-profile Areas of the 15 Experimentally Studied Reaction Solvents, Together with Their Corresponding Experimental Reaction Rate Constants for the Diels–Alder Reaction Between 1,3-Cyclopentadiene and Acrolein at 303 K

Solvent	Log k_{exp}	S_1	S_2	S_3	S_4	S_5	S_6
Acetic acid	−2.491	0.2805	0.5864	4.5564	1.9442	1.8717	0
Ethanol	−2.964	0.0135	0.6676	4.5887	2.3670	1.0127	0.1639
1-Propanol	−3.186	0.0163	0.6709	5.7535	3.2236	1.0294	0.1427
1-Butanol	−3.219	0.0122	0.6794	6.9118	4.0754	1.0128	0.1359
Methanol	−3.257	0.0190	0.7260	3.4825	1.3293	1.0064	0.1927
Chloroform	−3.383	0	0.8256	3.4240	7.5032	0	0
1,2-dichloroethane	−3.602	0	0.4368	4.7253	6.5480	0	0
Dimethylformamide	−3.640	0	0	8.0367	1.6913	1.7447	0.1161
Dichloromethane	−3.699	0	1.1114	2.8389	5.8986	0	0
Toluene	−3.745	0	0	7.1149	6.9400	0	0
Acetonitrile	−3.757	0	0.3018	4.5277	2.0646	1.3673	0
1,4-Dioxane	−3.827	0	0	8.2030	1.9772	1.8345	0.0311
Tetrachloromethane	−3.873	0	0	4.9773	8.4433	0	0
Acetone	−3.983	0	0	6.6600	1.9346	1.6555	0.0129
Ethyl acetate	−4.036	0	0	8.6090	2.9359	1.8029	0.0008

The Experimental Data Are Taken from Ref. 49.

One of the most interesting aspects of this reaction is its pronounced solvent dependency.[48] Blankenburg et al.[49] systematically studied the effects of various solvents on the DA reaction between 1,3-cyclopentadiene and various dienophiles. In this work, the reaction between 1,3-cyclopentadiene and acrolein at 303 K is considered as example reaction where 15 solvents with a wide variation of physical and chemical properties were experimentally investigated. In Table 1, the six σ-profile areas of these solvents are presented together with the corresponding reaction rate constants.

The following linear regression formula between the logarithm of the reaction rate constant and the six solvent area descriptors was determined

$$\log k = -4.0113 + 4.7103S_1 + 0.3012S_2 + 0.0050S_3 + 0.0274S_4 - 0.0292S_5 + 3.1087S_6 \quad (1)$$

The mean absolute percentage error (MAPE), a measure of accuracy in statistics, was used to analyze the regression

$$\mathrm{MAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{\log k_{\mathrm{exp},i} - \log k_{\mathrm{cal},i}}{\log k_{\mathrm{exp},i}}\right| \quad (2)$$

where $\log k_{\mathrm{exp},i}$ and $\log k_{\mathrm{cal},i}$ are the logarithms of the experimental and calculated reaction rate constants in solvent i, respectively, and N is the total number of solvent data points.

The calculated MAPE is 2.58% and the coefficient of determination, R^2, is 0.9231. These statistical indicators show that the COSMO area parameters are good solvent descriptors to quantify the kinetic solvent effects. For better illustration, we compared the calculated values, log k_{cal}, with the experimental values, log k_{exp} in Table 2. Despite some quantitative deviations, the regression model (Eq. 1) can be used for solvent ranking with respect to reaction rates. The correlation predicts the same first five solvents as the top five experimentally ranked solvents, although the order is not exactly correct. These findings suggest that the kinetic regression model can be further used for solvent prescreening and also for the molecular design of solvents.

Reaction Solvent Design

GCs to solvent descriptors

With the obtained reaction kinetic model that correlates reaction rate constants to solvent COSMO descriptors in a quantitative way, we now need a relationship that links these descriptors to solvent molecular structures. GC methods are widely used for the estimation of solvent physical properties based on their molecular structures.[50–52] The GC approach is developed on the assumption that the contributions of individual groups on solvent properties are additive. From Figure 3, we can see that the σ-profile of a molecule is closely related to the atom groups of the molecule. In other words, a certain group should contribute a specific part of the σ-profile. This inspires us to try to use the GC approach to estimate the solvent descriptors used in Eq. 1. As depicted in Eq. 3, S_i ($i = 1, 2 \ldots 6$) are the six solvent σ-profile areas, s_{ij} is the contribution of the j-th group to the i-th σ-profile area and s_{i0} are the fitting constants. n_j denotes the number of group j in the molecule while N is the total number of different groups. The contributions of individual groups to solvent COSMO descriptors were obtained from the multiple linear regression of the σ-profile areas of the 168 common solvents from the Minnesota Solvent Set.[53] The σ-profiles of these solvents were taken directly from the latest COSMO solvent database (C30-1201). As molecular structural groups, the UNIFAC (Dortmund) groups were used[54]

$$S_i = s_{i0} + \sum_{j=1}^{N} n_j s_{ij} \quad (i = 1, 2, \ldots, 6) \quad (3)$$

Table 2. Experimental Log k_{exp}, Calculated Log k_{cal}, and Associated Solvent Rankings of the 15 Solvents Used for the Diels–Alder Reaction

Solvent	$k \times 10^5$ L/(mol s)	Log k_{exp}	Log k_{cal}	Calculated Ranking
Acetic acid	322.8	−2.491	−2.492	1
Ethanol	108.7	−2.964	−3.179	3
1-Propanol	65.2	−3.186	−3.202	4
1-Butanol	60.4	−3.219	−3.210	5
Methanol	55.3	−3.257	−3.080	2
Chloroform	41.4	−3.383	−3.540	7
1,2-dichloroethane	25	−3.602	−3.677	9
Dimethylformamide	22.9	−3.640	−3.615	8
Dichloromethane	20	−3.699	−3.501	6
Toluene	18	−3.745	−3.786	11
Acetonitrile	17.5	−3.757	−3.881	13
1,4-Dioxane	14.9	−3.827	−3.873	12
Tetrachloromethane	13.4	−3.873	−3.755	10
Acetone	10.4	−3.983	−3.933	14
Ethyl acetate	9.2	−4.036	−3.938	15

Table 3. Seventy-One UNIFAC Structural Groups and Their Contributions to Solvent COSMO Descriptors

	s_{10}	s_{20}	s_{30}	s_{40}	s_{50}	s_{60}
	−0.00202	−0.15091	3.58025	−0.91960	0.38711	0.07294
Group	s_{1j}	s_{2j}	s_{3j}	s_{4j}	s_{5j}	s_{6j}
CH_3	0.00089	0.02073	0.53888	2.06187	−0.17281	−0.02667
CH_2	−0.00004	0.00742	1.05649	0.91999	−0.00238	−0.00132
CH	−0.00208	0.01217	1.14222	−0.18643	0.14520	0.02145
C	−0.00437	0.01606	1.51306	−1.85052	0.30961	0.07360
CH_2═CH	0.00419	0.11012	0.99885	2.94047	−0.10216	−0.05282
CH═CH	0.00028	0.10202	1.54385	1.85037	0.04173	−0.01827
CH═C	−0.00059	−0.60012	3.68092	−0.35402	0.25616	0.00906
ACH	0.00042	0.01990	0.34061	1.25012	−0.06168	−0.01246
AC	0.00001	−0.06217	0.98538	−0.25030	0.10503	0.01567
$ACCH_3$	0.00023	0.03093	1.54454	1.86466	−0.06579	−0.01173
$ACCH_2$	0.00358	0.01361	2.45084	0.24363	0.10191	−0.00404
ACCH	−0.00176	0.00379	2.69615	−1.08616	0.26908	0.04291
OH	0.01484	0.71461	−0.31120	0.31695	0.77063	0.08152
CH_3OH	0.02102	0.87691	−0.09780	2.24890	0.61929	0.11971
H_2O	0.02952	1.75981	−3.01045	1.61070	0.63294	0.31571
ACOH	0.21121	0.66711	−0.27753	1.26076	0.61421	−0.01135
CH_3CO	0.00121	0.11035	2.57475	0.77126	1.31332	−0.02939
CH_2CO	0.00056	0.07364	3.11906	−0.47464	1.43224	−0.01007
CHO	0.00121	0.11534	1.38480	0.13605	1.19788	−0.04322
CH_3COO	0.00121	0.11534	3.61839	0.72511	1.57978	−0.04302
CH_2COO	0.00026	0.10573	3.87689	−0.18063	1.62852	−0.01871
HCOO	0.00115	0.29929	1.66505	0.58384	1.44054	−0.04561
OCH_3	−0.00610	0.00842	1.61665	0.79532	0.87103	−0.03756
OCH_2	−0.00008	0.08996	2.55219	−0.17182	0.75078	0.03014
OCH	0.00054	0.05580	2.67201	−1.11185	0.79676	0.16094
CH_2NH_2	0.00121	0.49071	1.39382	1.34315	0.22805	0.34375
CH_3NH	−0.00006	0.71858	1.29259	1.83777	0.11208	−0.02630
CH_2NH	0.00032	0.23948	2.07790	0.64838	0.29887	0.27948
$ACNH_2$	−0.00005	1.24791	0.12337	1.10092	0.35031	0.01823
AC_2H_2N	0.00054	0.12960	1.19015	2.17970	0.49286	0.22744
AC_2HN	−0.00011	0.07577	1.77370	0.92383	0.58884	0.25392
AC_2N	−0.00100	0.05084	2.44362	−0.38454	0.73532	0.30697
CH_3CN	0.00202	0.45276	0.94750	2.98415	0.98014	−0.07294
CH_2CN	0.00115	0.19507	1.97855	1.56929	1.14896	−0.04561
COOH	0.26792	0.70732	0.53942	0.72350	1.59226	−0.04357
HCOOH	0.41207	0.94666	−1.07100	2.75785	1.28849	−0.07294
CH_2Cl	0.00096	0.20956	1.00841	3.42696	−0.19105	−0.03755
CHCl	0.00028	0.10202	1.65035	1.84147	0.01193	−0.01827
CH_2Cl_2	0.00202	1.26231	−0.74140	6.81825	−0.38711	−0.07294
CCl_2	0.00106	1.26480	−0.67746	5.75529	−0.19605	−0.03539
$CHCl_3$	0.00202	0.97646	−0.15620	8.42280	−0.38711	−0.07294
CCl_3	0.00113	0.18487	1.54257	6.44233	−0.21430	−0.04627
CCl_4	0.00202	0.15091	1.39700	9.36285	−0.38711	−0.07294
ACCl	0.00013	0.07639	0.58193	2.74455	−0.07234	−0.01136
CH_3NO_2	0.00202	0.93386	0.25205	4.20715	0.54659	−0.07294
CH_2NO_2	0.00115	0.48784	1.55762	2.63991	0.70929	−0.04561
$CHNO_2$	0.00024	0.13509	2.64094	1.11861	0.79805	−0.01960
$ACNO_2$	0.00004	0.12734	1.79831	1.47803	0.72343	−0.01099
CH_2SH	0.00113	0.23727	1.55512	2.73973	0.47355	−0.04627
I	0.00098	0.31142	0.28793	3.42427	−0.19957	−0.03852
Br	0.00122	0.29019	−0.07726	3.01428	−0.18491	−0.03482
CH≡C	0.00125	0.54523	0.51504	2.55271	−0.00825	−0.04230
DMSO	0.00202	0.33316	3.81845	2.46050	0.91024	0.67446
Cl─C═C	0.00087	0.43079	−0.80392	2.99129	−0.21442	−0.02733
C═C	−0.00146	−1.57226	5.73254	−2.39786	0.47058	0.03639
ACF	0.00033	0.02723	0.42266	1.84768	−0.06490	−0.01211
CF_3	0.16827	0.40553	−1.61744	5.88581	−0.34510	−0.15314
CY─CH_2	0.00033	0.02789	0.58224	1.20095	−0.06829	−0.01175
CY─CH	−0.00061	−0.02424	0.82277	−0.09242	0.09431	0.01277
$CONH_2$	0.01222	1.76786	−1.14500	2.61975	1.40994	0.04246
$HCONHCH_3$	0.00472	0.57090	2.48956	2.20841	1.36751	0.06107
ACCN	−0.00005	0.18321	1.73137	0.97682	1.03356	−0.01062
$CON(CH_3)_2$	0.00113	0.13017	4.74422	1.40633	1.39585	0.17573
AC_2H_2S	0.00119	0.26076	0.64009	4.78302	−0.26279	−0.04801
$(CH_2)_2SU$	0.00135	0.28258	4.35993	0.03511	2.91937	−0.04945
ACCHO	−0.00005	0.05471	2.65267	1.24612	0.44596	−0.01062
ACCOO	−0.00094	0.03238	3.42309	−0.62785	1.44816	0.01605
H_2COCH_2	0.00108	0.07939	2.41823	1.36229	0.72119	−0.00152
CH_2S	0.00028	0.10202	2.73840	0.81752	0.77933	−0.01827
CFH_2	0.00136	0.08568	1.60542	1.54163	0.04025	−0.03832
─S─S─	0.00024	0.12354	2.11729	2.52571	0.31405	−0.01960

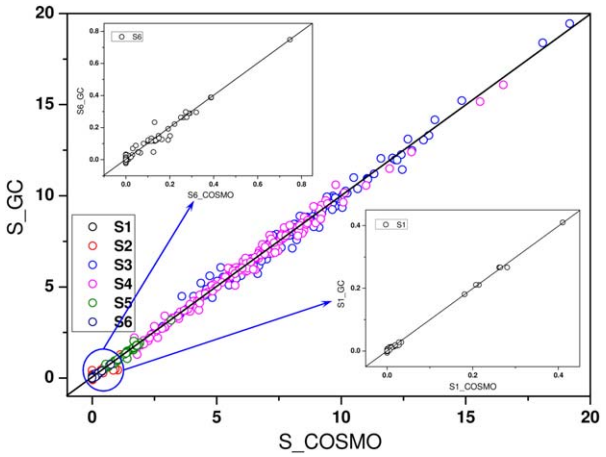

Figure 5. GC predicted σ-profile areas plotted with the original COSMO σ-profile areas for the 168 solvents.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The regressed GCs to each solvent COSMO descriptor, together with the fitting constants, are presented in Table 3.

For illustrating the quality of the regression, six σ-profile areas of the 168 solvents estimated from the GC method are plotted versus their original COSMO areas in Figure 5. The overall MAPE between the GC areas and the original COSMO areas is 7.28%. It can be concluded that the solvent COSMO descriptors can be estimated from the proposed GC method (Eq. 3) at relatively high accuracy.

Computer-aided solvent design for optimal reaction rates

With the kinetic regression model (Eq. 1) which connects the reaction rate constant to the six solvent descriptors and the GC method (Eq. 3) which can be used to predict the solvent descriptors, a CAMD method now can be established to design solvents for maximizing the reaction rate constant. The molecular design of solvents is based on the systematic combination of numerous functional groups to develop a feasible chemical structure of particular physical and/or chemical properties. With regard to environmental compatibility, only hydrocarbons, esters, ketones, alcohols, and ethers are considered as solvents in this work. Consequently, the group basis set G from which solvents can be built was limited to the following molecular groups: CH_3, CH_2, CH, C, OH, CH_3CO, CH_2CO, CHO, CH_3COO, CH_2COO, HCOO, OCH_3, OCH_2, OCH, and COOH. Every generated solvent molecule is represented by a composition vector $\mathbf{n} = [n_1, \ldots, n_j, \ldots, n_N]$ with the element n_j indicating the number of group j presented in the molecule.

The generated molecules need to satisfy several structural constraints. The octet rule (Eq. 4) is used to ensure that the molecule has zero valency and the modified bonding rule (Eq. 5) is used to ensure that two adjacent groups in a molecule are not linked by more than one bond. The maximum number of each group is specified to result in a commonly used solvent (Eq. 6). Considering the size and structural complexity of typical solvent molecules, the total number of groups making up a molecule and the total number of functional groups are also limited, as given in Eqs. 7 and 8, respectively. The group classes, valences, and maximum numbers are listed in Table 4.

The CAMD problem is formulated as follows

$$\max_{n_j(j\in G)} \{\log k\}$$

s.t.

- Reaction rate constant equations: Eqs. 1 and 3
- Chemical feasibility constraints

$$\sum_{j=1}^{N} (2-v_j) n_j - 2m = 0 \tag{4}$$

$$n_j(v_j-1) + 2m - \sum_{k=1}^{N} n_k \le 0 \quad (j=1\ldots N) \tag{5}$$

- Chemical complexity constraints

$$0 \le \text{integer}\, n_j \le N_{\max}[j], \quad (\text{group}\, j \in G1\text{–}G15) \tag{6}$$

$$\begin{gathered} 2 - \sum_{j=1}^{N} n_j \le 0 \\ \sum_{j=1}^{N} n_j \le 5 \end{gathered} \tag{7}$$

$$\begin{gathered} \sum_{j\in G_F} n_j \le 1 \quad \text{if} \sum_{j=1}^{N} n_j = 2 \quad or \sum_{j\in\{G_{12},G_{13},G_{14}\}} n_j \ge 1 \\ \sum_{j\in G_F} n_j \le 2 \quad \text{else} \end{gathered} \tag{8}$$

v_j stands for the valence of group j. N is the total number of different groups (in our case $N = 15$). $m = +1$, 0, -1 indicate acyclic, monocyclic, and bicyclic structures, respectively. Due to biocompatibility concerns only acyclic compounds are considered in this work. The solution of the above CAMD optimization problem is an optimal solvent structure having the highest reaction rate constant. This optimal solvent, shown in Table 5, was identified in a few iterations at low computational cost. However, the modest number of all feasible structures makes it possible to perform the full evaluation of all solvents in the design space. Thereby, 25 solvents were identified with larger reaction rate

Table 4. Group ID, Class, Valence, and Maximum Number of the 15 Investigated Groups

Group j	Group ID and Class	Group Valence v_j	Maximum Group Number, $N_{\max}[j]$
CH_3	$G1$, G_M	1	4
CH_2	$G2$, G_M	2	3
CH	$G3$, G_M	3	1
C	$G4$, G_M	4	1
OH	$G5$, G_F	1	2
CH_3CO	$G6$, G_F	1	1
CH_2CO	$G7$, G_F	2	1
CHO	$G8$, G_F	1	1
CH_3COO	$G9$, G_F	1	1
CH_2COO	$G10$, G_F	2	1
HCOO	$G11$, G_F	1	1
OCH_3	$G12$, G_F	1	1
OCH_2	$G13$, G_F	2	1
OCH	$G14$, G_F	3	1
COOH	$G15$, G_F	1	1

Table 5. Predicted Top 10 Reaction Solvents and Their Corresponding Log k_{pred} Values

Ranking	Group Combination	Solvent	Log k_{pred}
1	2 CH_3, 1 C, 1 OH, 1 COOH	$(CH_3)_2C(OH)COOH$	−1.875
2	3 CH_2, 1 OH, 1 COOH	$OHCH_2CH_2CH_2COOH$	−1.934
3	1 CH_3, 1 CH_2, 1 CH, 1 OH, 1 COOH	$CH_3CH(OH)CH_2COOH$	−1.942
4	2 CH_2, 1 OH, 1 COOH	$OHCH_2CH_2COOH$	−1.962
5	1 CH_3, 1 CH, 1 OH, 1 COOH	$CH_3CH(OH)COOH$	−1.971
6	1 CH_2, 1 OH, 1 COOH	$OHCH_2COOH$	−1.991
7	3 CH_3, 1 C, 1 COOH	$(CH_3)_3CCOOH$	−2.406
8	2 CH_3, 1 C, 1 CH_3CO, 1 COOH	$CH_3COC(CH_3)_2COOH$	−2.455
9	2 CH_3, 1 C, 1 HCOO, 1 COOH	$HCOOC(CH_3)_2COOH$	−2.462
10	1 CH_3, 3 CH_2, 1 COOH	$CH_3(CH_2)_3COOH$	−2.465

constants than the best experimentally found solvent acetic acid. The top 10 ranked solvents and their corresponding log k values are listed in Table 5. Carboxyl acids and alcohols are found to be the most promising solvent types for accelerating the here investigated DA reaction. This is in good agreement with reported kinetic solvent effects[49] where acetic acid was found to be the best solvent and the second to fifth solvents were all alcohols. The solvents in Table 5 can be targeted for further experimentations.

Integrated Solvent and Process Design

It has been proven that the choice of an appropriate solvent can significantly improve the performance of a chemical reaction. However, one should note that promising solvents for reactions may not be the best choices for the whole chemical process. From the conceptual design point of view, the selection of solvents should be integrated into a process-wide design framework and the criterion for making the final decision should be always based on the performance or cost of the overall process. The classical two-stage design method (molecular design followed by process design) and later developed decomposed solvent and process design approaches cannot fully reflect the strong interdependencies between solvent properties and the process performance. Thus, simultaneous solvent and process design is highly desirable to overcome this limitation. Typically, the integrated solvent and process design problem can be formulated as an MINLP optimization problem where solvent molecular structures are represented by integer variables and the process operation conditions appear as continuous variables. Thanks to recently developed efficient MINLP solution algorithms,[55,56] and steadily enhanced computer power, optimization-based programming methods can now successfully combine molecular design and process optimization for finding the optimal solvent structure and the best process operation conditions.

Solvent property estimation

Solvent effects on chemical processes are directly related to their physical properties. To develop an efficient chemical process, trade-offs between different solvent properties need to be addressed based on their effects on the process performance. In this work, several GC methods are used to estimate solvent physical properties. In most GC methods, the contribution toward any property made by a certain group is assumed to be independent of that made by another group. The only information required for estimating a certain property of a molecule is the number and type of structural groups in the molecule. The used GC methods are expressed by Eqs. 9–15. The GCs to different solvent properties are summarized in Table 6.

Table 6. A Summary of Different Group Contribution Methods Used in This Work

Group Contributions (GC) to Solvent Properties	Sources
GC to COSMO σ-profile areas s_{ij}	Table 3 in this work
GC to boiling point $t_{b,j}$, to critical temperature $t_{c,j}$, to critical pressure $p_{c,j}$, and to standard enthalpy of vaporization $\Delta h_{vap,j}$	Ref. 52
GC to liquid molar volume $v_{m,j}$	Ref. 51
GC to heat capacity constants a_j, b_j, c_j, d_j	Ref. 50

$$\sigma\text{-profile areas } (\mathrm{e/nm^2}): \quad S_i = s_{i0} + \sum_{j=1}^{N} n_j s_{ij} \quad (i = 1, 2, \ldots, 6) \tag{9}$$

$$\text{Boiling temperature (K)}: \quad T_b^S = 222.543 \ln\left(\sum_{j=1}^{N} n_j t_{b,j}\right) \tag{10}$$

$$\text{Critical temperature (K)}: \quad T_c^S = 231.239 \ln\left(\sum_{j=1}^{N} n_j t_{c,j}\right) \tag{11}$$

$$\text{Critical pressure (bar)}: \quad P_c^S = \left(0.108998 + \sum_{j=1}^{N} n_j p_{c,j}\right)^{-2} + 5.9827 \tag{12}$$

$$\text{Standard enthalpy of vaporization (kJ/mol)}: \quad \Delta H_{vap}^S = \left(11.733 + \sum_{j=1}^{N} n_j \Delta h_{vap,j}\right) \times 1000 \tag{13}$$

$$\text{Liquid molar volume (L/mol)}: \quad V_m^S = 0.01211 + \sum_{j=1}^{N} n_j v_{m,j} \tag{14}$$

$$\text{Heat capacity (cal/mol} \cdot \text{K)}: \quad C_p^S = \sum_{j=1}^{N} n_j a_j + \left(\sum_{j=1}^{N} n_j b_j\right) \cdot T + \left(\sum_{j=1}^{N} n_j c_j\right) \cdot T^2 + \left(\sum_{j=1}^{N} n_j d_j\right) \cdot T^3 \tag{15}$$

Process models

A simple flow sheet is constructed for a continuous DA reaction process (Figure 6). This process contains a continuous stirred tank reactor (CSTR), a heat exchanger, and a distillation column, including the condenser and the reboiler.

Table 7. Optimization Results of the Integrated Solvent and Process Design and the Reference Design Problems for the Diels–Alder Reaction Process

	Integrated Solvent and Process Design (MINLP)	Reference Case (NLP)
Solvent	Isopropanol	Acetic Acid
Reaction conversion X	0.755	0.872
Product recovery ξ_{nor}	0.0055	0.0223
Total profit (US $/year)	27918.1	22776.4
EP (US $/year)	65319.7	74118.8
CI_R (US $/year)	10145.0	5750.3
CI_{heater} (US $/year)	654.3	670.6
CI_{column} (US $/year)	7662.7	17133.2
CI_{reb} (US $/year)	705.2	712.7
CI_{cond} (US $/year)	725.1	705.9
UC_R (US $/year)	2340.9	2701.7
UC_{heater} (US $/year)	1261.4	2325.4
UC_{reb} (US $/year)	13399.3	20816.1
UC_{cond} (US $/year)	507.7	526.5
Log k (L/mol · s)	−3.294	−2.551
V_R (m^3)	1.90	1.08
N_T (number of trays)	4	15
Reflux ratio R	0.128	0.920

As depicted, the reactants acrolein (acro) and 1,3-cyclopentadiene (cyclo), together with certain amount of solvent, are fed into the reactor. After the reaction, the mixture is heated and fed into the distillation column. The unconverted reactants and solvents are separated from the product 5-Norbornane-2-carboxaldehyde (Norbo) and sent back into the reactor. Nearly pure product is collected from the bottom of the column. Detailed process models that are used to derive the MINLP optimization problem are presented in the following sections.

Reactor. According to the experimental reaction conditions, the component molar flow rates of the reactor inlet stream, $\dot{N}_{in,cyc}$, $\dot{N}_{in,acr}$, and $\dot{N}_{in,S}$, are set to 1, 1, and 2 mol/s, respectively. The reaction conversion not only decides the production rate but also affects the reactor volume as well as the size and utility costs (UCs) of the subsequent separation system. It must be carefully selected to optimally balance process revenue, reactor cost, and separation cost. In this work, the single-pass conversion of the reaction, X, is used as design variable to be optimized. As a result, the component molar flow rates of the reactor outlet stream can be determined as follows

$$\dot{N}_{out,cyc}=\dot{N}_{in,cyc}(1-X) \tag{16}$$

$$\dot{N}_{out,acr}=\dot{N}_{in,acr}(1-X) \tag{17}$$

$$\dot{N}_{out,nor}=\dot{N}_{in,cyc}X \tag{18}$$

$$\dot{N}_{out,S}=\dot{N}_{in,S} \tag{19}$$

The subscripts cyc, acr, nor, and S represent 1,3-cyclopentadiene, acrolein, 5-norbornane-2-carboxaldehyde, and the solvent, respectively. The concentration of compound i in the outlet of the CSTR can be calculated as

$$C_i=\dot{N}_{out,i}/\sum_{i\in COM}\dot{N}_{out,i}V_m^i \tag{20}$$

COM includes S, cyc, acr, and nor. The solvent molar volume V_m^S (L/mol) is estimated from Eq. 14. Molar volumes of the reactants and product V_m^i (i = cyc, acr, and nor) are presented in Supporting Information Table S1 of Appendix C. The reaction was reported to be an irreversible second-order reaction with respect to the reactants.[49] Therefore, the required reactor volume V_R (m^3) can be expressed by

$$V_R=\dot{N}_{out,nor}\times 0.001/\left(kC_{cyc}C_{acr}\right) \tag{21}$$

The heat duty of the reactor cooling water can be evaluated by

$$Q_R=\Delta H_R^{303.15}\dot{N}_{in,cyc}X \tag{22}$$

where the estimated reaction enthalpy change is $\Delta H_R^{303.15}$ = −165.75 kJ/mol.

Heat Exchanger. The composition of the feed stream of the distillation column can be calculated as follows

$$\dot{N}_{F,i}=\dot{N}_{out,i} \tag{23}$$

$$x_{F,i}=\dot{N}_{F,i}/\sum_{i\in COM}\dot{N}_{F,i} \tag{24}$$

The pressure of the feed stream is set to p_F = 1.1 bar. The bubble point feed condition can be written as

$$P_{sat}^j(T_F)=p_F/\sum_{i\in COM}\left(\alpha_{ij}^F\cdot x_{F,i}\right) \tag{25}$$

where the temperature-dependent relative volatility is calculated from the ratio of the vapor pressures of the two components

$$\alpha_{ij}^F=P_{sat}^i(T_F)/P_{sat}^j(T_F) \tag{26}$$

Methods for calculating the temperature dependences of vapor pressures of the solvent, reactants, and product are given in Appendix A. The heat duty of the heat exchanger can be evaluated after the feed temperature T_F is determined

$$Q_H=\int_{T_0}^{T_F}(C_p^S\dot{N}_{F,S}+\sum_{i\in Species}C_p^i\dot{N}_{F,i})dT \tag{27}$$

where Species includes cyc, acr, and nor, and T_0 is 303 K. Solvent heat capacity C_p^S is estimated from Eq. 15. Heat capacities of the reactants and product C_p^i are also expressed as functions of temperature with their heat capacity constants listed in Supporting Information Table S2

$$C_p^i=a_i+b_iT+c_iT^2+d_iT^3 \tag{28}$$

Distillation Column. Due to the complex nonlinear structure of the process design problem combined with the large discrete molecular design space, the resulted MINLP problem is computationally demanding. For simplifying the

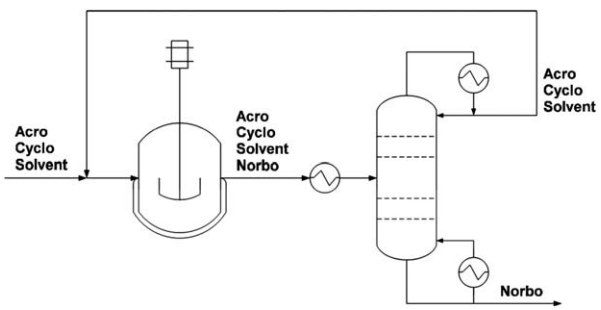

Figure 6. Flow sheet diagram of the DA reaction process.

The reaction is Acrolein (Acro) + 1,3-Cyclopentadiene (Cyclo) to 5-Norbornane-2-carboxaldehyde (Norbo).

calculation, the distillation column is represented by a shortcut model including several assumptions such as equilibrium stages, ideal vapor and liquid phases, and constant relative volatilities throughout the column. The pressure at the top of the column p is set to 1 bar and the total pressure drop in the column Δp is assumed to be 0.2 bar. To meet the product purity requirement, the recovery of the solvent (light key component) in the distillate ξ_S is specified to 99.99%. As a key process operation condition, the recovery of the product (heavy key component) ξ_{nor} in the distillate is set as an optimization control variable. The recoveries of the reactants, ξ_{cyc} and ξ_{acr}, are considered to be 100% due to the sharp split assumption. Moreover, to avoid azeotrope formation and to insure a sufficiently large relative volatility for the solvent and product pair, a minimum normal boiling point difference of 30 K between the solvent and the product is specified.

The molar compositions of the distillate and bottom streams are calculated as

$$\dot{N}_{D,i}=\dot{N}_{F,i}\cdot\xi_i \tag{29}$$

$$\dot{N}_{B,i}=\dot{N}_{F,i}-\dot{N}_{D,i} \tag{30}$$

$$x_{D,i}=\dot{N}_{D,i}/\sum_{i\in COM}\dot{N}_{D,i} \tag{31}$$

$$x_{B,i}=\dot{N}_{B,i}/\sum_{i\in COM}\dot{N}_{B,i} \tag{32}$$

Bubble point calculation at the bottom of the column

$$\alpha_{ij}^{B}=P_{sat}^{i}(T_{bub})/P_{sat}^{j}(T_{bub}) \tag{33}$$

$$P_{sat}^{j}(T_{bub})=(p+\Delta p)/\sum_{i\in COM}\left(\alpha_{ij}^{B}\cdot x_{B,i}\right) \tag{34}$$

Dew point calculation at the top of the column

$$\alpha_{ij}^{D}=P_{sat}^{i}(T_{dew})/P_{sat}^{j}(T_{dew}) \tag{35}$$

$$P_{sat}^{j}(T_{dew})=p\cdot\sum_{i\in COM}\left(\frac{x_{D,i}}{\alpha_{ij}^{D}}\right) \tag{36}$$

The average relative volatility of the solvent-product pair throughout the column is defined by

$$\alpha_{S,nor}=\sqrt{\alpha_{S,nor}^{D}\cdot\alpha_{S,nor}^{B}} \tag{37}$$

The Westerberg method[57] is used to provide a rough estimation of the column size and reflux ratio. The arbitrary weights γ_N and γ_R are set to 0.8 in this work

$$N_{lk}=12.3/\left(\left(\alpha_{S,nor}-1\right)^{2/3}\cdot\left(1-\xi_S\right)^{1/6}\right) \tag{38}$$

$$N_{hk}=12.3/\left(\left(\alpha_{S,nor}-1\right)^{2/3}\cdot{\xi_{nor}}^{1/6}\right) \tag{39}$$

$$R_{lk}=1.38/\left(\left(\alpha_{S,nor}-1\right)^{0.9}\cdot\left(1-\xi_S\right)^{0.1}\right) \tag{40}$$

$$R_{hk}=1.38/\left(\left(\alpha_{S,nor}-1\right)^{0.9}\cdot{\xi_{nor}}^{0.1}\right) \tag{41}$$

$$N_T=\gamma_N\max\{N_{lk},N_{hk}\}+(1-\gamma_N)\min\{N_{lk},N_{hk}\} \tag{42}$$

$$R=\gamma_R\max\{R_{lk},R_{hk}\}+(1-\gamma_R)\min\{R_{lk},R_{hk}\} \tag{43}$$

Enthalpy of vaporization of the stream mixture at the condenser and reboiler are expressed as molar weight sum of the enthalpies of vaporization of the individual compounds

$$\Delta H_{vap}^{D}=\sum_{i\in COM}x_{D,i}\Delta H_{vap}^{i}(T_{dew}) \tag{44}$$

$$\Delta H_{vap}^{B}=\sum_{i\in COM}x_{B,i}\Delta H_{vap}^{i}(T_{bub}) \tag{45}$$

where the temperature dependence of the enthalpy of vaporization[58] is expressed by

$$\Delta H_{vap}^{i}(T)=\Delta H_{vap,298.15}^{i}\left(\frac{1-T/T_c^i}{1-298.15/T_c^i}\right)^{0.375} \tag{46}$$

The standard enthalpy of vaporization (ξ) and critical temperature (K) of the solvent are estimated by GC methods, as given in Eqs. 13 and 11. These two properties of the reactants and product are provided in Supporting Information Table S1 in Appendix C.

Finally, the heat duties of the condenser and the reboiler can be determined

$$Q_{cond}=\Delta H_{vap}^{D}(R+1)\sum_{i\in COM}\dot{N}_{D,i} \tag{47}$$

$$Q_{reb}=\Delta H_{vap}^{B}(R+1)\sum_{i\in COM}\dot{N}_{D,i} \tag{48}$$

Cost models

The determination of the overall profit of a process requires the evaluation of the economic potential (EP), capital investment (CI), and UC of the process. The detailed calculation of these factors for the DA reaction process is presented in Appendix B.

MINLP optimization

The integrated solvent and process design problem is formulated as a MINLP. The integer variables n_j represent the number of group j in the solvent molecule while the continuous variables are the reaction conversion X and the product recovery in the distillate stream ξ_{nor}. The maximization of the annual profit of a process is applied as objective function in this process design task. The final MINLP optimization problem can be summarized as follows:

Maximize: profit = $EP-CI_{tot}-UC_{tot}$

Variables: Integer n_j (group $j \in G1 - G15$);
Reaction conversion X and product recovery ξ_{nor}

Subject to: Variable boundaries:

$0 \le \xi_{nor} \le 0.2$
$0.50 \le X \le 0.99$
$0 \le n_j \le N_{max}[j]$, group $j \in G1 - G15$

Solvent structure constraints:

Chemical feasibility: Eqs. 4 and 5
Chemical complexity: Eqs. 7 and 8

Solvent property constraints:

Reaction rate constant: Eq. 1
Physical property estimation: Eqs. 9–15
Property boundary: $330 \le T_b^S \le 416$; $\alpha_{S,nor} > 1$

Process models:

Reactor: Eqs. 16–22
Heat exchanger: Eqs. 23–28
Distillation column: Eqs. 29–48
Vapor pressure calculation: Eqs. A1–A7

Cost models:

Economic potential: Eq. B1
Capital investment: Eqs. B2–B9
Utility cost: Eqs. B10–B14

The solvent molecular structure, the reaction conversion X, and the product recovery ξ_{nor} were simultaneously optimized through the solution of the above MINLP problem using the KNITRO[59] solver in the AMPL[60] modeling environment. For better illustration of the results, a reference case is optimized where acetic acid, the best reaction solvent determined experimentally, is fixed as the solvent. Because the solvent is fixed, the reference design case becomes a nonlinear programming problem. Table 7 shows the optimization results of the integrated solvent and process design, and the reference design problems for the DA reaction process. Isopropanol is found to be the best solvent in this case. The corresponding optimal reaction conversion is 0.755 and product recovery is 0.0055. When comparing this integrated design case to the reference case, it is clear that both the lower process EP and the higher reactor investment cost CI_R are resulted from the much lower reaction rate of isopropanol. But, although isopropanol shows relatively low reaction efficiency, it dramatically improves the energetic performance of the solvent-product separation. In contrast, despite the excellent reaction performance of acetic acid, it is not the best choice for the continuous process because of its extremely high energy consumption in the distillation column. This result highlights the importance of investigating trade-offs among different solvent properties to achieve an overall highest process performance. It can be concluded from the data in Table 7 that the proposed integrated solvent and process design framework allows the identification of the optimal solvent and optimal process conditions which would lead to an improvement of the overall annual profit by more than 20%, compared to the reference case.

Finally, it should be mentioned that, due to the nonconvexity of the MINLP problem, the proposed methodology does not guarantee that the resulted solvent represents the global optimum, but a locally optimal solution. However, many initializations were used with the solution converging to the same optimum point.

Conclusion

A method for designing solvents for chemical reactions was developed. It can be quickly and reliably used for the identification of optimal solvents for chemical reactions without the need for extensive experiments. Considering the multiple and complicated effects of solvents on chemical processes, we propose a new and practical methodology for integrated reaction solvent and process design. The solvent molecular structure and the key process variables are simultaneously optimized by the formulation and solution of an MINLP problem whereby the economic performance of the entire process is maximized. The proposed integrated solvent and process design methodology is applicable beyond the example of the here considered DA reaction. Its application is straightforward for general solvent-involved process design problems once a small set of reaction kinetic data is available for a limited number of solvents.

To the best of our knowledge, this work is one of the first attempts to integrate reaction solvent design into a process-wide optimization framework. For simplifying the calculation, a relatively small number of solvents were considered and a simplified process model was used. A larger solvent group set and a more detailed process model would, of course, increase the reliability and applicability of the results. However, as the complexity and the search space of the optimization problem increase, computational efficiency is likely to become a limiting factor. The extension of the proposed methodology toward more complex molecular structures and process configurations will be attainable, once more efficient MINLP solution algorithms and more powerful computer systems are available.

Acknowledgments

This work was conducted partly in cooperation with the Collaborative Research Center SFB/TRR 63 "Integrated Chemical Processes in Liquid Multiphase Systems (InPROMPT)." The financial support from the Deutsche Forschungsgemeinschaft (DFG) is gratefully acknowledged. Moreover, the authors acknowledge the support from the International Max Planck Research School (IMPRS) in Magdeburg/Germany for Teng Zhou, the support of the Max Planck Society in Germany for the Max Planck Partner Group at the East China University of Science and Technology in Shanghai, and the support of the Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP 20120074110008) and the 111 Project in China (B08021).

Literature Cited

1. Reichardt C. *Solvents and Solvent Effects in Organic Chemistry*, 2nd ed. New York: Wiley-VCH, 1988.
2. Brignole EA, Bottini S, Gani R. A strategy for the design and selection of solvents for separation processes. *Fluid Phase Equilib.* 1986; 29:125–132.
3. Gani R, Nielsen B, Fredenslund A. A group contribution approach to computer-aided molecular design. *AIChE J*. 1991;37:1318–1332.
4. Pretel EJ, López PA, Bottini SB, Brignole EA. Computer-aided molecular design of solvents for separation processes. *AIChE J*. 1994;40:1349–1360.
5. Harper PM, Gani R, Kolar P, Ishikawa T. Computer-aided molecular design with combined molecular modeling and group contribution. *Fluid Phase Equilib.* 1999;160:337–347.
6. Harper PM, Gani R. A multi-step and multi-level approach for computer aided molecular design. *Comput Chem Eng*. 2000;24:677–683.
7. Chen B, Lei Z, Li Q, Li C. Application of CAMD in separating hydrocarbons by extractive distillation. *AIChE J*. 2005;51:3114–3121.
8. Karunanithi AT, Achenie LEK, Gani R. A new decomposition-based computer-aided molecular/mixture design methodology for the design of optimal solvents and solvent mixtures. *Ind Eng Chem Res.* 2005;44:4785–4797.
9. Karunanithi AT, Achenie LEK, Gani R. A computer-aided molecular design framework for crystallization solvent design. *Chem Eng Sci.* 2006;61:1247–1260.
10. Odele O, Macchietto S. Computer aided molecular design: a novel method for optimal solvent selection. *Fluid Phase Equilib.* 1993;82: 47–54.
11. Venkatasubramanian V, Chan K, Caruthers JM. Computer-aided molecular design using genetic algorithms. *Comput Chem Eng.* 1994;18:833–844.
12. Sinha M, Achenie LEK, Ostrovsky GM. Environmentally benign solvent design by global optimization. *Comput Chem Eng*. 1999;23: 1381–1394.
13. Maranas CD. Optimal molecular design under property prediction uncertainty. *AIChE J*. 1997;43:1250–1264.
14. Pistikopoulos EN, Stefanis SK. Optimal solvent design for environmental impact minimization. *Comput Chem Eng*. 1998;22:717–733.
15. Buxton A, Livingston AG, Pistikopoulos EN. Optimal design of solvent blends for environmental impact minimization. *AIChE J*. 1999; 45:817–843.

16. Giovanoglou A, Barlatier J, Adjiman CS, Pistikopoulos EN, Cordiner JL. Optimal solvent design for batch separation based on economic performance. *AIChE J*. 2003;49:3095–3109.
17. Marcoulaki EC, Kokossis AC. On the development of novel chemicals using a systematic synthesis approach. Part I. Optimisation framework. *Chem Eng Sci*. 2000;55:2529–2546.
18. Marcoulaki EC, Kokossis AC. On the development of novel chemicals using a systematic optimisation approach. Part II. Solvent design. *Chem Eng Sci*. 2000;55:2547–2561.
19. van Dyk B, Nieuwoudt I. Design of solvents for extractive distillation. *Ind Eng Chem Res*. 2000;39:1423–1429.
20. Karunanithi AT, Acquah C, Achenie LEK, Sithambaram S, Suib SL. Solvent design for crystallization of carboxylic acids. *Comput Chem Eng*. 2009;33:1014–1021.
21. Cheng HC, Wang FS. Computer-aided biocompatible solvent design for an integrated extractive fermentation–separation process. *Chem Eng J*. 2010;162:809–820.
22. Roughton BC, Christian B, White J, Camarda KV, Gani R. Simultaneous design of ionic liquid entrainers and energy efficient azeotropic separation processes. *Comput Chem Eng*. 2012;42:248–262.
23. Kossack S, Kraemer K, Gani R, Marquardt W. A systematic synthesis framework for extractive distillation processes. *Chem Eng Res Des*. 2008;86:781–792.
24. Lek-utaiwan P, Suphanit B, Douglas PL, Mongkolsiri N. Design of extractive distillation for the separation of close-boiling mixtures: solvent selection and column optimization. *Comput Chem Eng*. 2011;35:1088–1100.
25. Papadopoulos AI, Linke P. Multiobjective molecular design for integrated process-solvent systems synthesis. *AIChE J*. 2005;52:1057–1070.
26. Papadopoulos AI, Linke P. Efficient integration of optimal solvent and process design using molecular clustering. *Chem Eng Sci*. 2006;61:6316–6336.
27. Papadopoulos AI, Linke P. Integrated solvent and process selection for separation and reactive separation systems. *Chem Eng Process*. 2009;48:1047–1060.
28. Eden MR, Jorgensen SB, Gani R, El-Halwagi MM. A novel framework for simultaneous separation process and product design. *Chem Eng Process*. 2004;43:595–608.
29. Eljack FT, Eden MR, Kazantzi V, Qin X, El-Halwagi MM. Simultaneous process and molecular design—a property based approach. *AIChE J*. 2007;53:1232–1239.
30. Kim KJ, Diwekar UM. Integrated solvent selection and recycling for continuous processes. *Ind Eng Chem Res*. 2002;41:4479–4488.
31. Pereira FE, Keskes E, Galindo A, Jackson G, Adjiman CS. Integrated solvent and process design using a SAFT-VR thermodynamic description: high-pressure separation of carbon dioxide and methane. *Comput Chem Eng*. 2011;35:474–491.
32. Oyarzún B, Bardow A, Gross J. Integration of process and solvent design towards a novel generation of CO2 absorption capture systems. *Energy Procedia*. 2011;4:282–290.
33. Bardow A, Steur K, Gross J. Continuous-molecular targeting for integrated solvent and process design. *Ind Eng Chem Res*. 2010;49:2834–2840.
34. Gani R, Gomez PA, Folic M, Jimenez-Gonzalez C, Constable DJC. Solvents in organic synthesis: replacement and multi-step reaction systems. *Comput Chem Eng*. 2008;32:2420–2444.
35. Kamlet MJ, Abboud JLM, Abraham MH, 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*. 1983;48:2877–2887.
36. Abraham MH, Doherty RM, Kamlet MJ, Harris JM, Taft RW. Linear solvation energy relationships. Part 37. An analysis of contributions of dipolarity-polarisability, nucleophilic assistance, electrophilic assistance, and cavity terms to solvent effects on t-butyl halide solvolysis rates. *J Chem Soc Perkin Trans*. 1987:2;913–920.
37. Lowrey AH, Cramer CJ, Urban JJ, Famini GR. Quantum chemical descriptors for linear solvation energy relationships. *Comput Chem*. 1995;19:209–215.
38. Zhou T, Qi Z, Sundmacher K. Model-based method for the screening of solvents for chemical reactions. *Chem Eng Sci*. 2014;115:177–185.
39. Folić M, Adjiman CS, Pistikopoulos EN. Design of solvents for optimal reaction rate constants. *AIChE J*. 2007;53:1240–1256.
40. Folic´ M, Adjiman CS, Pistikopoulos EN. Computer-aided solvent design for reactions: maximizing product formation. *Ind Eng Chem Res*. 2008;47:5190–5202.
41. Tomasi J, Mennucci B, Cammi R. Quantum mechanical continuum solvation models. *Chem Rev*. 2005;105:2999–3093.
42. Klamt A, Schüürmann G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J Chem Soc Perkin Trans*. 1993:2;799–805.
43. Eckert F, Klamt A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J*. 2002;48:369–385.
44. Ortega J, Marrero E, Palomar J. Description of thermodynamic behavior of the systems formed by alkyl ethanoates with 1-chloroalkanes using the COSMO-RS methodology contributing with new experimental information. *Ind Eng Chem Res*. 2008;47:3253–3264.
45. Lapkin AA, Peters M, Greiner L, Chemat S, Leonhard K, Liauw MA, Leitner W. Screening of new solvents for artemisinin extraction process using ab initio methodology. *Green Chem*. 2010;12:241–251.
46. Palomar J, Torrecilla JS, Lemus J, Ferro VR, Rodriguez F. A COSMO-RS based guide to analyze/quantify the polarity of ionic liquids and their mixtures with organic cosolvents. *Phys Chem Chem Phys*. 2010;12:1991–2000.
47. Thormann M, Klamt A, Hornig M, Almstetter M. COSMOsim: bioisosteric similarity based on COSMO-RS sigma profiles. *J Chem Inf Model*. 2006;46:1040–1053.
48. Cativiela C, García JI, Mayoral JA, Salvatella L. Solvent effects on endo/exo- and regio-selectivities of Diels-Alder reactions of carbonyl-containing dienophiles. *J Chem Soc Perkin Trans*. 1994:2;847–851.
49. Blankenburg B, Fiedler H, Hampel M, Hauthal HG, Just G, Kahlert K, Korn J, Müller KH, Pritzkow W, Reinhold Y, Röllig M, Sauer E, Schnurpfeil D, Zimmermann G. DIELS-ALDER-Reaktionen. II. Über die Anwendung linearer Freier-Energie-Beziehungen auf DIELS-ALDER-Reaktionen. *J Prakt Chem*. 1974;316:804–816.
50. Rihani DN, Doraiswamy LK. Estimation of heat capacity of organic compounds from group contributions. *Ind Eng Chem Fund*. 1965;4:17–21.
51. Constantinou L, Gani R, O'Connell JP. Estimation of the acentric factor and the liquid molar volume at 298 K using a new group contribution method. *Fluid Phase Equilib*. 1995;103:11–22.
52. Marrero J, Gani R. Group-contribution based estimation of pure component properties. *Fluid Phase Equilib*. 2001;183–184:183–208.
53. Winget P, Dolney DM, Giesen DJ, Cramer CJ, Truhlar DG. Minnesota Solvent Descriptor Database. Available at: http://amsol.chem.umn.edu/solvation/mnsddb.pdf, 2010. Accessed June 6, 2010.
54. Gmehling J, Lohmann J, Jakob A, Li J, Joh R. A modified UNIFAC (Dortmund) model. 3. Revision and extension. *Ind Eng Chem Res*. 1998;37:4876–4882.
55. Belotti P, Kirches C, Leyffer S, Linderoth J, Luedtke J, Mahajan A. Mixed-integer nonlinear optimization. *Acta Numer*. 2013;22:1–131.
56. Grossmann IE. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optim Eng*. 2002;3:227–252.
57. Biegler LT, Grossmann IE, Westerberg AW. *Systematic Methods of Chemical Process Design*. Englewood Cliffs, NJ: Prentice Hall, 1997.
58. Poling BE, Prausnitz JM, O'Connell JP. *The Properties of Gases and Liquids*, 5th ed. New York: McGraw-Hill, 2001.
59. Byrd RH, Nocedal J, Waltz RA. KNITRO: An integrated package for nonlinear optimization, In: *Large-Scale Nonlinear Optimization*. Berlin/Heidelberg, Germany: Springer, 2006.
60. Fourer R, Gay DM, Kernighan BW. *AMPL: A Modeling Language for Mathematical Programming*. Belmont, CA: Duxbury Press, 1993.
61. Nannoolal Y, Rarey J, Ramjugernath D. Estimation of pure component properties: Part 3. Estimation of the vapor pressure of nonelectrolyte organic compounds via group contributions and group interactions. *Fluid Phase Equilib*. 2008;269:117–133.
62. Guthrie KM. Data and techniques for preliminary capital cost estimating. *Chem Eng*. 1969;76:114–142.

Appendix A: Vapor Pressure Calculation

- Solvent vapor pressure calculated from the Clapeyron equation

$$T_r^S = T/T_c^S \tag{A1}$$

$$T_{br}^S = T_b^S/T_c^S \tag{A2}$$

$$h = T_{br}^S \frac{\ln\left(P_c^S/1.01325\right)}{1-T_{br}^S} \tag{A3}$$

$$\ln P^{S}_{vpr} = h\left(1 - \frac{1}{T^{S}_{r}}\right) \quad \text{(A4)}$$

$$P^{S}_{sat}(T) = P^{S}_{vpr} \cdot P^{S}_{c} \quad \text{(A5)}$$

- Vapor pressure of reactants and product from the Nannoolal–Rarey method[61]

$$T^{i}_{rb} = T/T^{i}_{b} \quad \text{(A6)}$$

$$\log_{10}\left(\frac{P^{i}_{sat}(T)}{1\text{atm}}\right) = (4.1012 + dB_i)\left(\frac{T^{i}_{rb} - 1}{T^{i}_{rb} - (1/8)}\right), (i = \text{cyc, acr, and nor}) \quad \text{(A7)}$$

The normal boiling point T_b and equation parameter dB of the reaction species are provided in Supporting Information Table S1.

Appendix B: Process Economic Evaluation

To evaluate the process economics, the EP, CI, and UC of the process need to be calculated. Prices and cost model parameters are presented in Supporting Information Table S3 in Appendix C.

Economic potential ($/year)

The EP per year is calculated from

$$\text{EP} = 122 \times 60 \times 60 \times 24 \times 330 \times \dot{N}_{B,nor} \cdot \text{GI} \quad \text{(B1)}$$

Capital investment ($/year)

$$\text{The CI of the reactor is : } \text{CI}_R = V_R \cdot \psi_R/\text{PBT} \quad \text{(B2)}$$

The CI of the distillation column is composed of the cost for trays and the cost for the column vessel

$$\text{CI}_{column} = ((N_T/\xi_T) \cdot \psi_T + ((N_T/\xi_T - 1) \times 0.6 + 6) \cdot \psi_V)/\text{PBT} \quad \text{(B3)}$$

The heat transfer areas of the heat exchanger, condenser, and reboiler are given by

$$S_{heater} = Q_H \Big/ \left(\frac{T_F - T_0}{\ln((423 - T_0)/(423 - T_F))} \times 1420\right) \quad \text{(B4)}$$

$$S_{reb} = Q_{reb}/(1420 \times (493 - T_{bub})) \quad \text{(B5)}$$

$$S_{cond} = Q_{cond}/(567.8 \times 20/\ln((T_{dew} - 300)/(T_{dew} - 320))) \quad \text{(B6)}$$

Condensing steams of 423 K and 493 K were used as media in the heat exchanger and reboiler, respectively. The temperature of the cooling water in the condenser is assumed to increase from 300 to 320 K. The heat-transfer coefficients in the condensing steam heating and the room temperature water cooling processes are estimated as 1420 and 567.8 $W/(m^2K)$, respectively.

With the calculated heat equipment sizes, their CIs can be evaluated according to the bare module cost model,[62] where u = heater, cond, reb denote heat exchanger, condenser, and reboiler, respectively

$$\text{BC}_u = C_0(S_u/S_0)^{\alpha} \quad \text{(B7)}$$

$$\text{CI}_u = \text{BC}_u \times \text{UF} \times (\text{MF} + \text{MPF} - 1)/\text{PBT} \quad \text{(B8)}$$

Finally, the total CI can be calculated as

$$\text{CI}_{tot} = \text{CI}_R + \text{CI}_{column} + \text{CI}_{heater} + \text{CI}_{cond} + \text{CI}_{reb} \quad \text{(B9)}$$

Utility cost ($/year)

Assuming that the temperature increase of the cooling water in the reactor and in the condenser are 5 and 20 K, respectively, UCs of the reactor and the condenser can be evaluated as follows

$$\text{UC}_R = 330 \times 24 \times 60 \times 60 \times Q_R \cdot \psi_{CW}/(75.33 \times 5) \quad \text{(B10)}$$

$$\text{UC}_{cond} = 330 \times 24 \times 60 \times 60 \times Q_{cond} \cdot \psi_{CW}/(75.33 \times 20) \quad \text{(B11)}$$

The costs of hot steams in the heat exchanger and reboiler are calculated as follows

$$\text{UC}_{heater} = 330 \times 24 \times 60 \times 60 \times Q_H \cdot \psi_{HS}/38012.4 \quad \text{(B12)}$$

$$\text{UC}_{reb} = 330 \times 24 \times 60 \times 60 \times Q_{reb} \cdot \psi_{HS}/33303.6 \quad \text{(B13)}$$

The heat capacity of room temperature cooling water is 75.33 $J/(mol \cdot K)$. The enthalpies of vaporization of 423 and 493 K hot steams are 38012.4 and 33303.6 J/mol, respectively.

The total UC each year is

$$\text{UC}_{tot} = \text{UC}_R + \text{UC}_{heater} + \text{UC}_{cond} + \text{UC}_{reb} \quad \text{(B14)}$$

Appendix C: Supporting Data

Appendix C is found in the Supporting Information.

Manuscript received May 10, 2014, and revision received July 24, 2014.