


Rx-COSMO-CAMPD: Enhancing Reactions by Integrated Computer-Aided Design of Solvents and Processes based on Quantum Chemistry

Christoph Gertig^{1,*}, Lorenz Fleitmann¹, Johannes Schilling¹, Kai Leonhard¹,
and André Bardow^{1,2,3,*}

DOI: 10.1002/cite.202000112

 This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.



Supporting Information
available online

Solvents strongly affect reaction-based chemical processes. Process design, therefore, needs to integrate solvent design. For this purpose, the integrated computer-aided molecular and process design (CAMPD) method Rx-COSMO-CAMPD is proposed. It employs a hybrid optimization scheme combining a genetic algorithm to explore the molecular design space with gradient-based optimization of the process. To overcome limitations of molecular design based on group-contribution methods, reaction kinetics and thermodynamic properties are predicted using advanced quantum-chemical methods. Rx-COSMO-CAMPD is demonstrated in a case study of a carbamate-cleavage process where promising solvents are designed efficiently. The results show that the integrated solvent and process design with Rx-COSMO-CAMPD outperforms computer-aided molecular design without process optimization in the identification of solvents that enable optimal process performance.

Keywords: Computational chemistry, Computer-aided molecular design, Kinetics, Process design, Solvent selection

Received: May 31, 2020; *revised:* August 20, 2020; *accepted:* August 21, 2020

1 Introduction

The performance of chemical and energy processes is often strongly affected by molecules used in the process [1], e.g., by solvents for liquid-liquid extraction [2], washing agents for absorption [3] or working fluids for organic Rankine cycles [4, 5]. In particular, solvent molecules can have a major impact on chemical reactions [6]. Therefore, the identification of optimal solvents is an important task in the design of reaction-based chemical processes. However, the number of potential solvent molecules for a process is vast [7] and systematic design methods are required to identify the most promising molecules. To design optimal solvents *in silico*, computer-aided molecular design (CAMD) [8–10] methods have been developed. Such design methods systematically explore the chemical space and have to fulfill two key requirements to be reliable [11]: First, sound prediction methods are required to capture the impact of designed solvents on the reaction systems, especially on kinetics. Second, the evaluation criteria for the candidate solvents should be based on their performance in the chemical process. In the following, important developments in the field of CAMD methods for solvent design are discussed. The focus is on published literature on CAMD of reaction solvents considering the solvent impact on reaction kinetics.

This provides the background for the presentation of our own developments in subsequent sections of this paper.

For the evaluation criteria, early CAMD methods for the design of reaction solvents [12–14] relied on expert knowledge. Substantial previous knowledge was required to derive criteria to assess the candidate solvents. Commonly, the expert knowledge-based criteria defined target ranges for several properties (e.g., solubilities, melting and boiling points) leading to good candidates but do not allow for a quantitative ranking of the solvent candidates based on a

¹Christoph Gertig, Lorenz Fleitmann, Johannes Schilling, Prof. Dr. rer. nat. Kai Leonhard, Prof. Dr.-Ing. André Bardow
Christoph.Gertig@ltt.rwth-aachen.de
RWTH Aachen University, Institute of Technical Thermodynamics, Schinkelstraße 8, 52062 Aachen, Germany.

²Prof. Dr.-Ing. André Bardow
Forschungszentrum Jülich GmbH, Institute of Energy and Climate Research – Energy Systems Engineering (IEK-10), Wilhelm-Johnen-Straße, 52425 Jülich, Germany.

³Prof. Dr.-Ing. André Bardow
abardow@ethz.ch
ETH Zurich, Department of Mechanical and Process Engineering, Energy & Process Systems Engineering, Tannenstrasse 3, 8092 Zürich, Switzerland.

common metric. More importantly, the employed criteria do not capture process performance directly.

A popular evaluation criterion in CAMD methods for reaction solvents is the reaction rate constant or rate. Employing the reaction rate constant as criterion is generally applicable to any reaction-based process without prior knowledge about the specific process. To predict reaction rate constants in designed solvents, Folić et al. [15,16] fit the coefficients of the solvatochromic equation [17,18] to experimentally determined rate constants in a few solvents. Subsequently, the fitted equation is used to assess candidate solvents designed based on a predefined set of UNIFAC groups [19]. Required descriptors of the solvatochromic equation are obtained from group-contribution (GC) methods. Similar to Folić et al., also Struebing et al. [20,21] predict reaction rate constants during design based on GC methods and the solvatochromic equation. However, they fit the coefficients of the solvatochromic equation to predicted rate constants in a small set of solvents. These rate constants are predicted by transition state theory (TST) [22] and quantum-mechanical density-functional theory (DFT). Struebing et al. also perform the TST-based prediction for the top solvent of every design step and use the result to improve the coefficients in the solvatochromic equation. Thus, the prediction accuracy is expected to improve as the molecular design progresses.

Liu et al. [23] propose a CAMD method for the design of reaction solvents where the equation used for the prediction of rate constants is based on TST and the use of additional descriptors. Fitting of coefficients in this equation to experimental data is performed specifically for each reaction under consideration. The descriptors are obtained from the solvation model COSMO-SAC [24] and GC methods. Solvent molecules are designed based on the functional groups available in the used GC methods. In an extension to their approach, Liu et al. [25] take the inertness of designed solvents into account by adding constraints on the equilibrium constants of possible side reactions.

The previously discussed methods rely on equations with fitted coefficients to predict the rate constants such as the solvatochromic equation. The CAMD approach of Austin et al. [26,27] overcomes this reliance. In this approach, reaction rate constants are directly predicted with TST for every candidate solvent considered during design. DFT and the solvation model COSMO-RS [28] are used to predict the activation barriers required to calculate rate constants with TST. For quantitative predictions, one parameter is fitted to experimental data for each reaction considered. The reaction solvents are designed based on a predefined set of UNIFAC groups. Recently, the present authors [29] developed a method for CAMD of reaction solvents with a complete prediction of reaction rate constants based on TST. The prediction also splits the activation barrier required to calculate rate constants with TST into solvent-independent and solvent-dependent contributions. The term “solvent-independent contributions” refers to all contributions to the

activation barrier obtained by quantum chemical or thermochemical calculations independent of the solvent and the reaction mixture composition, e.g., contributions of electronic energies. “Solvent-dependent contributions” are the contributions that arise due to interactions of a molecule with its environment in a liquid reaction mixture. The solvent-dependent contributions are calculated by COSMO-RS as in other methods. The solvent-independent contributions are now determined by accurate quantum chemical post-Hartree-Fock (HF) methods. Despite the use of highly accurate prediction methods, this approach has been shown to be computationally feasible. The CAMD method employs a genetic optimization algorithm that designs solvent molecules based on 3D molecule fragments to explore the molecular design space [30]. The 3D molecular geometries provide direct input for the quantum-chemical prediction methods.

In summary, several CAMD methods for reaction solvents explore the molecular design space and evaluate candidate solvent molecules quantitatively based on their impact on the reaction rate constant or rate. Some approaches can even design reaction solvents without the need for experimental data to fit kinetic models [20,21,29]. However, as explained above, assessing solvent performance based on reaction rates only is not reliable and solvents should rather be assessed based on their performance in chemical processes to ensure reliable results. Optimal process performance is usually based on delicate trade-offs between different requirements [1, 31–33]. These trade-offs cannot be captured if solvents are solely evaluated based on rate constants [11]. CAMD methods have, therefore, been extended to integrated computer-aided molecular and process design (CAMPD) [34].

There are only few approaches for CAMPD of reaction solvents in the literature: Sioumkrou et al. [35] design gas-expanded liquids (GXL) as solvents for a Diels-Alder reaction and assess candidate solvent molecules based on process-related objectives such as the total annualized cost. Optimal GXLs are designed based on a small set of organic solvents and CO₂ as expansion gas. Reaction rate constants are predicted using the solvatochromic equation with coefficients regressed to experimental data. Also, Zhou et al. [36] demonstrate CAMPD of reaction solvents for Diels-Alder reactions. Candidate solvents are evaluated according to costs determined using process and cost models. Solvent molecules are designed based on a set of UNIFAC groups. Reaction rate constants are predicted using a quantitative structure-property relationship (QSPR) model with descriptors based on so-called σ -profiles [37]. These profiles are histograms of screening charge densities that characterize a molecule's electrostatic interaction with an infinitely polarizable environment. The coefficients of the QSPR model are fitted to experimentally determined rate constants in a few solvents. The σ -profiles of candidate solvents are predicted using a GC method. In another work, Zhou et al. [38] show how to account for prediction uncertainties in a robust design. The discussed CAMPD methods for integrated

design of reaction solvents and processes allow for the desired process-based assessment of candidate solvent molecules. Thus, trade-offs in the process are accounted for and the reliability of the design results is increased. However, the prediction of reaction kinetics relies on equations fitted to experimental data that needs to be available. Moreover, the design depends on the quality of the used QSPR [27].

In this work, a method for integrated design of reaction solvents and processes called Rx-COSMO-CAMPD is proposed that enables solvent design using prediction of reaction kinetics based on quantum chemical methods. No fitting of parameters to experimental data is required. Every considered candidate solvent molecule is evaluated based on process performance. A hybrid stochastic/deterministic optimization yields the optimal combination of reaction solvent and process conditions. Thus, the present work extends our COSMO-CAMPD framework [39,40] to reactive processes. Solvent molecules are designed using the genetic optimization algorithm LEA3D [30] that optimizes molecular structures based on a library of 3D molecule fragments. The availability of 3D structural information facilitates the use of quantum chemical prediction methods. Reaction rate constants are calculated using TST following previous work [29]. The activation barriers required by TST are split into solvent-independent contributions determined using high-level quantum chemical methods and solvent-dependent contributions determined based on COSMO-RS. Deterministic gradient-based optimization is used to optimize process conditions. The design with Rx-COSMO-CAMPD is demonstrated for the case study of a carbamate-cleavage process. This case study is part of a possible CO₂-based production route to industrially important isocyanates investigated in the research project Carbon2Chem [41].

This paper is structured as follows: in Sect. 2, the integrated reaction solvent and process design problem is formulated. Next, the methods used for the prediction of reaction kinetics and thermodynamic equilibrium properties as well as process modeling are explained. Subsequently, the solution approach of Rx-COSMO-CAMPD is detailed. In Sect. 3, Rx-COSMO-CAMPD is applied for the integrated solvent and process design for thermal carbamate cleavage. Conclusions are drawn in Sect. 4.

2 Rx-COSMO-CAMPD Method for Integrated Reaction Solvent and Process Design

The integrated reaction solvent and process design problem is formulated as an optimization problem based on the general formulation given by Gani [42]:

$$\begin{aligned}
 &\max_{x,y} f(x, \Theta, k) && \text{ } \} \text{process-based objective} \\
 &s.t. \quad k = h_1(x, \Theta) && \text{ } \} \text{kinetic model} \\
 &\quad \Theta = h_2(x, y) && \text{ } \} \text{thermodynamic property model} \\
 &\quad 0 = h_3(x, \Theta) && \text{ } \} \text{process model} \\
 &\quad g_1(y) = 0 && \text{ } \} \text{chemical feasibility of molecules} \\
 &\quad g_2(y) \leq 0 && \text{ } \} \text{chemical feasibility of molecules} \\
 &\quad c_1(\Theta) \leq 0 && \text{ } \} \text{constraints on thermodynamic properties} \\
 &\quad c_2(y) \leq 0 && \text{ } \} \text{constraints on molecular properties} \\
 &\quad c_3(x, \Theta, k) \leq 0 && \text{ } \} \text{constraints on the process} \\
 &\quad x \in X && \text{ } \} \text{process conditions} \\
 &\quad y \in Y && \text{ } \} \text{molecular structure of the solvent}
 \end{aligned} \tag{1}$$

In problem (1), $f(x, \Theta, k)$ represents a process-based objective function (e.g., yield or profit) that depends on the process conditions x , thermodynamic equilibrium properties Θ and reaction rate constants k . The process conditions x and the molecular structure y of the solvent are optimized to maximize the objective $f(x, \Theta, k)$. The process conditions x are contained in an allowed operating range of the process X and the molecular structures y of designed solvents in the molecular design space Y . Constraints $h_1(x, \Theta)$ include all equations used to calculate reaction rate constants k and constraints $h_2(x, y)$ all equations used to obtain thermodynamic equilibrium properties Θ . The equations of the process model are denoted by $h_3(x, \Theta)$. To ensure chemical feasibility of designed molecular structures y (e.g., correct valency), equality constraints $g_1(y)$ and inequality constraints $g_2(y)$ are used. Constraints $c_1(\Theta)$ and $c_2(y)$ are used on thermodynamic equilibrium properties (e.g., boiling points) and on molecular properties of the designed solvent molecules (e.g., the maximum number of atoms in designed molecules), respectively. Furthermore, constraints $c_3(x, \Theta, k)$ on the process are used that may depend on the process conditions x , thermodynamic equilibrium properties Θ and reaction rate constants k .

In the following, the methods used for the prediction of reaction rate constants k (Sect. 2.1) as well as thermodynamic equilibrium properties Θ (Sect. 2.2) are summarized and process modeling is shortly described (Sect. 2.3). Subsequently, the solution approach to problem (1) is described (Sect. 2.4).

2.1 Prediction of Reaction Kinetics Based on Quantum Chemistry

The reaction rate constants k are required to describe reaction kinetics in the designed solvents. We assume that the designed solvents are inert and, thus, do not take part in any reactions. This assumption may not hold in general but is justified by careful definition of the molecular design space and inspection of the obtained molecular structures after the design. The employed prediction method for the reaction rate constants k was described in detail in earlier work [29, 43] and is only summarized in the following. The

calculation of rate constants k of elementary reactions is based on conventional TST and the so-called Eyring equation [22, 44]:

$$k = \frac{k_B T}{h} V_m^{(n-1)} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) \quad (2)$$

In Eq. (2), T is the temperature, V_m the molar volume of the reaction phase and n the reaction order defined as the sum of the stoichiometric coefficients of the reactants. The activation barrier of the reaction ΔG^\ddagger is defined as the difference in Gibbs free energy between the so-called transition state [45] of the reaction and the state of the reactants. Planck's constant is denoted by h , the Boltzmann constant by k_B and the gas constant by R . Quantities such as the activation barrier ΔG^\ddagger can be divided into solvent-independent and solvent-dependent contributions to enable the use of appropriate methods for each contribution [46–49]. As derived in a earlier work [29] and shown in the Supporting Information (SI), Eq. (2) then takes the form

$$\begin{aligned} k &= \frac{k_B T}{h} (V_m^{i.G.})^{(n-1)} \exp\left(-\frac{\Delta G_{\ddagger,i.G.}^\ddagger}{RT}\right) \frac{\prod_i \gamma_i^{uN}}{\gamma_{\ddagger}^{uN}} \\ &\quad \cdot \exp\left(-\frac{\Delta \tilde{G}_{\ddagger}^{\text{solv}} - \sum_i \Delta \tilde{G}_i^{\text{solv}}}{RT}\right) \quad (3) \\ &= k^{i.G.} \frac{\prod_i \gamma_i^{uN}}{\gamma_{\ddagger}^{uN}} \exp\left(-\frac{\Delta \tilde{G}_{\ddagger}^{\text{solv}} - \sum_i \Delta \tilde{G}_i^{\text{solv}}}{RT}\right) \end{aligned}$$

In Eq. (3), $V_m^{i.G.}$ is the molar volume determined by the ideal gas law, $\Delta G_{\ddagger,i.G.}^\ddagger$ the activation barrier of the reaction and $k^{i.G.}$ the reaction rate constant in the ideal gas reference state, respectively. The unsymmetrically normalized activity coefficients of the transition state and the reactants are denoted by γ_{\ddagger}^{uN} and γ_i^{uN} , respectively. The solvation free energies of the transition state and the reactants are denoted by $\Delta \tilde{G}_{\ddagger}^{\text{solv}}$ and $\Delta \tilde{G}_i^{\text{solv}}$, respectively. The solvation free energies are computed using molar reference states for concentration terms and activity at infinite dilution, while the molar volume $V_m^{i.G.}$ and the rate constant in the ideal gas reference state $k^{i.G.}$ are computed at a reference pressure of 1 bar (see SI for details).

To compute the rate constant k using Eq. (3), the following procedure is applied:

- 1) The geometries of reactants and transition states are optimized employing the quantum-mechanical density-functional theory (DFT) method B3LYP [50] with empirical dispersion correction [51] (B3LYP-D3). Subsequently, vibrational analysis is performed. B3LYP is known to perform well in geometry optimizations and vibrational analyses in spite of the moderate computational cost [52, 53]. Possible conformers of reactants and transition states are sought using rotor scans. The rigid rotor harmonic oscillator [54] (RRHO) approxi-

mation is applied in the frequency analysis. The software Gaussian 09 [55] is used for geometry optimization and frequency analysis.

- 2) To guarantee that the identified transition states connect the correct reactants and products, intrinsic reaction coordinate [45] (IRC) scans are performed.
- 3) Accurate electronic energies of reactants and transition states are determined in single point (SP) calculations. The post-Hartree-Fock method DLPNO-CCSD(T) [56, 57] is used with auc-cc-pVTZ basis set and TightPNO settings implemented in ORCA [58] for SP calculations.
- 4) To obtain activation barriers in the ideal gas reference state $\Delta G_{\ddagger,i.G.}^\ddagger$, thermochemical calculations are performed with GoodVibes [59]. Grimme's quasi-harmonic treatment [60] is employed to reduce the error of the RRHO approximation for low frequencies.
- 5) Reaction rate constants in the ideal gas reference state $k^{i.G.}$ are calculated.
- 6) Tunneling corrections to $k^{i.G.}$ are computed based on Eckart tunneling [61] using the software package TAMkin [62].
- 7) Solvation free energies $\Delta \tilde{G}^{\text{solv}}$ are calculated for reactants and transition states using the solvation model COSMO-RS [28, 37, 63]. The software turbomole [64, 65] is used to optimize reactant and transition state geometries with the DFT method BP86 [66–68] and def2-TZVP basis set. Using the obtained geometries, COSMO [69] calculations are performed as single point calculations. Subsequently, the actual COSMO-RS calculations [70, 71] are performed to obtain the solvation free energies $\Delta \tilde{G}^{\text{solv}}$.
- 8) The unsymmetrically normalized activity coefficients γ^{uN} are calculated using either COSMO-RS directly or a suitable surrogate model fitted to activity coefficient data from COSMO-RS.
- 9) The reaction rate constant k is calculated using Eq. (3).

Conformers of reactants and transition states are accounted for as described in our previous work [72]. The expected uncertainty in the rate constants k predicted with the procedure described above was also discussed in detail elsewhere [29, 43, 72] and only the results are given here: Experimental and predicted rate constants k should agree within one order of magnitude at temperatures around 25 °C and the accuracy is expected to increase at higher temperatures. Moreover, some errors cancel if rate constants in different solvents are compared [29], which is beneficial in solvent design where rankings are more important than absolute values.

The procedure described above is also used to compute equilibrium constants K of reactions. The only differences are that no tunneling correction is required and that the calculations performed for the transition states are performed for the products instead.

It is important to note that steps 1–3 are the most time-consuming steps in the procedure described above. As these

steps do not depend on the used solvent, they can be performed in advance of the actual solvent design. Consequently, using the described procedure in solvent design is computationally feasible despite employing high-level quantum chemical methods like DLPNO-CCSD(T). Furthermore, only the unsymmetrically normalized activity coefficients γ_i^{UN} depend on the reaction mixture composition and must be evaluated during process simulation, while the solvation free energies $\Delta\tilde{G}_i^{\text{solv}}$ can be computed before process simulations are started.

2.2 Prediction of Thermodynamic Equilibrium Properties

For the prediction of thermodynamic equilibrium properties, the vapor phases are assumed to show ideal behavior. COSMO-RS is used to compute solvation free energies $\Delta\tilde{G}_i^{\text{solv}}$ and enthalpies $\Delta\tilde{H}_i^{\text{solv}}$ as well as to obtain parameters for the Antoine equation [19] used to calculate vapor pressures p_i^{S} . Henry coefficients \tilde{H} are obtained from COSMO-RS as well to describe the vapor-liquid equilibrium (VLE) behavior of supercritical components.

To compute activity coefficients γ_i efficiently during process simulation and optimization, the NRTL [73] model is used as surrogate model. NRTL parameters are fitted to activity coefficient data from COSMO-RS.

Molar isobaric ideal gas heat capacities $c_{p,i}^{\text{p.i.G.}}$ are computed with TAMkin [62] based on RRHO and geometry optimization and vibrational analysis with Gaussian 09 [55] using AM1 [74]. Molar heat capacities of liquids c_p^{L} are determined from molar isobaric ideal gas heat capacities $c_{p,i}^{\text{p.i.G.}}$ and 2nd derivatives of solvation free energies $\Delta\tilde{G}_i^{\text{solv}}$ (see SI). Enthalpies of gaseous and liquid mixtures are determined based on heat capacities $c_{p,i}^{\text{p.i.G.}}$ and c_p^{L} , respectively, activity coefficients γ_i and solvation enthalpies $\Delta\tilde{H}_i^{\text{solv}}$ as described in the SI.

2.3 Process Modeling

Process models are formulated as differential-algebraic systems of equations (DAE) based on mole and energy balances. Reaction kinetics are described by power laws [75] including the predicted reaction rate constants k . Moreover, equations are included that describe the VLE behavior of mixtures (see SI). Process models are implemented in MATLAB [76]. The solver ode15i for systems of ordinary differential equations (ODE) is used to solve the process models.

2.4 Solution Approach of Rx-COSMO-CAMPD

The proposed Rx-COSMO-CAMPD method for integrated reaction solvent and process design is based on our

Rx-COSMO-CAMD method [29] for design of reaction solvents and our COSMO-CAMPD [39,40] framework for non-reactive systems. A preliminary version of Rx-COSMO-CAMPD was already described in a conference paper [77].

Rx-COSMO-CAMPD follows a hybrid optimization scheme where the molecular structure y is optimized using the genetic algorithm LEA3D [30] and gradient-based optimization is used to optimize the process conditions x . LEA3D employs libraries with 3D molecule fragments that are combined to molecular structures. The optimal structure is determined by genetic optimization. The genetic optimization is based on a sequence of generations of candidate molecules, where the candidates of a new generation are suggested based on the performance of the candidates of the previous generation. As LEA3D works with 3D structures, its output can directly be used as input for the quantum chemical methods for property prediction. In particular, the 3D information ensures that isomers are distinguished and specific interactions of functional groups depending on the 3D geometry are accounted for. Moreover, the explored molecular design space is not limited to certain sets of functional groups fitted previously to experimental data.

The whole procedure of Rx-COSMO-CAMPD used to solve problem (1) is shown in the flowchart in Fig. 1 and described in the following:

- 1) The specifications for quantum chemical calculations (geometry optimizations, frequency analyses, SP calculations) are defined as described in Sect. 2.1. Furthermore, the reaction network is specified.
- 2) Subsequently, the rate constants in the ideal gas reference state $k^{\text{i.G.}}$ are calculated. As already noted in Sect. 2.1, all contributions to $k^{\text{i.G.}}$ are independent of the used solvent and the computation of $k^{\text{i.G.}}$ is not repeated during rest of the design procedure.
- 3) Prior to the start of the hybrid optimization, several specifications are required:
 - specification of the fragment library containing the 3D fragments for molecular design; here: specific molecular fragments could be excluded that would inevitably lead to the design of solvents that undergo side reactions,
 - settings for the genetic optimization of molecular structures (e.g., maximum number of generations, number of candidate molecules per generation, probabilities for genetic operations),
 - specification of constraints $c_1(\Theta)$ on thermodynamic equilibrium properties, $c_2(y)$ on molecular properties of the designed solvent molecules and constraints $c_3(x,\Theta,k)$ on the process,
 - specification of the process model $h_3(x,\Theta)$,
 - choice of the process-based objective function $f(x,\Theta,k)$.
- 4) Based on the specified fragment library, LEA3D generates 3D molecular structures of candidate solvents. For

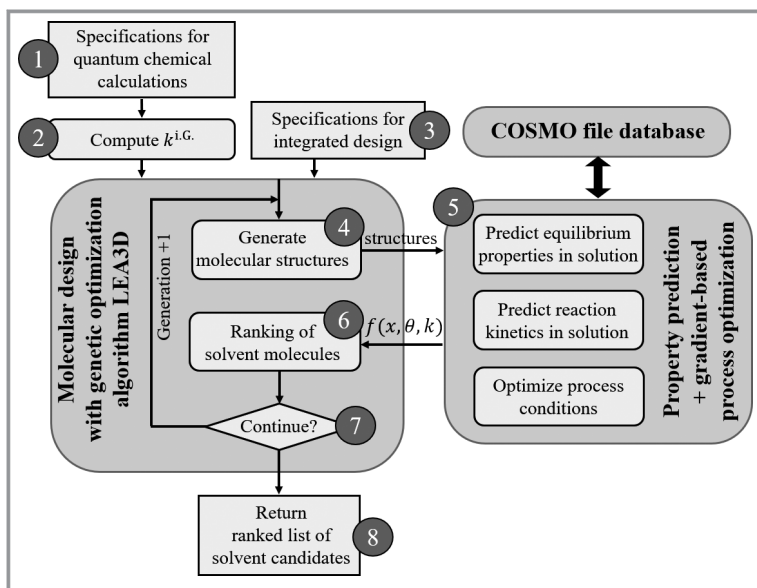


Figure 1. Flowchart of Rx-COSMO-CAMPD for integrated reaction solvent and process design.

the initial generation, the fragments are combined randomly. Candidates of subsequent generations are obtained by genetic operations like cross-over or mutation of promising candidates from the previous generation. The constraints $g_1(y)$ and $g_2(y)$ are respected by LEA3D in these operations to ensure chemical feasibility of the designed candidate solvent molecules.

- 5) Next, all necessary computations for the determination of the process-based objective $f(x, \Theta, k)$ are performed. For this purpose, reaction rate constants k as well as thermodynamic equilibrium properties Θ are calculated as described in Sect. 2.1 and 2.2. The software COSMO-conf [78] is used to perform the COSMO calculations required as basis for COSMO-RS. The COSMO calculations are the most time-consuming and computationally expensive part of evaluating the process objective $f(x, \Theta, k)$ during design. Therefore, the results of COSMO calculations performed for newly designed molecules are stored in a COSMO-file database. In case LEA3D suggests candidate molecules that are already available in the database, COSMO calculations are skipped. Gradient-based optimizations of the process conditions x are performed for each candidate solvent. The interior-point algorithm available in the MATLAB [76] function `fmincon` for constrained optimization is used for these optimizations. The solver `ode15i` for systems of ordinary differential equations (ODE) is used to solve the process models.
- 6) The values of the objective $f(x, \Theta, k)$ determined for the candidate solvent molecules of a generation are the basis for the current ranking of the candidates.

- 7) Based on promising candidate solvent molecules from the current generation, LEA3D generates a new generation of candidates using genetic operations (step 4). Steps 4 to 7 are repeated until the maximum number of generations specified in step 3 is reached in step 7.
- 8) The result of the integrated reaction solvent and process design with Rx-COSMO-CAMPD is a list of combinations of solvent molecules and corresponding optimal process conditions ranked according to the values of the objective $f(x, \Theta, k)$. From this list, the user may choose an appropriate solvent. This choice might include further criteria like commercial availability and environmental properties that have not been part of the optimization procedure. Moreover, any available expert knowledge on the reaction system that is not explicitly used in the design may be useful to make the best possible final choice.

3 Case Study: Integrated Design of Solvent and Carbamate-Cleavage Process

The integrated solvent and process design with Rx-COSMO-CAMPD is demonstrated in the case study of the thermal cleavage reaction of methyl phenyl carbamate (MPC) to phenyl isocyanate and methanol (MeOH) (Figure 2).

Carbamate-cleavage reactions are part of production routes to industrially important isocyanates [79]. In particular, these cleavage reactions are part of a CO₂-based route [80] investigated in the research project Carbon2Chem [41]. The process design is challenging because carbamate-cleavage reactions are strongly endothermic and have a very unfavorable reaction equilibrium that lies almost completely on the reactant's side [81]. Typical reaction temperatures are above 200 °C.

To avoid the fast back-reaction, methanol must be continuously removed from the liquid reaction phase. For this purpose, the high volatility of methanol compared to the carbamate and isocyanate can be utilized. One option for methanol removal from the reaction phase is stripping with the inert gas nitrogen [82]. Fig. 3 shows a possible process

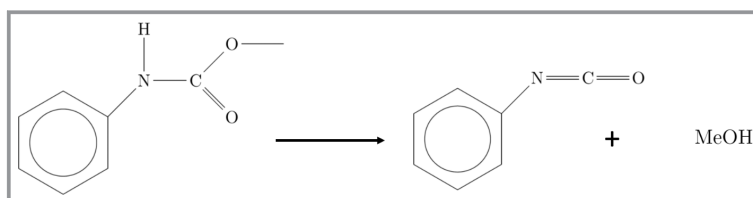


Figure 2. Thermal cleavage reaction of methyl phenyl carbamate (MPC) to phenyl isocyanate and methanol (MeOH).

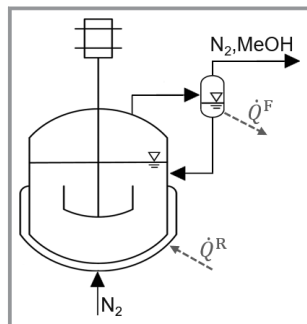


Figure 3. Flowsheet of the semi-batch process for carbamate cleavage. Material streams are depicted with black solid and heat flows with gray dashed arrows.

flowsheet of a batch process that was already introduced in a similar case study with a preliminary version of Rx-COSMO-CAMPD [77]. The reaction is carried out in the liquid phase in a semi-batch reactor. The reaction temperature T^R is kept constant by a heat flow \dot{Q}^R provided by pressurized steam. The formed methanol is stripped out of the reactor by a nitrogen stream. A flash is used as cooling trap to condense stripped solvent, isocyanate and carbamate that are then recycled to the reactor. A heat flow \dot{Q}^F is removed from the flash for cooling and condensation. Cooling water is used to take up the heat flow \dot{Q}^F . Nitrogen and methanol leave the process in the gaseous outlet stream of the flash.

The process model includes mole and energy balance equations, kinetic equations and equations accounting for vapor-liquid equilibrium (VLE) as discussed in Sect. 2.3. The derivation of the full set of equations $h_3(x, \Theta)$ as well as the assumptions made are given in the SI. Reaction rate constants and thermodynamic equilibrium properties are computed as discussed in Sect. 2.1 and 2.2. It was shown in a previous work [72] that the scheme and methods presented in Sect. 2.1 are suited to predict the kinetics of MPC cleavage quantitatively.

Due to the challenging properties of the reaction system, it is difficult to produce appreciable product amounts in a carbamate-cleavage batch process. Thus, a first integrated reaction solvent and carbamate-cleavage process design is performed aiming to maximize the isocyanate yield.

3.1 Design Specifications for Yield Optimization

The yield is defined as the final moles of isocyanate in the reactor divided by the initial moles of carbamate and chosen as objective function $f(x, \Theta, k)$ in problem (1) for the first case study. The objective function is maximized by identifying optimal process conditions x^* and the optimal solvent molecular structure y^* . The process conditions x that are degrees of freedom in the optimization are: the temperature in the flash T^F , the volume flow of nitrogen \dot{V}_{N_2} and the pressure p . This choice of degrees of freedom allows the investigation of various influences and trade-offs in the integrated solvent and carbamate-cleavage process design. However, different choices are also possible allowing for further process optimization. The reaction temperature T^R

in the reactor is fixed to 220 °C, which is in the typical temperature range used for carbamate cleavage in the liquid phase [72, 83]. Higher reaction temperatures could enhance the cleavage reaction but also lead to undesired side reactions. The batch process is run for a reaction time of 12 h. A total volume of the reactor of $V^R = 1 \text{ m}^3$ is chosen.

The constraints $g_1(y)$ and $g_2(y)$ are respected by LEA3D in the design of molecular structures (Sect. 2.4). Constraints $c_1(\Theta)$ are used to limit the normal boiling points of designed solvents to values between 171 °C and 600 °C to avoid the design of solvents that are more volatile than the produced isocyanate or that have extreme boiling points. Furthermore, a minimal solubility of 20 wt % carbamate in the solvent is included in the constraints $c_1(\Theta)$. Constraints $c_2(y)$ on molecular properties are used to limit the number of non-hydrogen atoms in designed molecules to a maximum of 16 and to avoid the design of protic solvents that would possibly undergo side reactions with the produced isocyanate.

The operation range X of the process is set to allow flash temperatures of $10^\circ\text{C} < T^F < 60^\circ\text{C}$, nitrogen volume flows of $0 < \dot{V}_{N_2} < 0.1 \text{ m}^3\text{s}^{-1}$ and pressures of $1 \text{ bar} < p < 10 \text{ bar}$. The lower bound on T^F is set based on the assumption that cooling water with a temperature T^C of 3 °C is available and a minimal temperature difference $T^F - T^C$ of 2 °C is required. The lower bound on \dot{V}_{N_2} is used as no negative flows are possible. The lower bound on p of 1 bar approximately corresponds to the ambient pressure. Thus, no operation under partial vacuum is allowed. The upper bounds are set to define a search space of reasonable size. It is expected that favorable process conditions lie well below these upper bounds.

According to the literature [84–86], various aprotic solvents are suitable for carbamate cleavage. To define a sufficiently rich molecular design space Y , a library of alkane, aromatic, ether, ester, amide, keto, aldehyde, nitro, sulfone and sulfoxide fragments is provided. The full list of fragments is given in the SI. For the genetic optimization with LEA3D, a maximum number of 35 generations and 40 candidate molecules per generation are specified.

3.2 Results of the Design for Yield Optimization

The integrated reaction solvent and process design as specified in Sect. 3.1 identifies 2-phenyl butane as optimal solvent y^* . The corresponding optimal process conditions x^* are determined as the flash temperature $T^F = 10.6^\circ\text{C}$, the nitrogen volume flow $\dot{V}_{N_2} = 0.050 \text{ m}^3\text{s}^{-1}$ and the pressure $p = 1.65 \text{ bar}$. The optimal value of the objective function $f(x, \Theta, k)$ is a yield of 38.5 %. The computations took about 37 h on the compute cluster of RWTH Aachen University using 48 processor cores of type Intel Skylake Platinum 8160. In total, 329 designed solvents meet all constraints. Fig. 4 presents the yields achieved by optimized processes with the designed solvents plotted versus the rate constants k and equilibrium constants K .

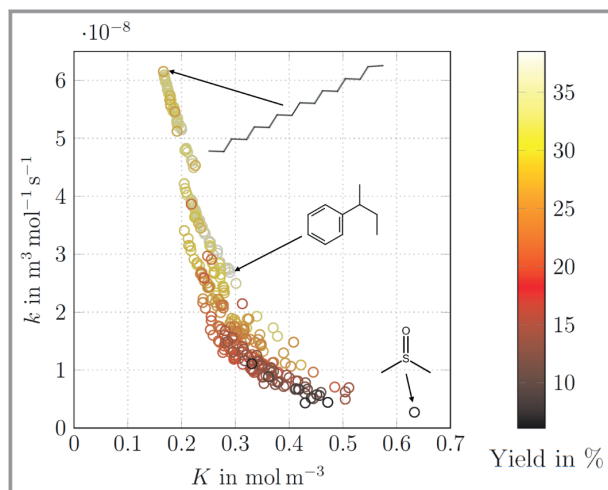


Figure 4. Optimized yields of carbamate cleavage processes with 329 designed solvents plotted versus the rate constants k and equilibrium constants K of the cleavage reaction in the solvents. The optimal solvent 2-phenyl butane is shown as well as the solvents with the highest reaction rate (n -hexadecane) and the highest equilibrium constant (DMSO).

Fig. 4 shows that the optimal solvent 2-phenyl butane neither features the highest rate constant k nor the highest equilibrium constant K . Trends of increasing yield with both, increasing k and increasing K are visible. However, a trade-off between rate constant k and equilibrium constant K is observed: If a reaction solvent provides a high rate constant, the equilibrium constant tends to be low, and vice versa. This observation can be explained on the molecular level based on the influence of polarity: The transition state of the carbamate-cleavage reaction is less polar than the reactant methyl phenyl carbamate. In contrast, the products (isocyanate and methanol) are more polar than the reactant. Accordingly, increasing polarity of the reaction solvent increases the relative stabilization of the products and decreases the relative stabilization of the transition state, leading to opposite trends for the equilibrium constant K and the reaction rate constant k . A design with the rate constant k as objective f identifies n -hexadecane as optimal solvent, which allows a maximum yield of 32.8 %. A design with the equilibrium constant K as objective f identifies dimethyl sulfoxide (DMSO) as optimal solvent that only allows for a yield of 6.8 % in the corresponding optimal process. The reason why the simpler design criteria such as high rate or equilibrium constants do not identify the solvent that enables the optimal yield is that only the integrated reaction solvent and process design correctly accounts for trade-offs (e.g., between k and K) as well as other criteria like activities and relative volatilities that, e.g., influence the selective methanol removal. This example clearly shows the strength of integrated solvent and process design in the identification of optimal solvents.

To further study the behavior of the carbamate-cleavage process, the achievable yield was computed as function of

the flash temperature T^F and the volume flow of nitrogen \dot{V}_{N_2} for the process with the optimal solvent 2-phenyl butane (Fig. 5).

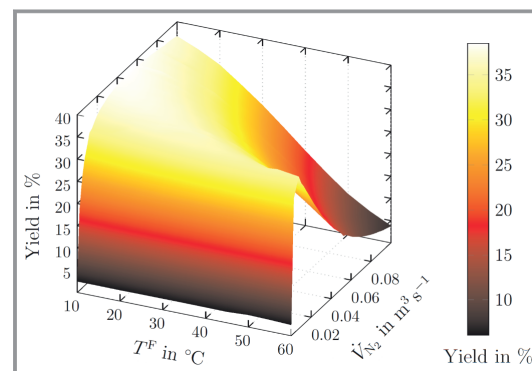


Figure 5. Yield versus the flash temperature T^F and the volume flow of nitrogen \dot{V}_{N_2} at a constant pressure of $p = 1.65$ bar for the process with the optimal solvent 2-phenyl butane.

As can be seen in Fig. 5, only very small yields are achieved without stripping ($\dot{V}_{N_2} = 0$) and the yield first increases strongly with increasing \dot{V}_{N_2} . This strong increase can be attributed to a large benefit of methanol removal from the reactor. Favorable values of the nitrogen volume flow \dot{V}_{N_2} are related to the flash temperature T^F : the higher T^F , the lower are favorable values of \dot{V}_{N_2} . The reason for this behavior is that the dew point of the gaseous reactor outlet stream depends on its nitrogen fraction. While higher volume flows of nitrogen generally remove more methanol from the reactor, the flash temperature must be low enough to condense and recycle the co-stripped isocyanate. This condensation and recycling is only ensured if the volume flow of nitrogen \dot{V}_{N_2} is not too high compared to the flash temperature T^F . Consequently, at $T^F = 10^\circ\text{C}$, the yield increases with increasing \dot{V}_{N_2} up to high values of \dot{V}_{N_2} . In contrast, at $T^F = 60^\circ\text{C}$, the yield strongly increases with increasing \dot{V}_{N_2} up to $\dot{V}_{N_2} \approx 0.0075 \text{ m}^3 \text{ s}^{-1}$ and decreases again at higher nitrogen flows.

Another observation that can be made in Fig. 5 is that the yield as function of T^F and \dot{V}_{N_2} exhibits a wide plateau of near-optimal yield values around the optimum. This finding raises the question of whether process conditions x exist that significantly reduce the cost of operation of the process at almost optimal yield. Moreover, the optimal solvent might change if the cost of operation is considered as second objective besides the yield. To answer these questions, a bi-objective design is performed with Rx-COSMO-CAMPD using both the yield and the cost of operation as objectives. Such a bi-objective design with two independent objectives is best suited to study the trade-off between yield and cost. The specifications of this bi-objective design are given in the next section.

3.3 Design Specifications for Bi-Objective Yield and Cost Optimization

A bi-objective design is based on two objective functions $f_1(x, \Theta, k)$ and $f_2(x, \Theta, k)$: the yield as defined in Sect. 3.1 and the costs of operation per batch, respectively. The total costs of operation TOC include the cost C_{steam} of steam to heat the reactor, the cost C_{cw} of cooling water to cool the flash and the cost C_{N_2} of the used amount of the inert gas nitrogen:

$$TOC = C_{\text{steam}} + C_{\text{cw}} + C_{N_2} \quad (4)$$

For the calculation of the individual contributions to TOC , the reader is referred to the SI.

Again, the temperature in the flash T^F , the volume flow of nitrogen \dot{V}_{N_2} and the pressure p are selected as process conditions x that are optimized. The upper bound on T^F is increased to 220 °C in order to include process conditions x where no cooling is required in the design space. The other settings are identical to those given in Sect. 3.1.

The hybrid optimization in the bi-objective design proceeds as follows: For every candidate solvent molecule suggested by the genetic algorithm LEA3D, a first process optimization maximizes $f_1(x, \Theta, k)$ (yield). Next, the process is optimized with objective function $f_2(x, \Theta, k)$ (cost of operation) using a process constraint $c_3(x, \Theta, k)$ that ensures a minimum yield of 5 %. Subsequently, the Pareto front is scanned employing constraints $c_3(x, \Theta, k)$ in the form of normal constraints as described by Ismail-Yahaya and Messac [87]. Finally, a Pareto ranking [88] is determined to calculate scores of the candidate solvent molecules that are returned to LEA3D. This step is required because LEA3D expects a scalar objective function value. In the Pareto ranking, the score of a candidate solvent f^P is calculated as:

$$f^P = \frac{1}{1 + N^P} \quad (5)$$

where N^P indicates the number of solvents from the current or previous generations that completely dominate the performance of the candidate solvent along the whole Pareto front. Thus, the candidate solvents of a generation are scored based on the currently known Pareto front. When the integrated design is completed, the final Pareto front arises from the set of Pareto optimal points. This procedure is much more time-efficient in the hybrid optimization compared to performing separate integrated designs for each point on the Pareto front.

3.4 Results of the Bi-Objective Design for Yield and Cost Optimization

The bi-objective design with Rx-COSMO-CAMPD specified in Sect. 3.3 took about 80 h on the compute cluster of

RWTH Aachen University using 48 processor cores of type Intel Skylake Platinum 8160. Fig. 6 shows the determined Pareto front.

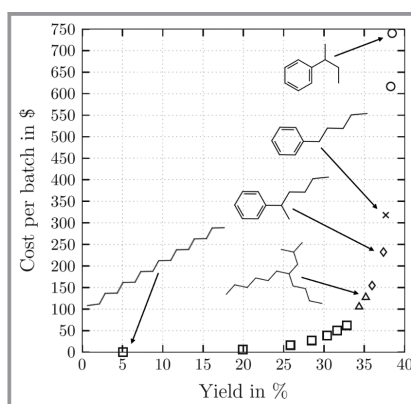


Figure 6. Pareto front determined by the bi-objective design for optimal yield and cost of operation. The optimal solvent changes along the pareto front: while 2-phenyl butane enables the highest yield, *n*-hexadecane enables the lowest cost. Furthermore, an *i*-pentadecane, 2-phenyl hexane and pentyl benzene are found along the Pareto front and enable compromises between optimal yield and optimal cost.

Several conclusions can be drawn from the results of the bi-objective design (Fig. 6): the optimal yield of 38.5 % is achieved at high cost of operation (740 \$ per batch) by the process with 2-phenyl butane as solvent. The main reason for the high achieved yield is the superior removal of methanol from the reactor. Thus, the back-reaction of the formed isocyanate to carbamate is avoided. The optimal cost of operation at 5 % yield (0.21 \$ per batch) is achieved by a process using *n*-hexadecane as solvent. *n*-Hexadecane and similar molecules feature much lower volatilities compared to 2-phenyl butane. Thus, only small amounts of solvent are stripped and recycled. Consequently, the cost of steam and cooling water is reduced. Along the Pareto front, an *i*-pentadecane, 2-phenyl hexane and pentyl benzene are found as solvents that enable compromises between optimal yield and optimal cost.

The cost of operation depends most strongly on the volume flow of nitrogen \dot{V}_{N_2} , which is not only due to the cost of the nitrogen itself: High volume flows of nitrogen carry large amounts of solvent out of the reactor, causing high cost for evaporation in the reactor and for cooling and condensation in the flash. Fig. 7 shows the contributions of different costs to the total cost of operation for two processes: The process with optimal cost at 5 % yield (solvent *n*-hexadecane) and the process with optimal yield (solvent 2-phenyl butane). In both processes, the cost for the steam required to drive the endothermic carbamate-cleavage reaction itself has only a minor share. In contrast, the cost for steam required for heating of streams and for evaporation in the reactor has the largest share. However, the ratio of these costs is much more extreme in case of the process

with optimal yield. Thus, the cost of operation required to reach the optimal yield is disproportionately high. The reason is that the superior methanol removal required to reach very high yields is only possible if large amounts of material are evaporated in the reactor and recycled after partial condensation. This finding demonstrates the benefits of the bi-objective design: candidate solvents are found that enable processes with relatively low costs of operation at still very high yields.

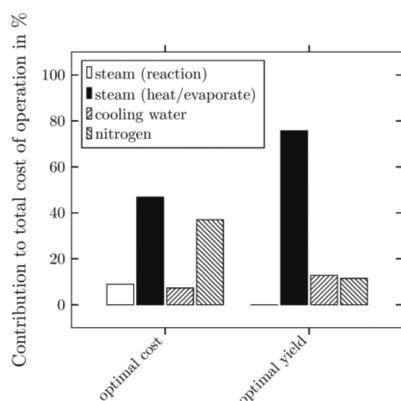


Figure 7. Contribution of different costs (%) to the total cost of operation: cost of steam required for driving the endothermic carbamate cleavage reaction itself; cost of steam required for heating and evaporation; cost of cooling water; cost of nitrogen. Left: process with optimal cost at 5 % yield (solvent: *n*-hexadecane); right: process with optimal yield (solvent: 2-phenyl butane).

In summary, the results demonstrate that the design with Rx-COSMO-CAMPD is able to account for trade-offs between different process-based objectives in a bi-objective design. Trends can be analyzed to learn about different influences on the objectives. Thus, design with Rx-COSMO-CAMPD not only identifies promising reaction solvents, but also facilitates a better understanding of reaction systems and processes.

4 Conclusions

The computer-aided molecular and process design (CAMPD) method Rx-COSMO-CAMPD is proposed for the integrated design of reaction solvents and processes. The method uses a hybrid optimization scheme: the molecular design space is explored by the genetic optimization algorithm LEA3D and process conditions are optimized by gradient-based deterministic optimization. As LEA3D designs molecular structures based on 3D molecule fragments, full 3D structural information is available during design. The availability of 3D structures enables the direct use of quantum chemical prediction methods for the calculation of reaction kinetics and thermodynamic equilibrium properties. Thus, simplified property prediction methods and the need for experimental data are avoided. In the pre-

diction of reaction rate constants, solvent-independent and solvent-dependent contributions are separated. This separation ensures computational feasibility despite using high-level quantum chemical methods in the prediction.

The use of Rx-COSMO-CAMPD for integrated reaction solvent and process design is demonstrated in a case study of a thermal carbamate-cleavage batch process. Promising reaction solvents allowing to maximize product yield are identified efficiently in a single-objective design. The best candidate solvent is 2-phenyl butane and allows to achieve a predicted yield of 38.5 %, which is satisfactory for this highly challenging batch process. In a bi-objective design, trade-offs between yield and cost of operation as objectives are studied and it is shown that the optimal solvent changes along the Pareto front. Promising candidate solvents allowing for compromises between yield and cost are identified.

Despite the use of high-level quantum chemical methods in the prediction, there is no guarantee that no candidate solvent's performance is overestimated. Moreover, the stability and inertness of solvents should always be confirmed before their use in chemical processes. Thus, it is recommended to test the promising candidates identified by Rx-COSMO-CAMPD experimentally.

Currently, Rx-COSMO-CAMPD may be used to design inert solvents for reaction-based chemical processes provided all possible reaction mechanisms of occurring reactions are known. Methods for the automated identification of unknown reaction mechanisms and for the generation of reaction networks are subject of ongoing research and will hopefully become broadly applicable in the near future [89, 90]. Another major challenge is the quantitative prediction of solvation effects for ions that occur in many reaction networks, although the development of appropriate prediction methods is progressing [91].

In summary, the results show that the integrated reaction solvent and process design is superior to computer-aided molecular design (CAMPD) with rate constants or equilibrium constants as objectives. Solvents facilitating a higher process performance are identified by the integrated design. Furthermore, Rx-COSMO-CAMPD allows to study trends and trade-offs, thus, enabling a better understanding of reaction systems and processes.

Supporting Information

Supporting Information for this article can be found under DOI: <https://doi.org/10.1002/cite.202000112>.

The authors thank the German Federal Ministry of Education and Research for funding the project Carbon2Polymers (03EK30442C). Moreover, L.F. gratefully acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – Cluster of Excellence 2186 „The Fuel Science Center” – ID: 390919832. Furthermore, the authors are grateful to Dr. J. Langanke and E. Erdkamp for valuable discussions. Simulations were performed with computing resources granted by RWTH Aachen University under projects rwth0284 and rwth0478.

References

- [1] C. S. Adjiman, A. Galindo, G. Jackson, *Comput.-Aided Chem. Eng.* **2014**, *34*, 55–64.
- [2] C. Redepenning, W. Marquardt, *ACS Sustainable Chem. Eng.* **2017**, *5*, 3382–3389.
- [3] F. E. Pereira, E. Keskes, A. Galindo, G. Jackson, C. S. Adjiman, *Comput. Chem. Eng.* **2011**, *35*, 474–491.
- [4] J. Schilling, D. Tillmanns, M. Lampe, M. Hopp, J. Gross, A. Bardow, *Mol. Syst. Des. Eng.* **2017**, *2*, 301–320.
- [5] A. Papadopoulos, M. Stijepovic, P. Linke, *Appl. Therm. Eng.* **2010**, *30*, 760–769.
- [6] C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, 2nd ed., VCH, Weinheim **1990**.
- [7] T. Fink, H. Bruggesser, J.-L. Reymond, *Angew. Chem., Int. Ed.* **2005**, *44*, 1504–1508.
- [8] N. D. Austin, N. V. Sahinidis, D. W. Trahan, *Chem. Eng. Res. Des.* **2016**, *116*, 2–26.
- [9] A. I. Papadopoulos, I. Tsivintzelis, P. Linke, P. Seferlis, *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier, Amsterdam **2018**. DOI: <https://doi.org/10.1016/B978-0-12-409547-2.14342-2>
- [10] A. S. Alshehri, R. Gani, F. You, *Comput. Chem. Eng.* **2020**, *141*, 107005. DOI: <https://doi.org/10.1016/j.compchemeng.2020.107005>
- [11] C. Gertig, K. Leonhard, A. Bardow, *Curr. Opin. Chem. Eng.* **2020**, *27*, 89–97.
- [12] R. Gani, C. Jiménez-González, D. J. C. Constable, *Comput. Chem. Eng.* **2005**, *29*, 1661–1676.
- [13] R. Gani, P. A. Gómez, M. Folić, C. Jiménez-González, D. J. C. Constable, *Comput. Chem. Eng.* **2008**, *32*, 2420–2444.
- [14] M. Folić, R. Gani, C. Jiménez-González, D. J. Constable, *Chin. J. Chem. Eng.* **2008**, *16*, 376–383.
- [15] M. Folić, C. S. Adjiman, E. N. Pistikopoulos, *AIChE J.* **2007**, *53*, 1240–1256.
- [16] M. Folić, C. S. Adjiman, E. N. Pistikopoulos, *Ind. Eng. Chem. Res.* **2008**, *47*, 5190–5202.
- [17] M. H. Abraham, R. M. Doherty, M. J. Kamlet, J. M. Harris, R. W. Taft, *J. Chem. Soc., Perkin Trans. 2* **1987**, 913–920.
- [18] M. H. Abraham, R. M. Doherty, M. J. Kamlet, J. M. Harris, R. W. Taft, *J. Chem. Soc., Perkin Trans. 2* **1987**, 1097–1101.
- [19] A. Pfennig, *Thermodynamik der Gemische*, Springer Verlag, Berlin **2004**.
- [20] H. Struebing, Z. Ganase, P. G. Karamertzanis, E. Sioukrou, P. Haycock, P. M. Piccione, A. Armstrong, A. Galindo, C. S. Adjiman, *Nat. Chem.* **2013**, *5*, 952–957.
- [21] H. Struebing, S. Obermeier, E. Sioukrou, C. S. Adjiman, A. Galindo, *Chem. Eng. Sci.* **2017**, *159*, 69–83.
- [22] H. Eyring, *J. Chem. Phys.* **1935**, *3*, 107–115.
- [23] Q. Liu, L. Zhang, L. Liu, J. Du, Q. Meng, R. Gani, *Chem. Eng. Sci.* **2019**, *202*, 300–317.
- [24] S.-T. Lin, S. I. Sandler, *Ind. Eng. Chem. Res.* **2002**, *41*, 899–913.
- [25] Q. Liu, L. Zhang, K. Tang, Y. Feng, J. Zhang, Y. Zhuang, L. Liu, J. Du, *Chem. Eng. Res. Des.* **2019**, *152*, 123–133.
- [26] N. D. Austin, N. V. Sahinidis, D. W. Trahan, *Chem. Eng. Sci.* **2017**, *159*, 93–105.
- [27] N. D. Austin, N. V. Sahinidis, I. A. Konstantinov, D. W. Trahan, *AIChE J.* **2018**, *64*, 104–122.
- [28] A. Klamt, F. Eckert, W. Arlt, *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101–122.
- [29] C. Gertig, L. Kröger, L. Fleitmann, J. Scheffczyk, A. Bardow, K. Leonhard, *Ind. Eng. Chem. Res.* **2019**, *58*, 22835–22846.
- [30] D. Douguet, H. Munier-Lehmann, G. Labesse, S. Pochet, *J. Med. Chem.* **2005**, *48*, 2457–2468.
- [31] A. Bardow, K. Steur, J. Gross, *Ind. Eng. Chem. Res.* **2010**, *49*, 2834–2840.
- [32] J. Schilling, C. Horend, A. Bardow, *AIChE J.* **2020**, *66* (5), e16903.
- [33] P. Linke, A. Papadopoulos, P. Seferlis, *Energies* **2015**, *8*, 4755–4801.
- [34] C. Adjiman, A. Bardow, *Chem. Eng. Sci.* **2017**, *159*, 1–2.
- [35] E. Sioukrou, A. Galindo, C. S. Adjiman, *Chem. Eng. Sci.* **2014**, *115*, 19–30.
- [36] T. Zhou, K. McBride, X. Zhang, Z. Qi, K. Sundmacher, *AIChE J.* **2015**, *61*, 147–158.
- [37] A. Klamt, *J. Phys. Chem.* **1995**, *99*, 2224–2235.
- [38] T. Zhou, Z. Lyu, Z. Qi, K. Sundmacher, *Chem. Eng. Sci.* **2015**, *137*, 613–625.
- [39] J. Scheffczyk, L. Fleitmann, A. Schwarz, M. Lampe, A. Bardow, *Chem. Eng. Sci.* **2017**, *159*, 84–92.
- [40] J. Scheffczyk, P. Schäfer, L. Fleitmann, J. Thien, C. Redepenning, K. Leonhard, W. Marquardt, A. Bardow, *Mol. Syst. Des. Eng.* **2018**, *3*, 645–657.
- [41] www.thyssenkrupp.com/carbon2chem/de/carbon2chem
- [42] R. Gani, *Comput. Chem. Eng.* **2004**, *28*, 2441–2457.
- [43] L. Kröger, W. Kopp, K. Leonhard, *J. Phys. Chem. B* **2017**, *121*, 2887–2895.
- [44] L. Vereecken, D. Glowacki, M. Pilling, *Chem. Rev.* **2015**, *115*, 4063–4114.
- [45] J. Foresman, Æ. Frisch, *Exploring Chemistry with Electronic Structure Methods*, 3rd ed., Gaussian, Inc., Wallingford, CT **2015**.
- [46] P. Deglmann, I. Müller, F. Becker, A. Schäfer, K.-D. Hungenberg, H. Weiß, *Macromol. React. Eng.* **2009**, *3*, 496–515.
- [47] M. Peters, L. Greiner, K. Leonhard, *AIChE J.* **2008**, *54*, 2729–2734.
- [48] A. Hellweg, F. Eckert, *AIChE J.* **2017**, *63*, 3944–3954.
- [49] M. L. Coote, *Macromol. Theory Simul.* **2009**, *18*, 388–400.
- [50] J. Stephens, F. Devlin, C. Chabalowski, M. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [51] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [52] J. Zheng, Y. Zhao, D. Truhlar, *J. Chem. Theory Comput.* **2009**, *5*, 808–821.
- [53] H. Gottschalk et al., *J. Chem. Phys.* **2018**, *148*, 014301.
- [54] P. Atkins, R. Friedman, *Molecular Quantum Mechanics*, 5th ed., Oxford University Press, Oxford **2011**.
- [55] M. Frisch et al., *Gaussian 09, Revision D.01*, **2013**.
- [56] C. Riplinger, F. Neese, *J. Chem. Phys.* **2013**, *138*, 034106.
- [57] C. Riplinger, B. Sandhoefer, A. Hansen, F. Neese, *J. Chem. Phys.* **2013**, *139*, 134101.
- [58] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8*, e1327.

- [59] I. Funes-Ardoiz, R. Paton, *GoodVibes: GoodVibes 2.0.3*, **2018**. DOI: <https://doi.org/10.5281/zenodo.595246>
- [60] S. Grimme, *Chem. Eur. J.* **2012**, *18*, 9955–9964.
- [61] C. Eckart, *Phys. Rev.* **1930**, *35*, 1303–1309.
- [62] A. Ghysels, T. Verstraelen, K. Hemelsoet, M. Waroquier, V. van Speybroeck, *J. Chem. Inf. Model.* **2010**, *50*, 1736–1750.
- [63] A. Klamt, V. Jonas, T. Bürger, J. Lohrenz, *J. Phys. Chem. A* **1998**, *102*, 5074–5085.
- [64] *TURBOMOLE V7.0.1*, University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, **2015**.
- [65] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- [66] A. Becke, *Phys. Rev. A* **1988**, *38*, 3098.
- [67] J. Perdew, *Phys. Rev. B* **1986**, *33*, 8822.
- [68] J. Perdew, *Phys. Rev. B* **1986**, *34*, 7406.
- [69] A. Klamt, G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799–805.
- [70] *COSMOtherm, C3.0, release 1701*, COSMOlogic GmbH & Co KG, Leverkusen.
- [71] F. Eckert, A. Klamt, *AIChE J.* **2002**, *48*, 369–385.
- [72] C. Gertig, E. Erdkamp, A. Ernst, C. Hemprich, L. Kröger, J. Langanke, A. Bardow, K. Leonhard, *Reaction Mechanisms and Rate Constants of Auto-Catalytic Urethane Formation and Cleavage Reactions*, submitted.
- [73] H. Renon, J. Prausnitz, *AIChE J.* **1968**, *14*, 135–144.
- [74] M. Dewar, E. Zebisch, E. Healy, J. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- [75] O. Levenspiel, *Chemical Reaction Engineering*, 3rd ed., John Wiley & Sons, New York **1999**.
- [76] *MATLAB R2019b*, The MathWorks, Natick, MA **2019**.
- [77] C. Gertig, K. Leonhard, A. Bardow, *Comput. Aided Chem. Eng.* **2019**, *46*, 415–420.
- [78] *COSMOconf, 4.0*, COSMOlogic GmbH & Co KG, Leverkusen.
- [79] C. Six, F. Richter, Isocyanates, organic, in *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley-VCH, Weinheim **2000**.
- [80] T. Kaiser, A. Rathgeb, C. Gertig, A. Bardow, K. Leonhard, A. Jupke, *Chem. Ing. Tech.* **2018**, *90*, 1497–1503.
- [81] W. Leitner, G. Franciò, M. Scott, C. Westhues, J. Langanke, M. Lansing, C. Hussong, E. Erdkamp, *Chem. Ing. Tech.* **2018**, *90*, 1504–1512.
- [82] Y. Cao, H. Li, N. Qin, G. Zhu, *Chin. J. Chem. Eng.* **2015**, *23*, 775–779.
- [83] P. Wang, S. Liu, Y. Deng, *Chin. J. Chem.* **2017**, *35*, 821–835.
- [84] S. Ephraim, A. Woodward, R. Mesrobian, *J. Am. Chem. Soc.* **1958**, *80*, 1326–1328.
- [85] A. E. Oberth, R. S. Bruenner, *J. Phys. Chem.* **1968**, *72*, 845–855.
- [86] F. Merger, G. Nestler, R. Platz, F. Towae, H. Hellbach, *Patent DE3142627*, **1983**.
- [87] A. Ismail-Yahaya, A. Messac, *40th AIAA Aerospace Sciences Meeting & Exhibit*, Reno, NV, January **2002**.
- [88] A. Konak, D. W. Coit, A. E. Smith, *Reliab. Eng. Syst. Saf.* **2006**, *91*, 992–1007.
- [89] M. Döntgen, F. Schmalz, W. A. Kopp, L. C. Kröger, K. Leonhard, *J. Chem. Inf. Model.* **2018**, *58* (7), 1343–1355.
- [90] A. L. Dewyer, A. J. Argüelles, P. M. Zimmerman, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2018**, *8* (2), e1354.
- [91] L. C. Kröger, S. Müller, I. Smirnova, K. Leonhard, *J. Phys. Chem. A* **2020**, *124* (20), 4171–4181.
- [92] P. Atkins, J. de Paula, *Physical Chemistry*, 9th ed., Oxford University Press, Oxford **2010**.
- [93] G. Towler, R. Sinnott, *Chemical Engineering Design*, 2nd ed., Elsevier, Amsterdam **2012**.
- [94] K. Lucas, *Thermodynamik*, 7th ed., Springer Verlag, Berlin **2008**.