Chemical Engineering Journal 548 (2026) 180680

Contents lists available at ScienceDirect

Chemical Engineering Journal

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

Computer-aided bioprocess and solvent design for industrial methyl ketone production

Henry Hilker [a,b], Lukas Raßpe-Lange [c], Lukas Polte [a], Anita L. Ziegler [d], Clemens Kortmann [d], Alexander Mitsos [d,e,f], Andreas Jupke [a,b], Kai Leonhard [c,*]

[a] Fluid Process Engineering (AVT.FVT), RWTH Aachen University, Forckenbeckstr. 51, 52074 Aachen, Germany
[b] Institute for Bio- and Geosciences, Plant Science, IBG-2, Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Straße, 52428 Jülich, Germany
[c] Institute of Technical Thermodynamics, RWTH Aachen University, Schinkelstr. 8, 52062 Aachen, Germany
[d] Process Systems Engineering (AVT.SVT), RWTH Aachen University, 52074 Aachen, Germany
[e] JARA-ENERGY, 52056 Aachen, Germany
[f] Institute of Climate and Energy Systems: Energy Systems Engineering (ICE-1), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

ARTICLE INFO

Keywords:
COSMO-RS
CAMPD
COSMO-CAMPED
Methyl ketones
Solvent design
SimulKnock
OptKnock
Bioprocess design
Techno-economic analysis

ABSTRACT

Methyl ketones have recently been identified as a promising candidate for biofuels among other uses. They can be produced via biomass fermentation using *Pseudomonas taiwanensis* VLB120 Δ6 pProd. However, the literature does not contain a complete early stage design proposal including upstream bioreactor design and downstream product purification at industrial scale. We present a model-based design approach for methyl ketone production based on renewable resources, including fermentation and product purification. We apply OptKnock to propose beneficial gene knockouts in the strain and adapt our SimulKnockReactor methodology to optimize the bioreactor for the production of methyl ketones. We then apply an enhanced version of our COSMO-CAMPED framework for the downstream purification of the fermentation broth. Extraction-based separations offer substantial potential for efficient purification of chemical products, particularly from the dilute aqueous streams of biotechnological fermentations. Finding a solvent with the required thermodynamic properties, bio-compatibility, and benign environmental properties is vital for the overall process performance. We propose three gene deletions to *P. taiwanensis* VLB120 Δ6 pProd, namely, inorganic diphosphatase, polyphosphate kinase, and inorganic diphosphatase. We identify isopentyl isopropyl ether as most suitable extraction solvent among several others that outperform our literature benchmark of Di-n-butylether. We present an optimized layout for the combined upstream and downstream process that produces methyl ketones at a bulk production capacity of $1 \times 10^7\,\mathrm{kg\,a^{-1}}$ for a total cost of $4.2\,\mathrm{USD\,kg^{-1}}$.

1. Introduction

Climate change, energy crisis, and political uncertainties have prompted research towards sustainable processes replacing fossil-based production [1,2]. Methyl ketones, a valuable class used as solvents for paints, coatings, and adhesives [3] and promising diesel-range biofuels [4], represent prime candidates for renewable production pathways.

Of particular interest for these applications are medium length methyl ketones like 2-undecanone. In this work we present a methodology for the early stage design of a bioreactor process to produce medium chain methyl ketones alongside the subsequent purification of the methyl ketones in an extraction distillation process.

Several studies have investigated their production via fermentation using various microorganisms [5–8]. Hanko et al. [6] report a titer of $315\,\mathrm{mg\,L^{-1}}$ using an optimized strain of *Yarrowia lipolytica*. Furthermore, Goh et al. [5] produced methyl ketones with *Escherichia coli* and report a titer of $380\,\mathrm{mg\,L^{-1}}$. In a separate study, Dong et al. [8] achieved a titer of $1.1\,\mathrm{g\,L^{-1}}$ using *Pseudomonas putida*. In recent studies, an altered strain of *Pseudomonas taiwanensis* VLB120 was identified as a suitable microorganism for the production of methyl ketones [4,9]. In this context, co-feeding with ethanol has emerged as a promising strategy to enhance yield and reduce competition with the food industry by decreasing the glucose consumption [4,10]. This may become more important in the future as Power to X pathways for the production of ethanol are becoming more important [11]. However, in this study we will consider glucose as only substrate.

* Corresponding author.
E-mail address: kai.leonhard@ltt.rwth-aachen.de (K. Leonhard).

https://doi.org/10.1016/j.cej.2026.180680
Received 2 June 2026; Received in revised form 31 July 2026; Accepted 11 August 2026
Available online 9 September 2026
1385-8947/© 2026 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Although laboratory-scale production is feasible, three major barriers must be overcome for economically-viable industrial-scale production:

1. **Upstream scale-up**: Laboratory-scale titers of the product ($1\,g\,L^{-1}$ to $2\,g\,L^{-1}$) must be increased to achieve a space–time yield that enables industrial, continuous production on a ton-scale and the attainment of economic productivity goals.
2. **Downstream purification**: Energy-efficient separation processes are required to recover highly dilute methyl ketones from complex aqueous fermentation broths.
3. **Economic process optimization**: Minimum total cost process designs must be identified by integrating optimal fermentation performance with efficient downstream purification.

Upstream scale-up. Commercial bioreactor simulators are available for simulation-based evaluation of the fermentation step [12–14]; however, holistic optimal design requires coupled numerical optimization of operating conditions, reactor sizing [15], and microbial performance [16,17]. With this aim of coupling, Mahadevan et al. [18] proposed a framework known as "dynamic flux balance analysis" (dFBA). In dFBA, a model of cellular metabolism — more specifically, flux balance analysis (FBA) [19] — is embedded into the mass balances of batch fermentation. The dFBA framework was extended by Hjersted and Henson [20] to fed-batch fermentation and by Zhao et al. [21] to nonlinear cellular objectives. As a next step, Chang et al. [22] and Ploch et al. [23] incorporated the aspect of optimal fermentation design in dFBA, namely, for nonlinear model predictive control, using a simultaneous solution approach for fed-batch fermentation [22] and a sequential solution approach for batch fermentation [23]. Both worked with a metabolic network comprising fewer than 100 metabolites. To represent information from genome-scale metabolic networks for model predictive control of fed-batch fermentation, de Oliveira et al. [24] replaced the FBA with a surrogate model. Going beyond, we recently introduced a framework that extends fermentation design with embedded FBA to bioreactor design with embedded FBA. This framework is referred to as the "sequential SimulKnockReactor" approach [25] and can directly process genome-scale metabolic models in the embedded FBA. In the first step of this approach, an established strain design formulation, e.g., OptKnock [26], is applied to identify optimal gene deletions. In the second step, our framework SimulKnockReactor proposes an optimal bioreactor design for the modified strain based on the embedded FBA, mass and energy balances of continuous fermentation, as well as cost correlations.

Downstream purification. Purification of dilute methyl ketones from the aqueous fermentation broth requires energy-efficient separation, for which extraction-distillation is particularly suitable in biotechnological applications [27]. In extraction-distillation, the choice of solvent critically influences overall process performance [28]. Computer-aided molecular and process design (CAMPD) addresses solvent selection by systematically searching the molecular design space, predicting thermodynamic properties required for process simulation, and evaluating solvent candidate performance via mixed-integer nonlinear programs (MINLPs) [29,30]. Recent CAMPD strategies encompass screening [31–35] and simultaneous process/equipment/solvent design and optimization [36–41]. We developed [42] COSMO-CAMPED, an integrated framework for extraction-distillation that simultaneously optimizes molecular structure, process topology, equipment sizing, and economic performance. Herein, we extend COSMO-CAMPED to biotechnological applications by addressing solvent–water cross-solubility through three key enhancements: (a) integration of quantitative structure–activity relationship (QSAR) models to predict environmental, health, and safety properties of residual solvent in the raffinate recycled to upstream fermentation, (b) an advanced rectification model accounting for water entrainment in the organic extract phase which affects distillation efficiency and final product purity, and (c) a warm-start strategy that accelerates optimization convergence.

Economic process optimization. We combine upstream and downstream optimization in a sequential optimization approach for the production process: first using the sequential SimulKnockReactor approach to determine optimal gene deletions and optimal upstream bioreactor performance, followed by the extended COSMO-CAMPED framework to identify optimal solvents and downstream purification design. This strategy considers both upstream productivity and downstream separation efficiency, yielding minimum total cost production processes for industrial-scale methyl ketone manufacturing. The extended COSMO-CAMPED approach also introduces several new features, including automated EHS and toxicological assessment of solvent candidates and fully automated topology layout for the distillation columns in the process model.

2. Methods

2.1. Tailoring the sequential SimulKnockReactor framework for methyl ketone fermentation

To predict the optimal design of the microbial strain and the bioreactor, we use the sequential approach described in our previous publication [25]: In the first step, the microbial strain is optimized using OptKnock [26], and in the second step, the bioreactor is optimized using SimulKnockReactor [25]. We opted for this sequential approach since the simultaneous approach, which is also presented in Ziegler et al. [25], proved to be too computationally intensive for this study, especially due to the large envisaged production capacity (for further explanation, see the Supplementary Information). In general, one potential drawback of the sequential approach is that the proposed genetic modifications may not be optimal for the proposed bioreactor design. However, we found that this is unlikely in the case of methyl ketones, as a case study for a small production capacity of $1.0\times10^{2}\,kg\,a^{-1}$ showed that the proposed genetic modifications of the simultaneous approach were consistent with those of the sequential approach. Fig. 1 depicts the procedure of the sequential approach.

In the first step, we use an established strain design formulation, namely, OptKnock [26]. OptKnock is a bilevel optimization formulation for predicting optimal gene deletions in strains. Targeted gene deletions prevent the formation of enzymes that catalyze specific reactions. Without catalyst, these reactions can no longer take place, which we refer to as a reaction knockout in the following. OptKnock focuses on the microorganism: The upper level of OptKnock maximizes the product yield, while the lower level maximizes growth. The lower level is based on flux balance analysis [19], which is one of the most established frameworks in metabolic modeling. In flux balance analysis, a metabolic network of the microorganism is used to calculate the fluxes inside the cell. These metabolic networks include all possible reactions within the cell, their stoichiometry, and the corresponding enzymes catalyzing the reaction. Fluxes represent the molar flow through each reaction in the metabolic network and are calculated based on mass balances.

We execute OptKnock several times with a varying number of knockouts. The resulting product flux is expected to increase with increasing number of knockouts until a plateau is reached. We then select the number of knockouts where the product flux reaches the plateau and transfer the corresponding knockouts to a second step. By handing the knockouts over from step 1 to step 2, we assume that the knockouts do not depend on the production capacity. This assumption applied to acetate and formate production using the *E. coli* core network [25].

In the second step of the sequential approach, we apply the SimulKnockReactor [25] optimization framework to estimate the minimal costs and optimal operating conditions of the upstream process. This step involves bioreactor optimization with embedded flux balance analysis [19] and was first introduced in SimulKnockReactor. SimulKnockReactor is a bilevel optimization formulation that minimizes the total costs of a continuous stirred-tank bioreactor and maximizes growth

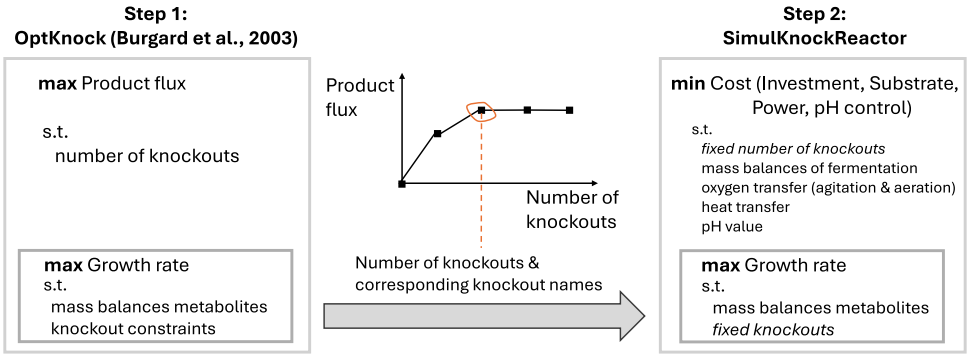


Fig. 1. Overview of the sequential approach. The first step is repeated with varying number of knockouts. With increasing number of knockouts, the product flux is expected to reach a plateau. The number of knockouts and corresponding knockout names of the point at which the plateau is reached are transferred to the second step.

of the microorganism. The upper level of the program represents the bioreactor based on mass balances, energy balances, and cost correlations including capital expenditure (CAPEX) and operating expense (OPEX). The lower level represents the microbial metabolism and, alike OptKnock, is based on flux balance analysis. Indeed, we observed a correlation between OptKnock and SimulKnockReactor in our previous publication [25]. Namely, OptKnock and SimulKnockReactor proposed similar knockouts. We therefore chose OptKnock to predict the reaction knockouts in step 1 and SimulKnockReactor to predict bioreactor cost and dimensioning in step 2. Thus, in this second step, the bioreactor is optimized subject to the metabolism with fixed reaction knockouts but variable fluxes. This way, the metabolic level still serves to refine the bioreactor prediction.

In the methyl ketone study, the two steps of the sequential approach are the following.

- **Step 1:** We perform OptKnock [19] several times for zero to five knockouts. We choose the smallest number of knockouts where maximum methyl ketone flux is reached and hand over the knockout number and names to step 2.
- **Step 2:** We solve SimulKnockReactor with fixed knockouts for a methyl ketone production capacity of 1×10^6, 2.5×10^6, and $5 \times 10^6\,\mathrm{kg\,a^{-1}}$ to find the optimal reactor setup.

In both steps, we use the genome-scale metabolic model of *P. taiwanensis* VLB120 Δ6 pProd, namely *i*JN1463_MK Δ6 pProd from our previous publication [10]. *P. taiwanensis* VLB120 Δ6 pProd produce methyl ketones as a blend of molecules of different chain length. To decrease numerical complexity, however, we choose 2-undecanone, one of the dominating compounds in the methyl ketone blend [9], as target chemical and representative compound for methyl ketones. Furthermore, we set glucose as substrate.

The adapted SimulKnockReactor formulation in step 2 has the following structure:

min costs (CAPEX, OPEX)
subject to CAPEX: reactor, compressor,
OPEX: substrate, cooling & pH-control agents, power,
aeration: oxygen uptake & transfer, agitation & compression power,
cooling: energy balance, heat transfer, pumping power,
pH: adjusting, growth inhibition,
bioreactor sizing, mass balances using Monod kinetics,
process parameters (cf. Table S3),
fixed knockouts,
max growth rate
subject to mass balances using *i*JN1463_MK Δ6 pProd,
fixed knockout constraints,
metabolic parameters (cf. Table S4).

The required power, air, and cooling streams, as well as the size of the reactor, are calculated without selecting specific devices, e.g., a specific type of stirrer. Tables S3 and S4 in the Supplementary Information depict the bioreactor and the metabolic parameters, respectively. For example, in the pH growth correlation, the pH and the temperature that are optimal for cell growth are set to 7 and 30 °C, respectively [9,43]. For setting minimum and maximum pH for cell growth, we assumed that the strain *P. taiwanensis* VLB120 Δ6 pProd behaves towards pH similar to *E. coli*. In the absence of strain-specific data, the affinity constant for Monod kinetics is retrieved from Choi et al. [44] for a *P. putida* strain. The optimal growth rate of *P. taiwanensis* VLB120 Δ6 pProd is determined from growth data published by Grütering [43]. We fit the data linearly between 28 °C and 32 °C (with a rather low coefficient of determination $R^2 = 0.31$ due to the wide dispersion of the data). Using the linear fit, we calculate the optimal growth rate at 30 °C as $0.31\,\mathrm{h^{-1}}$. Additionally, we use the substrate uptake of $3.05\,\mathrm{mmol\,g_{CDW}^{-1}\,h^{-1}}$ from Nies et al. [9] and set the upper bound of the substrate feed concentration to $100\,\mathrm{g\,L^{-1}}$. Furthermore, in accordance with the conclusion in Ziegler et al. [25] that large reactor volumes should be preferred over a large number of reactors, we omit numbering-up of reactors and set the possible reactor size to $1 \times 10^3\,\mathrm{m^3}$ during the optimization. A full description of the underlying economic method is presented in Ziegler et al. [25].

2.2. Enhancing the COSMO-CAMPED framework and adapting to methyl ketone purification

The fermentation broth obtained from the optimization results of SimulKnockReactor defines the separation problem to be solved with the COSMO-CAMPED Framework. Purification of the highly diluted methyl ketones from the aqueous fermentation broth is achieved initially by using a counter-current extraction column to transfer the methyl ketones into an organic solvent phase. Subsequently, rectification of the organic extract is performed to obtain high-purity methyl ketones and to recycle the solvent. The raffinate is assumed to be recycled, thus solvent-loss into the raffinate phase is neglected. The overall purification target is a recovery of 98 % methyl ketone at 99 % purity. Since the COSMO-CAMPED framework is limited to ternary mixtures, we make assumptions to apply the framework.

Experiments show that the fermentation broth is a complex aqueous solution, consisting of a range of different methyl ketones, water, biomass, by-products, and unconverted reactants. We simplify this complex mixture by selecting 2-undecanone as the product and water as the carrier component. We apply this simplification for the following reasons: First the substrate concentration is very low compared to the product concentration. Second, since glucose is used as a substrate, and is a very hydrophilic molecule with a LogP value of 2.8. Its transfer to the organic phase will be negligible.

Third, we neglect the presence of biomass in the organic phase by assuming that it can be filtered out prior to extraction. In practice, such a filtration step is technically feasible because it would primarily remove suspended biomass from the fermentation broth without affecting the dissolved methyl ketones. However, the filtration step is not modeled explicitly in this study. Although filtration of biomass is possible, the influence of biomass on liquid-liquid extraction is not straightforward. Eberz et al. [45] show that the presence of biomass can either increase or decrease phase separation performance depending on the solvent. No general trend could be identified since combined effects of cells, solvent and further dissolved components in the fermentation broth affect this behavior.

As already performed for the upstream calculation, we again choose 2-undecanone as the representative for all methyl ketones since this molecule has been experimentally detected as one of the most dominant methyl ketones in the mixture [9]. The bioreactor optimization results specify the binary feed stream into the extraction section.

An extraction distillation process enables significant energy savings for highly diluted aqueous product mixtures, but the performance of this separation sequence is critically dependent on the solvent choice. Thus, to find the optimal solvent, the COSMO-CAMPED framework is applied to simultaneously optimize the molecular solvent structure and the downstream process. The downstream process is optimized for minimal total costs, incorporating operating and investment costs, by using the feed-to-solvent ratio as a variable. The optimization is formulated as a Mixed-Integer Non-Linear Problem (MINLP). The objective function describes the total cost per unit product and is minimized. The MINLP depends on the process variables x and solvent structure y which is part of the available molecular design space Y. Equality constraints $\mathbf{g_1}$ and $\mathbf{g_2}$ describe the process model and thermodynamic model, respectively. Inequality constraints $\mathbf{h}$ describe the operating limits, and inequality constraints $\mathbf{c}$ account for the molecular constraints.

$$
\begin{aligned}
&\textbf{min costs } (\textbf{CAPEX},\ \textbf{OPEX};\ \mathbf{F}(\mathbf{x},\mathbf{y})\\
\text{subject to } &\textbf{CAPEX}:\ \textbf{process model } g_1(x,y)=0\\
&\quad \text{feed to flow ratio, topology, column height,}\\
&\quad \text{column diameter, heat exchanger area, pump head}\\
&\quad \textbf{OPEX}:\ \textbf{thermodynamic model } g_2(x,y)=0\\
&\quad \text{feed to flow ratio, steam, cooling water,}\\
&\quad \text{electricity, labor-related operations}\\
&\quad \text{maintenance, overhead, property taxes and insurance}\\
&\quad \textbf{operating limits } h(x,y)\le 0\\
&\quad \text{product purity, maximum column height, recovery}\\
&\quad \textbf{molecular constraints } c(y)\le 0\\
&\quad \text{atom count, boiling point, interfacial tension,}\\
&\quad \text{LogP, Flash Point, EHS Score}\\
&\quad x\in\mathbf{R}^n, y\in Y
\end{aligned}
\tag{1}
$$

For full insights into the structure of the COSMO-CAMPED framework and the calculation methods to obtain thermodynamical properties, the reader is referred to Polte et al. [42]. In general, the framework consists of three main building blocks, i.e., the design space exploration, the property prediction, and the performance evaluation using a process model simulating the purification process.

For the design space exploration, the Genetic Algorithm (GA) generates solvent candidates and checks them for user-specified molecular constraints. Next, the thermodynamic model COSMO-RS [46] is applied for thermo-physical property prediction. Using these thermophysical properties, the process model subsequently simulates possible extraction-distillation variations and selects the optimal variant while the process parameters are optimized for minimal combined investment and operating costs. After a whole generation of solvent candidates has been evaluated within the process model, each solvent candidate's performance in terms of total costs is returned to the GA. The GA builds a ranking based on the calculated total costs and creates another generation of molecular structures on the basis of the previous results. The framework repeats these steps iteratively until a user-defined termination criterion, such as a certain number of generations, is fulfilled.

In the following, we introduce three extensions to the COSMO-CAMPED framework: (i) an efficient initialization method for the genetic algorithm, (ii) prediction of environmental, health, and safety (EHS) properties of the solvent in the raffinate phase, and (iii) an advanced rectification model to describe water entrainment in the organic extract.

Initializing the genetic algorithm

To improve the efficiency, robustness, and effectiveness of the GA, we implement a screening-guided warm start approach to initialize the GA. Instead of randomly creating the first generation of solvent candidates, we use existing knowledge about the target molecule and incorporate it into the first generation. One way to gather knowledge about the target molecule is, for instance, to either screen suitable molecules known from literature using process simulations or screen a database with a large number of know molecules. The best candidates from the screening will then be incorporated into the first generation of the design study. This approach has been shown to reduce the computing time of a CAMD study by up to 70 % while discovering a multitude of top-performing solvent candidates [47]. Since the COSMO-CAMPED framework applies the fragment-based GA LEA3D [48], the molecular structures are decomposed to create a database of the most frequently observed fragments within the best screening molecules. Therefore, we apply the fragmentation tool developed by Wang et al. [47].

The procedure to use the fragmentation tool for initializing the GA consists of three steps:

1. Perform a solvent screening study to search a database for promising candidates. The best performing candidates are then passed to the fragmentation tool.
2. Utilize the fragmentation tool [47] to generate a fragment library from the fragments of the evaluated candidates.
3. Perform a solvent design study using the fragment library from the fragmentation tool as input as well as the ranking of top candidates from the screening as initial generation.

Predicting environmental, health, and safety properties

COSMO-CAMPED can generate completely unknown solvent candidates. These candidates, therefore, have unknown Environmental, Health and Safety (EHS) properties. While these properties often do not have a direct impact on the performance in the process simulation they are detrimental to the decision whether the candidate is suitable or not.

For instance, a highly toxic solvent might perform well in separation but would be discarded due to complex and costly raffinate treatment requirements. Therefore, an important extension to the COSMO-CAMPED framework is to incorporate EHS factors directly into the solvent design. Thus, we can steer the design of the process already at an early stage towards benign solvents that comply with the Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) guidelines, as published by the European Chemicals Agency [49]. We evaluate how benign a potential solvent is, based on three criteria: Flammability of the potential solvent, its compatibility with the bio organism in the fermentation, and the potential solvent's toxicity. The flash point is an indicator for a molecule's flammability [50], the logP value is the logarithmic partition coefficient of the solvent in octanol and water and can indicate whether a proposed solvent is harmful to the bio-organism used in the fermentation [4]. Bio-compatibility is particularly relevant, as the process assumes a recycle of the raffinate phase into the fermentation step, which requires the solvent to be

Table 1
QSAR models, respective predicted property, toolbox where they are usable, and reference.

QSAR Model	Property predicted	Toolbox	Reference
Fish acute toxicity	LC50	Janus	[58]
Fish chronic toxicity	LC50	Janus	[58]
Daphnids acute toxicity	LC50	Janus	[58]
Daphnids chronic toxicity	LC50	Janus	[58]
Algae acute toxicity	LC50	Janus	[58]
Algae chronic toxicity	LC50	Janus	[58]
CAESAR model	BCF	Janus	[59]
Meylan model	BCF	Janus	[60]
kNN-read across	BCF	Janus	[60]
CATMoS	LD50	Opera	[56]
OPERA	BCF	Opera	[55]

compatible with the microorganisms involved. Toxicity is measured in different ways depending over which pathways a substance enters the environment. We consider three values to describe the toxicity of a potential solvent. First, the Lethal Dose 50 (LD50) measures the toxicity of a substance when ingested. It describes the point at which a substance kills 50 percent of the test population and expresses this as dose per body weight of the test subject. Second, the Lethal Concentration 50 (LC50) analogously measures the concentration in an environment, usually water, when 50 percent of a test species die. Finally, the Bioaccumulation Factor (BCF) measures how much a substance accumulates in an organism compared to the environment [51].

We calculate the flash point and LogP value as properties directly from COSMO-RS [52,53]. For the prediction of LD50, LC50, and BCF, we implement two EHS property prediction toolboxes, namely, the Janus models that are part of the Vega Toolbox [54] and Opera [55–57]. Both are publicly available software tools and comprise several Quantitative Structure-Activity Relationship (QSAR) models that are trained on experimental data. For the LD50, LC50, and BCF, both toolboxes offer multiple models, as shown in Table 1.

In order to concatenate our results into one single score we need to normalize the predicted value as BCF, LD50 and LC50 are measured on different scales. Therefore, we apply the method proposed Pizzo et al. [61] that normalizes all predicted values into a range between 0 and 1. For BCF, LD50 and LC50 a value of one represents a very hazardous substance and a value of 0 represents a very benign substance. For an indepth description of the normalization we refer the reader to Pizzo et al. [61]. Furthermore, the QSAR models for the toxicity prediction also yield a value for the reliability of the prediction. The reliability value depends on how similar the molecule under consideration is to the training data. [57]. We introduce a correction term of our normalized predicted BCF, LD50 and LC50 values with the reliability of the corrections as proposed by Pizzo et al. [61], for more information on the method we would also like to refer the reader to the original paper.

Finally, we accumulate all BCF, LD50, and LC50 predictions into one EHS score, S_{EHS}, as proposed by Linke et al. [34]:

$$S_{\mathrm{EHS}} = \frac{O_{\mathrm{LD50}} + O_{\mathrm{BCF}} + J_{\mathrm{LC50}} + J_{\mathrm{BCF}}}{4} \quad ,$$

where O_{LD50} and O_{BCF} are the accumulated and normalized values of all QSAR Opera models for LD50 and BCF that have been adjusted for uncertainty. Analogously, J_{LC50} and J_{BCF} represent the LC50 and BCF predictions from the Janus QSAR models.

To take Environmental Health and Safety (EHS) properties into account and to sort out solvent candidates that are physically infeasible or cannot be predicted accurately, we define a number of constraints that must be met (see Table 2). In case the constraints are not satisfied, the respective molecule is discarded from the design process. The constraints for the EHS properties are based on thresholds specified within the REACH legislation [49]. Molecular constraints are included

Table 2
Molecular and environmental, health and safety constraints applied in the screening and design studies.

Property	Symbol	Constraint for rejection
Atom count	n_{A}	$n_{\mathrm{A}} > 50$
Boiling point	T_{b}	$T_b < 299\,\mathrm{K} T_{\mathrm{b}} > 550\,\mathrm{K}$
Interfacial tension	σ	$\sigma \leq 0\,\mathrm{N\,m^{-1}}$
Octanol-water partition coefficient	Log P	$0 < \mathrm{Log\,P} < 3.4$
Flash point	T_{F}	$T_{\mathrm{F}} < 296\,\mathrm{K}$
Median lethal dose	LD_{50}	$LD_{50} < 2000\,\mathrm{mg\,kg^{-1}}$
Bioconcentration factor	BCF	$\mathrm{BCF} > 2000\,\mathrm{L\,kg^{-1}}$,
Aquatic toxicity	LC_{50}	$LC_{50} < 0.1\,\mathrm{mg\,L^{-1}}$

to check whether the boiling point and the molecule size are within the scope of the design study. The interfacial tension is calculated to determine whether the solvent phase and the carrier phase show a liquid-liquid equilibrium. A calculated value of zero – or a negative value resulting from COSMO-RS estimation errors – indicates complete miscibility [62]. Such cases are therefore excluded from further calculations. All applied constraints and their respective boundary conditions are listed in Table 2. The pure component properties for water and 2-undecanone, which are used in this study, are shown in Tables S6 and S7. All parameters that are used for the extraction-distillation simulation are given in Table S8.

Upgrading the process model

The COSMO-CAMPED approach assesses solvent performance by designing, simulating, and optimizing extraction-distillation sequences to minimize total downstream costs. The overall framework architecture of the COSMO-CAMPED process model, consisting of an extraction column model and a rectification model, remains as previously described [42]. Specifically, the hybrid rate-based extraction model, which accounts for fluid dynamics and column dimensioning, was retained without modification. In this work, we substantially revised the rectification model for product purification following extraction and the associated cost calculation methodology to extend the framework's applicability to more complex, non-ideal, multi-component mixtures. In doing so, the updated process model advancement accounts for the water cross-solubility in organic phases and its effect on the distillation sequence, enabling accurate prediction of energy demands and column dimensions for a realistic economic evaluation of downstream processes in biotechnological fermentation broths. The previously implemented Fenske-Underwood-Gilliland (FUG) method assumed ideal vapor–liquid equilibria and was restricted to binary mixtures. The revised process model integrates the Rectification Body Method (RBM), a pinch-based approach that enables rigorous modeling of multi-component, non-ideal mixtures, including those exhibiting azeotropic behavior [63].

The RBM constructs geometric approximations of all feasible vapor–liquid equilibrium profiles within the column composition space. The method identifies all pinch points (compositions where net driving force for mass transfer approaches zero), classifies them according to their thermodynamic and geometric stability and connects them via thermodynamic consistent paths using eigenvalue stability analysis [63].

For standard distillation columns, the RBM separately analyzes the rectifying (above feed) and stripping (below feed) sections. For each of these column sections, linear rectification bodies are constructed by connecting identified pinch points with the product and feed compositions. Each rectification body represents a polyhedral approximation of the feasible vapor–liquid equilibrium (VLE) region in composition space, within which overall and component mass as well as energy balances are simultaneously satisfied. Iterative reflux adjustment until the rectification bodies of adjacent column sections geometrically overlap

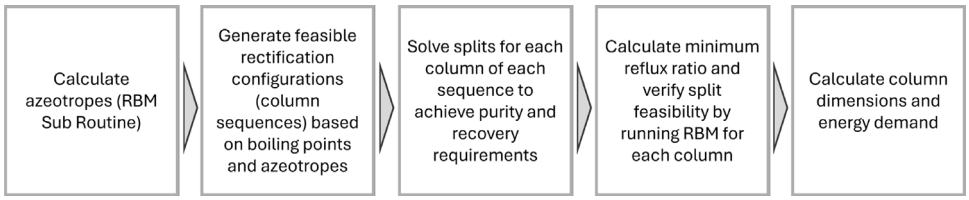


Fig. 2. Overview of the new approach for finding and simulating feasible rectification sequences to purify the product from the extract stream. The approach utilizes Rectification Body Method subroutines for identifying existing azeotropes first, and after generating possible configurations in terms of column sequences, simulating each rectification column to calculate the minimum reflux ratio and very split feasibility.

defines minimum reflux ratio, minimum energy demand (MED), and separation feasibility [63].

Unlike the boundary value method (restricted to four components) [64], the RBM has been fully automated through computational geometry algorithms that check polyhedron intersections in multidimensional composition space, enabling arbitrary component counts. An open-access implementation is available via AVT.SVT, RWTH Aachen University [65] which provides C-coded functions for rigorous thermodynamic calculations, particularly for finding homogeneous and heterogeneous azeotropes.

The RBM identifies the minimum energy demand and verifies split feasibility for a predefined split, which is specified by a combination of feed flow and composition with either the bottom or the distillate flow and composition. Since each RBM evaluation requires a specified split as input, extensive research has been conducted on automated process synthesis approaches to systematically generate feasible distillation sequences, i.e., combinations and orderings of individual columns whose corresponding splits achieve a desired overall separation while respecting the thermodynamic topology of the mixture. For comprehensive automated process synthesis approaches, see [66–68]. The focus of this work, however, is not to develop a fully automated process synthesis tool, but rather to implement a practical approach for applying the RBM to the treatment of the extract phase. To this end, a hierarchical screening framework for possible distillation sequences based on boiling temperatures and azeotropic points was implemented, followed by a subsequent algorithm for split calculation (see Fig. 2).

The procedure begins with azeotrope determination, implemented as a subroutine of the RBM software package. The method systematically identifies homogeneous and heterogeneous azeotropes by finding stationary points, using rigorous VLE data from non-random two-liquid (NRTL) model parameters, Antoine equations, and phase stability tests. Based on the identified azeotropes and the mixture topology, possible column sequences are generated in a hierarchical manner, i.e., by successively refining from simple single-column configurations to more complex sequences only if required by the thermodynamic constraints (mid boiling target components or presence of azeotropes). Mixture systems with more than one heterogeneous or homogeneous azeotrope in the distillation region of interest are penalized and thus excluded by the genetic algorithm. Note that this exclusion of solvents represents a conservative assessment that may occasionally exclude solvents for which sequences could be found through an advanced distillation process synthesis. For all remaining mixture systems with at most one relevant azeotrope, four principal configurations are implemented.

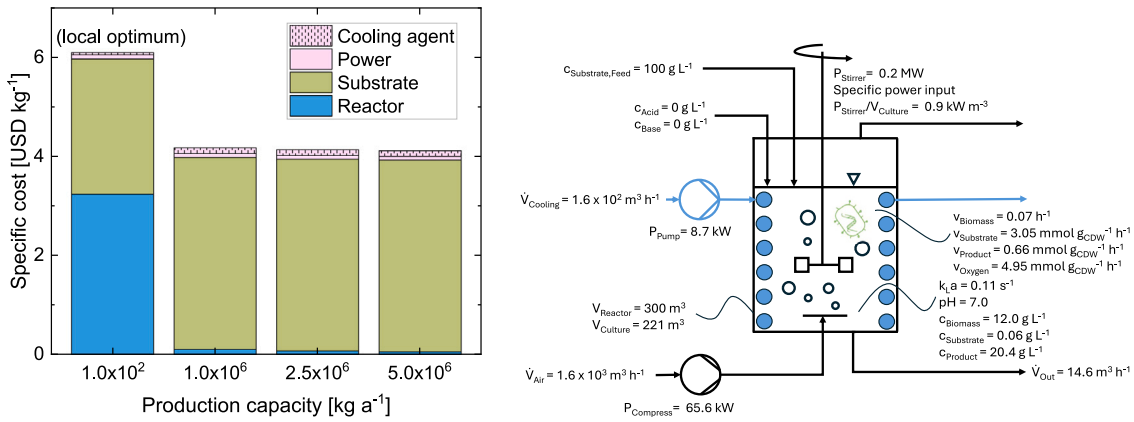
1. Single-Column Rectification (One Column, Ambient Pressure): Applied when the product is the lightest or heaviest boiling component and no azeotrope exists. Split factors for each component are determined by solving coupled mass balances with inequality constraints ensuring product purity.
2. Two-Column Sequential Rectification (Two Columns, Ambient Pressure): Employed when the product is the medium-boiling component, preventing single-column separation. Two configurations are evaluated: Light-boiling component removed first, or heavy-boiling component removed first. Split factors for both columns are solved simultaneously through coupled nonlinear equation systems, with mass balance closure verified within predefined error thresholds.
3. Pressure-Swing Distillation (Three Columns, Variable Pressure): Utilized when a homogeneous binary azeotrope between solvent and product prevents separation at atmospheric pressure. The method identifies pressure combinations where azeotropic composition shifts sufficiently to enable separation through recycle streams. RBM calculations are performed independently for each column at its assigned pressure, with feasibility checks ensuring rectification bodies remain realizable in composition space.
4. Heteroazeotropic Rectification with Decanter (Three Columns plus Decanter): Applied when a heterogeneous azeotrope (two immiscible liquid phases at vapor equilibrium) exists. A decanter exploits the miscibility gap to separate the azeotropic mixture into two distinct liquid phases, each fed to a dedicated rectification column. Component compositions for each decanter outlet are determined directly from the binary phase diagram; no additional composition calculation is required.

After selecting all feasible configurations, a system of equations is solved for each variant to determine the splits for all columns within that configuration. This system of equations accounts for the required product purity of 99.0 % mol, a product recovery of 99.0 %, the boiling order of components, and the presence of azeotropes. Using these calculated splits as input, the RBM is then applied to each column of each configuration to compute the minimum reflux ratio and verify the thermodynamic feasibility of every individual split.

Following RBM calculation of minimum reflux ratio $R_{\min}$, the actual operating reflux is set conservatively to $R = 1.3 \cdot R_{\min}$. The minimum number of theoretical stages $N_{\min}$ for the two key components (i.e., light and heavy boiling) is calculated using the Fenske equation, with the relative volatility taken as the geometric mean of the values computed at distillate and bottoms conditions [69]. For ternary mixtures, the third component is grouped with the heavy boiling component, and combined compositions and equilibrium constants are used. The Gilliland correlation then yields the actual number of theoretical stages N_{th} from $N_{\min}$ and the reflux ratio [70]. Column height and diameter are determined from the stage number, plate spacing, and flooding correlations, accounting for liquid and vapor loads, as described in [28]. Note that all detailed technical specifications are already described in Polte et al. [42]. Energy requirements for the reboiler and condenser are calculated from energy balances at operating conditions. Heat transfer areas are derived from overall heat transfer coefficients and temperature driving forces, accounting for the temperature-dependent properties.

Column dimensions derived with the RBM-based rectification model and the rate-based extraction model, together with the corresponding energy duties, are used as input for the economic evaluation. In contrast to previous COSMO-CAMPED studies, the cost calculation is based on the approach by Seider et al. [71] for estimating both investment and operating costs.

Investment costs are obtained from free-on-board (FOB) equipment costs (cost of equipment ready for shipment from supplier) using



(a) Cost factors (b) Process conditions for a methyl ketone production of 2.5×10^6 kg a^{-1}

Fig. 3. Bioreactor cost factors and operating conditions. The embedded network is *i*JN1463_MK Δ6 pProd [10]. (a) Cost factors for 1×10^2, 1×10^6, 2.5×10^6, and 5×10^6 kg a^{-1} methyl ketone production. Blue, green, and pink denote investment costs, raw materials, and utilities, respectively. For 1×10^2 kg a^{-1} methyl ketone production, the knockouts are not fixed beforehand and substrate uptake is set to 10 mmol g_{CDW}^{-1} h^{-1}. In the other three cases, the knockouts are fixed beforehand and substrate uptake is set to 3.05 mmol g_{CDW}^{-1} h^{-1} [9]. (b) Methyl ketone production conditions for 2.5×10^6 kg a^{-1}. Figure adapted from Ziegler et al. [25]. c: concentration, P: power, USD: US-Dollar, v: metabolic flux, V: volume, V̇: volume flow.

equipment-specific cost correlations and size factors, which depend on column height and diameter, heat exchanger area, pump head, and flow rates. These FOB costs are adjusted for specific material, pressure, and design factors, and updated to the target year via the Chemical Engineering Plant Cost Index. The resulting CAPEX is annualized over a depreciation period of 10 years at 8400 operating hours per year and normalized by the product mass flow to yield specific investment costs [71].

The operating cost model accounts for steam at multiple pressure levels, cooling water, electricity, maintenance, overhead, and property taxes and insurance. Utility costs are calculated from the RBM-derived reboiler and condenser duties and the pump power demand. Operating overhead represents the indirect costs of plant operation, such as administration, employee services, safety, logistics, and general site services. Overhead scales with the number of units rather than their size, as it is calculated as a percentage of maintenance and operations salaries, wages, and benefits. Solvent make-up costs are neglected under the assumption of complete solvent recovery. Total process costs are finally obtained as the sum of specific investment and operating costs and serve as the objective function within the COSMO-CAMPED optimization framework. Further information on the applied cost correlations are documented in 13.

3. Results

3.1. Fermentation costs and conditions

In this section, we present the results for bioreactor costs and optimal fermentation conditions. In our first step of strain design using OptKnock [26], we find that there is no methyl ketone production without knockouts. This finding matches the fact that the production of methyl ketones is not growth-coupled in *P. taiwanensis* VLB 120 Δ6 pProd and, in coherence with that, also in the metabolic network *i*JN1463_MK Δ6 pProd. Growth-coupled production implies that the organism necessarily produces the target chemical during its growth phase. Strains with growth-coupled production are generally less susceptible to metabolic evolvement, which makes them attractive for continuous fermentation. Note that, since we only consider growth-coupled production, we assume that metabolic evolvement is not significant in our case.

Also with one and two knockouts, no methyl ketone production is observed. In fact, growth-coupled production is achieved with three or more knockouts. Increasing the number of knockouts to more than three, however, does not lead to a further increase in methyl ketone production. Hence, we fix the number of knockouts to three and hand over the following knockouts to the next step: PPA (inorganic diphosphatase), PPK (polyphosphate kinase), and IDPh_1 (inorganic diphosphatase). With these knockouts in place, there is no direct pathway towards biomass growth anymore; all remaining pathways towards biomass growth also include the methyl ketone reaction.

In the next step, we determine the optimal bioreactor conditions, costs, and metabolic state at different production scales. Fig. 3(a) depicts the specific total costs and the cost factors for a methyl ketone production of 1×10^6 kg a^{-1} to 5.0×10^6 kg a^{-1}. For comparison with small-scale production and as a case study of the simultaneous approach, we added the cost factors for a methyl ketone production of 1.0×10^2 kg a^{-1}. The optimization for the whole envisaged production capacity of 1×10^7 kg a^{-1} proved to be infeasible with the set variable bounds, e.g., maximum reactor size of 1×10^3 m^3.

Overall, for the regarded production capacities, total production costs range from 4.1 to 6.1 USD kg^{-1}, depending on the production capacity. There is a steep decrease in specific costs when increasing the production capacity from 1×10^2 to 1×10^6 kg a^{-1}; then the costs gradually slightly decrease with increasing capacity. This behavior can be explained by an altered distribution of costs. At a low production capacity of 1×10^2 kg a^{-1}, investment costs (3.2 USD kg^{-1}) dominate over raw material costs (2.7 USD kg^{-1}). Investment costs, in turn, comprise almost entirely of reactor costs in all regarded cases; the results show that compressor costs are negligible (maximum 1.9×10^{-2} USD kg^{-1}). Furthermore, raw material costs consist entirely of substrate costs; pH adjustment costs are zero, since no pH adjustment agent is used in the optimal solution (see Fig. 3(b) for an example).

Increasing production capacity results in economies of scale in investment costs; we do not account for economies of scale for raw material costs and power. When increasing the production capacity from 1×10^2 kg a^{-1} to 1×10^6 kg a^{-1}, reactor investment costs decrease from 53 % to 3 % of the total costs. A further increase in production capacity to 2.5×10^6 kg a^{-1} and 5×10^6 kg a^{-1} also reduces specific reactor costs, i.e., reactor costs per kilogram product, but this has only a negligible impact on the overall cost distribution, where the share of reactor cost decrease to 2 %.

For industrial-scale production, substrate costs are the largest factor (93 % to 94 %), which aligns with common knowledge. Substrate costs

are followed by costs for the cooling agent, power costs, and reactor investment costs. In our framework, power comprises compressor power, stirrer power, and power for the cooling agent pump. Among these three, stirrer power is the largest contributor (see Fig. 3(b) for example values). .

In the optimization run with a production capacity of $1.0\times10^2\,\mathrm{kg\,a^{-1}}$, we use the simultaneous approach of SimulKnockReactor, i.e., the knockouts are not fixed beforehand. This optimization run is a case study where we aim to compare the simultaneous approach with the sequential approach. The simultaneous approach results in a program with approximately 3000 binary variables additional to the 1.4×10^4 continuous variables, which makes the program numerically demanding. Therefore, within a given runtime of 10 h on 48 cores of a high-performance computing cluster, the program returns a local solution rather than the global one. For comparison, OptKnock and SimulKnockReactor with fixed knockouts solve within seconds or minutes on a desktop computer. Also for numerical reasons, the substrate uptake is set to $10\,\mathrm{mmol\,g_{CDW}^{-1}\,h^{-1}}$ in the run without fixed knockouts. This elevated uptake makes the program easier to solve but may also result in decreased costs, since, for example, smaller reactors are required. Furthermore, note that the power of the cooling water pump is rated higher than in the other runs by a factor of 1×10^3 due to an initial setting; however, with negligible effect on the final cost scheme due to the low impact of power. Despite these discrepancies between the cases, we still consider a comparison to be useful, since the knockouts are ultimately identical. Namely, the knockouts proposed by OptKnock are the same as the ones proposed by the simultaneous approach of SimulKnockReactor for a production capacity of $100\,\mathrm{kg\,a^{-1}}$ and the number of knockouts fixed to three. This agreement in results once again demonstrates the correlation between SimulKnockReactor and OptKnock with regard to the proposed knockouts.

Considering the fermentation conditions in Fig. 3(b), it is apparent that the substrate concentration is low, as would be expected in continuous fermentation. Furthermore, the pH value is 7.0, which means that growth is not inhibited.

In general, a methyl ketone production of $2.5\times10^6\,\mathrm{kg\,a^{-1}}$ is a reasonable production capacity per reactor in order to achieve the full capacity of $1\times10^7\,\mathrm{kg\,a^{-1}}$. The results are also realistic in terms of reactor size [72–74] ($300\,\mathrm{m^3}$, with a corresponding height of 16.4 m). In order to reach the envisaged total production capacity of $1\times10^7\,\mathrm{kg\,a^{-1}}$, we assume that four reactors, each producing $2.5\times10^6\,\mathrm{kg\,a^{-1}}$ methyl ketones, operate in parallel.

On the basis of the presented upstream results, in a next step, the downstream process is designed. The binary feed stream into the extraction section with a volume flow of $58.4\,\mathrm{m^3\,h^{-1}}$ and a product concentration of $20.4\,\mathrm{g\,L^{-1}}$, which defines the separation problem.

3.2. Solvent design and downstream process optimization

To apply the screening-guided warm start method, we perform two successive CAMPED studies. In the first step, a screening study is conducted to screen a literature-derived database for potential solvents and their respective optimized downstream process design. Subsequently, a design study is conducted in which novel solvent structures are created based on the screening study results.

3.2.1. Screening study results

We assemble a molecular database from literature data [4] and add further solvent-like molecule structures. The literature data includes 22 pre-screened solvents which have been experimentally tested for their ability to extract methyl ketones out of the fermentation broth [4]. Adding further solvent structures to increase the number and variety of the available fragments and structures, we perform the screening study with 33 solvents as listed in Table S1.

From the initial population, 16 candidates finish the process evaluation, and the minimum total costs of the purification process could be evaluated. Thirteen candidates are rejected due to molecular or EHS constraints, and five further candidates show topologies that cannot be separated by the implemented process variations. The results of all successful solvent candidates are sorted for their minimal total costs in Table 3. From 16 solvents that finish the process simulation by successfully determining their minimal total costs, the process model selects the two-column rectification variant in eight cases. A two-column rectification variant is necessary because the respective candidates have higher boiling points compared to the product, which is therefore the medium-boiling component. For the remaining eight candidates, the process model selects the one-column rectification variant. In all two-column variant cases, the total costs of the purification process exceed $100\,\mathrm{USD\,t^{-1}}$ while the majority of all one-column variations stay below $100\,\mathrm{USD\,t^{-1}}$. This is mainly due to higher investment costs and higher steam costs that are required to operate two columns. The best performing candidate from the screening study is Di-n-butylether with total costs of $37.04\,\mathrm{USD\,t^{-1}}$. It serves as a benchmark solvent in the following design study. Using the fragmentation tool, introduced in Section 2.2, we derive a fragment database that is shown in Table S2.

The fragment database is composed of 22 fragments. Although only carbon oxygen and hydrogen molecules are occurring in the fragments this fragment database is sufficiently diverse to span a large design space.

The resulting ranking in Table 3, as well as the fragment database, is used to initialize the GA in the following design study.

3.2.2. Design study results

Applying the results from the previous screening study to initialize the GA, the design study runs for 30 generations, each creating a population of 20 solvent candidates. This makes it theoretically possible to investigate a maximum of 600 solvent structures per design study; however, this number is not reached because duplicates are possible. To assess the reproducibility of the results, we perform five independent runs of the design study. When conducting design studies on a high-performance cluster, approximately 7000 core hours per study are required on average. The investment is justified by the results obtained. Early-stage design of a fermentation process typically requires substantial resources, and the cost of simulation is undoubtedly lower than that of experimental investigation. Access to high performance computing systems becomes more available which makes the usage of this method relatively easily accessible. Furthermore, on average, more than 300 unique solvent structures are designed in each study, of which about one-third are successfully evaluated in the process optimization step by determining their minimum total costs. The residual candidates are discarded due to constraint violations or complex topologies, e.g., due to solvent structures that form azeotropes with the other components that cannot be separated by the process model. A detailed overview of the distribution of solvent candidates for which the minimum total costs could be successfully determined, as well as those that were discarded, is provided in Table S5.

The solvent candidates with minimal total costs of less than $50\,\mathrm{USD\,t^{-1}}$ are presented in Fig. 4 for each study conducted. They are specified for the optimal feed-to-solvent ratio, which represents the operating point of the extraction column determined by the framework. All these candidates utilize the one-column configuration for rectification. Solvent candidates for which the two-column configuration was selected are not included among the best candidates, as their costs exceed $100\,\mathrm{USD\,t^{-1}}$, a finding consistent with the previous screening study in Section 3.2.1. We use Di-n-butylether as our benchmark as it was the best performing candidate of the previously conducted screening study that included experimentally tested solvent candidates. Compared to the benchmark solvent Di-n-butylether, multiple solvent structures identified in each study achieved lower minimum total costs, demonstrating the framework's ability to reliably discover outstanding solvent candidates.

Table 3
Screening study results sorted for minimal total costs under consideration of the optimal feed-to-solvent ratio and number of rectification columns.

Candidate solvent	Total costs [USD t^{-1}]	$\dot{V}_{feed}/\dot{V}_{solvent}$	No. of rect. columns
1. Di-n-butylether	37.04	16.8	1
2. Dipentylether	43.85	19.4	1
3. n-Decane	50.74	20.6	1
4. n-Butylbenzene	52.12	26	1
5. Undecanal	69.68	22.8	1
6. Ethyl octanoate	77.59	3.1	1
7. Dodecane	89.99	8.6	1
8. Decyl acetate	112.99	2.9	2
9. Hexyl hexanoate	120.25	21.7	2
10. Octyl acetate	127.58	26.6	1
11. α-Farnesene isomer	127.86	23.5	2
12. 2-Butyl-1-octanol	141.48	13.9	2
13. Methyl decanoate	148.43	16.9	2
14. Hexadecane	150.37	7.8	2
15. 1-Undecanol	152.41	16.4	2
16. Ethyl decanoate	207.18	2.7	2

Table 4
20 best candidate solvents over all five design studies. The reduction of minimal total costs is referred to the benchmark Di-n-butyl ether.

Candidate solvent [InChi]	Total costs [USD t^{-1}]	$\dot{V}_{feed}/\dot{V}_{solvent}$	Study	Reduction
1. 1S/C8H18O/c1-7(2)5-6-9-8(3)4/h7-8H,5-6H2,1-4H3	28.2	40.5	5	23.87 %
2. 1S/C8H18O/c1-4-5-6-7-9-8(2)3/h8H,4-7H2,1-3H3	29.83	35.8	4	19.47 %
3. 1S/C9H18O2/c1-4-5-6-7-11-9(10)8(2)3/h8H,4-7H2,1-3H3	30.21	2.8	4	18.44 %
4. 1S/C8H18O/c1-4-6-7-9-8(3)5-2/h8H,4-7H2,1-3H3	30.57	30.6	2 & 3	17.47 %
5. 1S/C8H18O/c1-4-5-6-9-7-8(2)3/h8H,4-7H2,1-3H3	30.81	30.8	4	16.82 %
6. 1S/C9H18O/c1-5-8(2)6-9(3)7-10-4/h5,8-9H,1,6-7H2,2-4H3	31.16	48.7	4	15.87 %
7. 1S/C8H16O/c1-3-5-6-8-9-7-4-2/h3H,1,4-8H2,2H3	32.52	30.6	5	12.2 %
8. 1S/C9H18O/c1-4-7-10-8-6-9(3)5-2/h5,9H,2,4,6-8H2,1,3H3/t9-/m0/s1	32.58	30.6	1	12.04 %
9. 1S/C11H24O/c1-6-10(4)12-11(5)8-7-9(2)3/h9-11H,6-8H2,1-5H3	33.01	31.7	3	10.88 %
10. 1S/C8H18O/c1-5-6-8(2,3)7-9-4/h5-7H2,1-4H3	33.05	20.6	5	10.77 %
11. 1S/C8H18O2/c1-7(2)5-6-9-10-8(3)4/h7-8H,5-6H2,1-4H3	33.66	29.4	5	9.13 %
12. 1S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3	33.96	16	1	8.32 %
13. 1S/C10H20O/c1-4-6-9(3)7-10(5-2)8-11/h8-10H,4-7H2,1-3H3	34.17	17.2	5	7.75 %
14. 1S/C8H18O/c1-5-6-8(4)9-7(2)3/h7-8H,5-6H2,1-4H3	34.17	12.6	1	7.75 %
15. 1S/C10H22O/c1-5-8-11-9-10(4,6-2)7-3/h5-9H2,1-4H3	34.75	27	1	6.18 %
16. 1S/C8H18O/c1-5-8(4)6-9-7(2)3/h7-8H,5-6H2,1-4H3	35.22	14.4	4	4.91 %
17. 1S/C8H18O/c1-3-5-6-7-8-9-4-2/h3-8H2,1-2H3	35.4	18.3	3	4.43 %
18. 1S/C8H18O/c1-5-8(4)9-6-7(2)3/h7-8H,5-6H2,1-4H3	35.42	13.7	5	4.37 %
19. 1S/C10H22O/c1-5-8-11-10(4)7-6-9(2)3/h9-10H,5-8H2,1-4H3	35.54	30.2	3	4.05 %
20. 1S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3/t8-/m0/s1	35.55	16.2	2	4.02 %

Considering only solvent candidates that outperform the benchmark, the majority of all high-performing structures were distinct to each study, showing little overlap in the specific structures identified.

The exception is the candidate 1-(1-methylpropoxy)-butane (labeled InChI=1S/C8H18O/c1-4-6-7-9-8(3)5-2/h8H,4-7H2,1-3H3), which is the best candidate in studies two and three, with a cost of 30.57 USD t^{-1}. Given that the cost of the best candidate per design lies within a range of 28.20 USD t^{-1} in study five to 32.58 USD t^{-1} in study one, we conclude that the framework identifies a series of local optima that are located within a specific region of the design space, rather than a single global optimum.

The best candidate over all studies is isopentyl isopropyl ether (labeled InChI=1S/C8H18O/c1-7(2)5-6-9-8(3)4/h7-8H,5-6H2,1-4H3), with minimal total costs of 28.20 USD t^{-1}, which is a reduction of about 24 % compared to Di-n-butylether.

Out of the 20 best solvent candidates, shown in Table 4, a majority of 17 candidates show at least one ether conjunction. In addition, each solvent from this group has an aliphatic structure, often including alkyl groups of three to six carbon atoms. This is consistent with the applied fragment library derived from the screening study shown in S2, which consists mainly of aliphatic fragments. Although the database also contains several long-chained aliphatic fragments, such as tetradecane (C14) or decane (C10), the framework tends to select shorter fragments in the C3 to C6 range. Aromatics or other cyclic groups do not occur in the group of the 20 best-performing candidates, despite two cyclic structures appearing in the fragment database. The similarity of the structures indicates that the framework can narrow down the desired solvent to a specific group of structures that perform well and simultaneously show low risk to health, environment, and safety. In experimental in-situ extraction of methyl ketones, Nies et al. [9] utilize n-decane as solvent while Grütering et al. [4] determine 2-undecanone as the optimal solvent for in-situ extraction of a mixture of medium-length chained methyl ketones. Although both structures are also aliphatic, they are significantly larger than the structures identified in this study (C10 and C11). Since we take rectification into account, the framework favors solvent structures that have comparably lower boiling points and therefore smaller molecular weights to minimize the energy demand for separation. Therefore, purification processes utilizing long-chained solvents like n-decane are too expensive to compete with our best-performing candidates. Furthermore, none of the 20 best-performing candidate solvents across all five design studies are commercially available or part of the COSMObase database.

For the top 20 candidates, variables originating from the process model are shown in Table 5 to demonstrate the impact of different commonly used performance parameters for each solvent, e.g., enthalpy of vaporization (Δh_V) and distribution coefficient (K), and how they translate into the cost categories, e.g., investment and operating costs, eventually determining the solvent ranking. Within this subset of the top 20 candidates, CAPEX range from 11.74 USD t^{-1} to 17.81 USD t^{-1} and OPEX from 13.36 USD t^{-1} to 22.32 USD t^{-1}) revealing comparable cost spans across all candidates. Consequently, neither CAPEX nor OPEX alone are decisive factors in selecting the solvent; and rankings

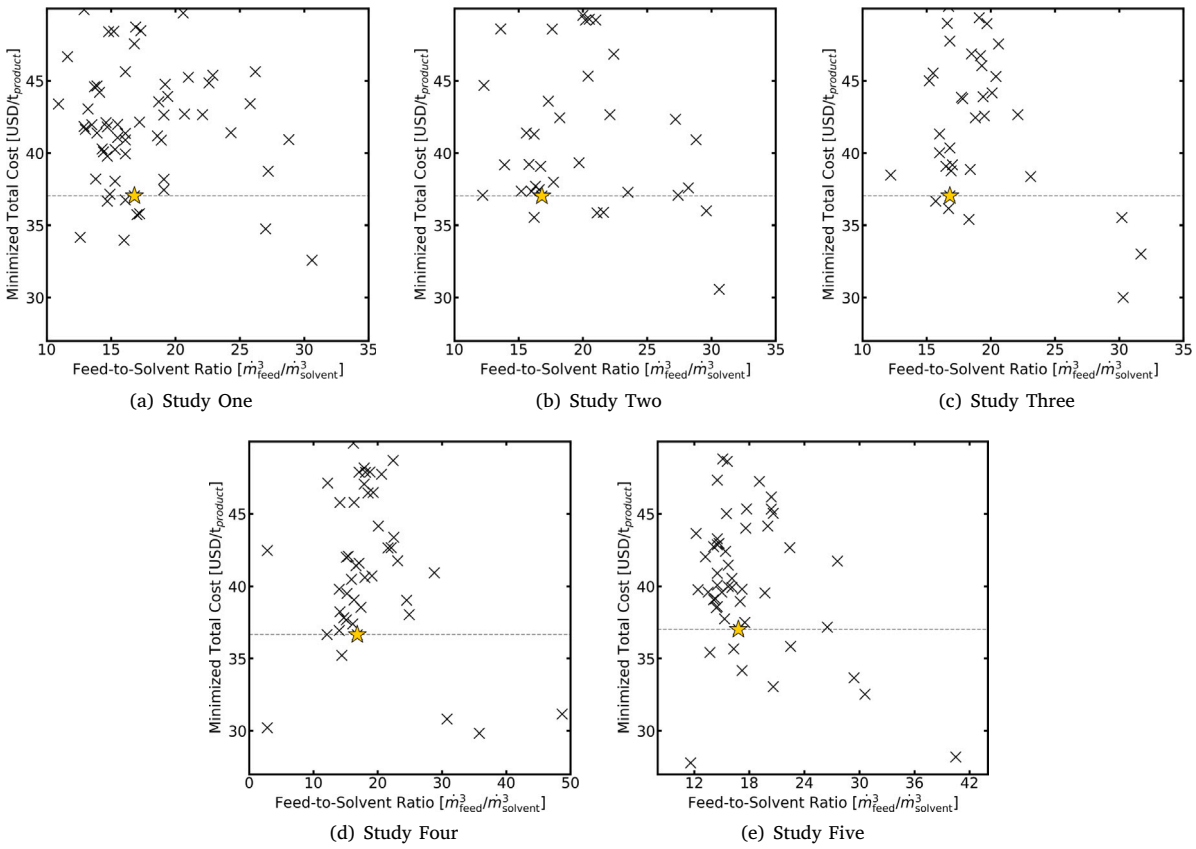


Fig. 4. Results of five conducted design studies. The figures show all candidates with total minimal process costs below 50 USD t^{-1}, marked with a cross (×). The benchmark Di-n-butylether is marked by a star (★) and a dashed line.

Table 5
Twenty best candidate solvents over all five design studies with extended performance parameter and simulation results from the process model. TC: Total costs.

Candidate solvent InChi	$\dot{V}_{solvent}/\dot{V}_{feed}$ [–]	TC [USD t^{-1}]	Rank TC [–]	OPEX [USD t^{-1}]	Rank OPEX [–]	CAPEX [USD t^{-1}]	Rank CAPEX [–]
1. 1S/C8H18O/c1-7(2)5-6-9-8(3)4/h7-8H,5-6H2,1-4H3	40.5	28.2	1	13.75	2	14.44	10
2. 1S/C8H18O/c1-4-5-6-7-9-8(2)3/h8H,4-7H2,1-3H3	35.8	29.83	2	15.1	3	14.73	11
3. 1S/C9H18O2/c1-4-5-6-7-11-9(10)8(2)3/h8H,4-7H2,1-3H3	2.8	30.21	3	18.47	11	11.74	1
4. 1S/C8H18O/c1-4-6-7-9-8(3)5-2/h8H,4-7H2,1-3H3	30.6	30.57	4	15.82	5	14.75	12
5. 1S/C8H18O/c1-4-5-6-9-7-8(2)3/h8H,4-7H2,1-3H3	30.8	30.81	5	15.75	4	15.06	13
6. 1S/C9H18O/c1-5-8(2)6-9(3)7-10-4/h5,8-9H,1,6-7H2,2-4H3	48.7	31.16	6	13.36	1	17.81	20
7. 1S/C8H16O/c1-3-5-6-8-9-7-4-2/h3H,1,4-8H2,2H3	30.6	32.52	7	16.45	7	16.07	16
8. 1S/C9H18O/c1-4-7-10-8-6-9(3)5-2/h5,9H,2,4,6-8H2,1,3H3/t9-/m0/s1	30.6	32.58	8	16.84	8	15.74	15
9. 1S/C11H24O/c1-6-10(4)12-11(5)8-7-9(2)3/h9-11H,6-8H2,1-5H3	31.7	33.01	9	18.66	12	14.35	9
10. 1S/C8H18O/c1-5-6-8(2,3)7-9-4/h5-7H2,1-4H3	20.6	33.05	10	17.94	10	15.12	14
11.1S/C8H18O2/c1-7(2)5-6-9-10-8(3)4/h7-8H,5-6H2,1-4H3	29.4	33.66	11	16.15	6	17.51	19
12. S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3	16	33.96	12	21.01	15	12.95	3
13. 1S/C10H20O/c1-4-6-9(3)7-10(5-2)8-11/h8-10H,4-7H2,1-3H3	17.2	34.17	13	20.53	14	13.64	7
14. 1S/C8H18O/c1-5-6-8(4)9-7(2)3/h7-8H,5-6H2,1-4H3	12.6	34.17	13	22.32	19	11.86	2
15. 1S/C10H22O/c1-5-8-11-9-10(4,6-2)7-3/h5-9H2,1-4H3	27	34.75	15	17.52	9	17.24	18
16. 1S/C8H18O/c1-5-8(4)6-9-7(2)3/h7-8H,5-6H2,1-4H3	14.4	35.22	16	21.7	17	13.52	5
17. 1S/C8H18O/c1-3-5-6-7-8-9-4-2/h3-8H2,1-2H3	18.3	35.4	17	21.82	18	13.58	6
18. 1S/C8H18O/c1-5-8(4)9-6-7(2)3/h7-8H,5-6H2,1-4H3	13.7	35.42	18	21.64	16	13.77	8
19. 1S/C10H22O/c1-5-8-11-10(4)7-6-9(2)3/h9-10H,5-8H2,1-4H3	30.2	35.54	19	19.47	13	16.08	17
20. 1S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3/t8-/m0/s1	16.2	35.55	20	22.4	20	13.14	4

(*continued on next page*)

based solely on either cost category are not indicative of the overall cost ranking. The enthalpies of vaporization of all top-20 solvents are low, with values well below 0.3 [kJ kg^{-1}] and a spread of approximately ±30%. This variation does not translate into a systematic trend in the overall ranking. In particular, one of the best-performing solvents exhibits one of the highest enthalpies of vaporization within the top-20 set, confirming that this property is not a primary discriminator under the given process conditions. The product, 2-undecanone, is highly nonpolar, leading to very high distribution coefficients 5. The reported values are the ratios of the product concentration in the organic and aqueous phases at the phase interface, averaged over

Table 5 (*continued*).

Candidate solvent InChi	Δh_V [MJ kg^{-1}]	K [–]	Extraction Height [m]	Extraction Diameter [m]	HETS [m]
1. 1S/C8H18O/c1-7(2)5-6-9-8(3)4/h7-8H,5-6H2,1-4H3	0.2805	1.15e+05	20.39	1.07	10.24
2. 1S/C8H18O/c1-4-5-6-7-9-8(2)3/h8H,4-7H2,1-3H3	0.2792	1.19e+05	20.99	1.06	10.54
3. 1S/C9H18O2/c1-4-5-6-7-11-9(10)8(2)3/h8H,4-7H2,1-3H3	0.2243	6.21e+03	7.9	1.65	3.97
4. 1S/C8H18O/c1-4-6-7-9-8(3)5-2/h8H,4-7H2,1-3H3	0.28	1.08e+05	21.22	1.04	10.65
5. 1S/C8H18O/c1-4-5-6-9-7-8(2)3/h8H,4-7H2,1-3H3	0.2797	1.05e+05	22.36	1.03	11.22
6. 1S/C9H18O/c1-5-8(2)6-9(3)7-10-4/h5,8-9H,1,6-7H2,2-4H3	0.2555	1.07e+05	33.94	0.99	16.98
7. 1S/C8H16O/c1-3-5-6-8-9-7-4-2/h3H,1,4-8H2,2H3	0.285	1.26e+05	25.16	1.03	12.63
8. 1S/C9H18O/c1-4-7-10-8-6-9(3)5-2/h5,9H,2,4,6-8H2,1,3H3/t9-/m0/s1	0.2555	1.21e+05	24.25	1.03	12.17
9. 1S/C11H24O/c1-6-10(4)12-11(5)8-7-9(2)3/h9-11H,6-8H2,1-5H3	0.2054	1.03e+05	20.5	1.05	10.28
10. 1S/C8H18O/c1-5-6-8(2,3)7-9-4/h5-7H2,1-4H3	0.282	1.05e+05	20.63	1.07	10.36
11.1S/C8H18O2/c1-7(2)5-6-9-10-8(3)4/h7-8H,5-6H2,1-4H3	0.2482	9.96e+04	30.41	1.01	15.25
12. S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3	0.2804	1.16e+05	10.93	1.18	5.49
13. 1S/C10H20O/c1-4-6-9(3)7-10(5-2)8-11/h8-10H,4-7H2,1-3H3	0.2291	8.03e+04	17.04	1.1	8.56
14. 1S/C8H18O/c1-5-6-8(4)9-7(2)3/h7-8H,5-6H2,1-4H3	0.2819	9.90e+04	6.26	1.34	3.14
15. 1S/C10H22O/c1-5-8-11-9-10(4,6-2)7-3/h5-9H2,1-4H3	0.2282	1.04e+05	27.43	1.08	13.75
16. 1S/C8H18O/c1-5-8(4)6-9-7(2)3/h7-8H,5-6H2,1-4H3	0.2811	1.10e+05	12.98	1.11	6.52
17. 1S/C8H18O/c1-3-5-6-7-8-9-4-2/h3-8H2,1-2H3	0.2788	1.09e+05	13.42	1.13	6.74
18. 1S/C8H18O/c1-5-8(4)9-6-7(2)3/h7-8H,5-6H2,1-4H3	0.2818	1.02e+05	13.56	1.11	6.81
19. 1S/C10H22O/c1-5-8-11-10(4)7-6-9(2)3/h9-10H,5-8H2,1-4H3	0.2246	1.09e+05	26.03	1.03	13.06
20. 1S/C8H18O/c1-4-5-6-7-8(2)9-3/h8H,4-7H2,1-3H3/t8-/m0/s1	0.2803	1.31e+05	11.51	1.19	5.78

the column height. These coefficients span two orders of magnitude, yet, this variation again shows no decisive influence on the total-cost ranking. Notably, the solvent ranked third exhibits the lowest distribution coefficient among the top-20 candidates. High distribution coefficients enable correspondingly high volume phase flow ratios. The maximum phase ratio was limited to 50, which is already close to the upper bound of practical operability, to provide the genetic algorithm with sufficient freedom to exploit advantageous thermodynamic properties during optimization, e.g., leveraging high distribution coefficients against high enthalpies of vaporization. The calculated column heights further illustrate the importance of incorporating fluid-dynamic calculations. A comparison of the solvents ranked second, third, and fifth demonstrates that the calculated column height does not directly correlate with the distribution coefficient. At the very high phase ratios leading to optimal total costs, the column diameter is predominantly determined by the fixed aqueous-phase flow. In contrast, the column height is controlled by kinetic effects, such as the residence times of both phases required to approach phase equilibrium and the interfacial mass-transfer rates. The property combinations of the solvents ranked third, twelfth, and fourteenth are particularly favorable with respect to these kinetic effects, resulting in smaller height equivalents to a theoretical stage (HETS) and, consequently, smaller extraction columns than for the other candidates.

Impact of the toxicity predictions on the downstream design

The top 20 solvent candidates identified are molecules absent from existing literature databases and posses unknown EHS properties. Consequently, incorporating early-stage EHS property evaluation significantly enhances the COSMO-CAMPED framework, as it now assesses whether novel candidates may be highly toxic or otherwise hazardous and thus do not merit further investigation. The upgrade also steers the design of new candidates towards more benign solvents for the chemical industry. However, integrating the EHS evaluation into the design framework introduces a notable limitation. It further constricts the available design space, thereby reducing the number of solvent candidates that can be assessed. Excessive constraint of the design space may result in too few successful candidates for the genetic algorithm (GA) to function effectively. In a preliminary study, we set individual thresholds for all QSAR models separately, and did not account for the uncertainty of the predictions. As a result, many potential solvent candidates where discarded based on a single QSAR prediction with a high degree of uncertainty, which caused the described problem of a too tightly constrained search space. To alleviate this problem, we introduced the overall EHS score under consideration of uncertainties we present in this study. Accumulating the EHS predictions into one score has proven to be a successful strategy in comparison to preliminary studies. When we compare the approaches of a combined EHS score and separate EHS constraints, of the 455 solvent candidates we could evaluate with the combined EHS score, 215 candidates would have failed due to exceeding a single threshold. Of these, 143 would have only exceeded one of the four thresholds marginally. This leads to the conclusion that implementing the combined EHS score avoids the problem of over constraining the design space and at the same time still designing molecules that are likely to comply with REACH regulations. To further enhance the method an automated LCA for each solvent can also be incorporated as has already been shown in Fleitmann et al. [37].

3.3. Total costs of the process

The result of the combined application of SimulKnockReactor and the enhanced COSMO-CAMPED framework is an optimized layout for a bioreactor production process of methyl ketones with a subsequent downstream purification of the fermentation broth with a bulk production capacity of 1×10^7 kg a^{-1}. The full bioprocess design is shown in Fig. 5 for the best solvent candidate isopentyl isopropyl ether (A) and the benchmark solvent Di-n-butylether (B). A detailed breakdown of all the individual cost contributions is provided in Fig. 6 for the bioprocess that utilizes isopentyl isopropyl ether as solvent. The optimal solution proposes to use four bioreactors that operate in parallel in order to produce the targeted production capacity. Each reactor has a size of 300 m^3 and houses a culture with a volume of 221 m^3. A culture of *P. taiwanensis* VLB120 Δ6 pProd with three gene knockouts (PPA, PPK, and IDPh_1) is used to produce the methyl ketones and fed with a substrate of glucose concentrated at 100 g L^{-1}. Each reactor has a volume outflow of 14.6 m^3 h^{-1} with a product concentration of 20.4 g L^{-1}. The production cost of methyl ketones before the downstream processing amount to 4.14 USD kg^{-1} of which the largest part is the substrate cost at 3.87 USD kg^{-1}. Apart from the dominating substrate costs, power (0.075 USD kg^{-1}), reactor investment (0.074 USD kg^{-1}), and cooling agent costs (0.12 USD kg^{-1}) are of the same order of magnitude and contribute similarly to the total bioreactor costs.

The purification of the methyl ketones is achieved with an extraction distillation process. COSMO-CAMPED proposes isopentyl isopropyl ether as optimal extraction solvent. The extraction column has a total volume of 18.253 m^3 and a column height of 20.39 m. The rectification is done in a single column with a volume of 1.29 m^3 and a height of 6.33 m. The total costs of the downstream process in this configuration amount to 0.0282 USD kg^{-1}, thus, being significantly lower

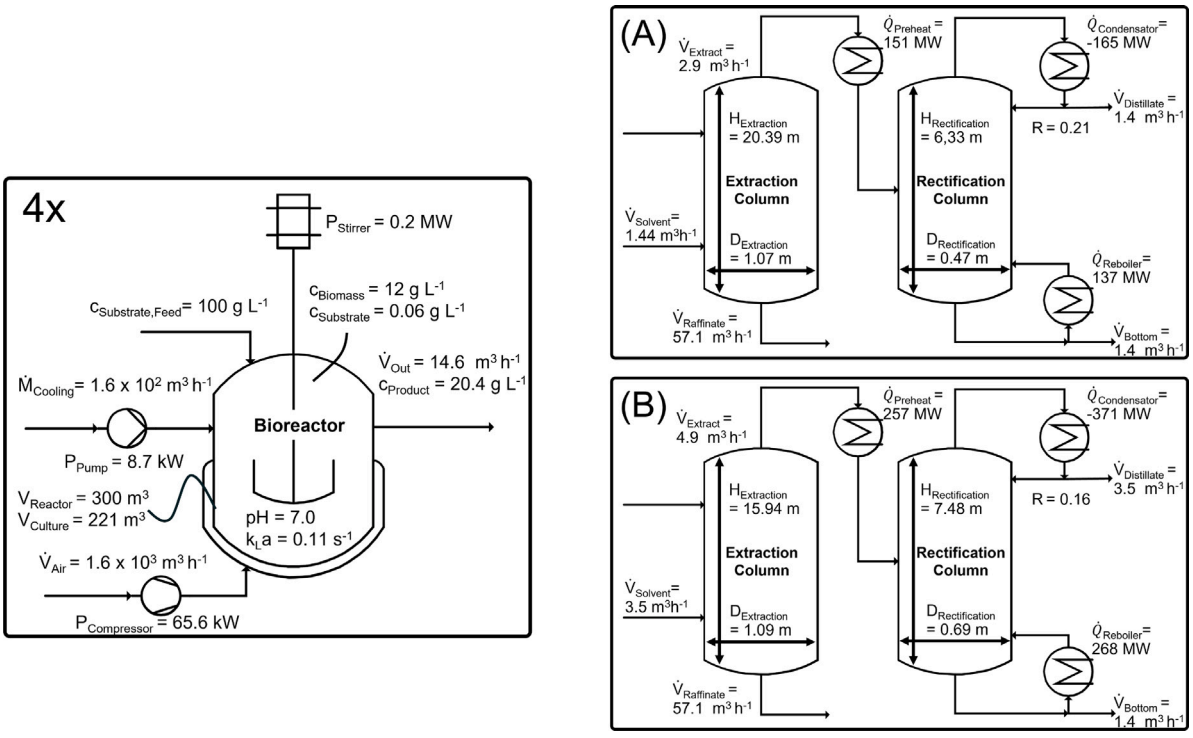


Fig. 5. The optimal bioreactor design is shown on the left. To reach a production capacity of $1\times10^7\,\mathrm{kg\,a^{-1}}$, four bioreactors of the same design operate in parallel. The right side shows two different designs of the purification section that are designed to handle the full annual production capacity. An additional filtration step prior to the purification section is assumed in which biomass is filtered out. This filtration is not modeled in the COSMO-CAMPED framework and is therefore not shown in this figure. (A) shows the downstream process design that utilizes isopentyl isopropyl ether as solvent. (B) shows the downstream process design that applies the benchmark solvent Di-n-butylether.

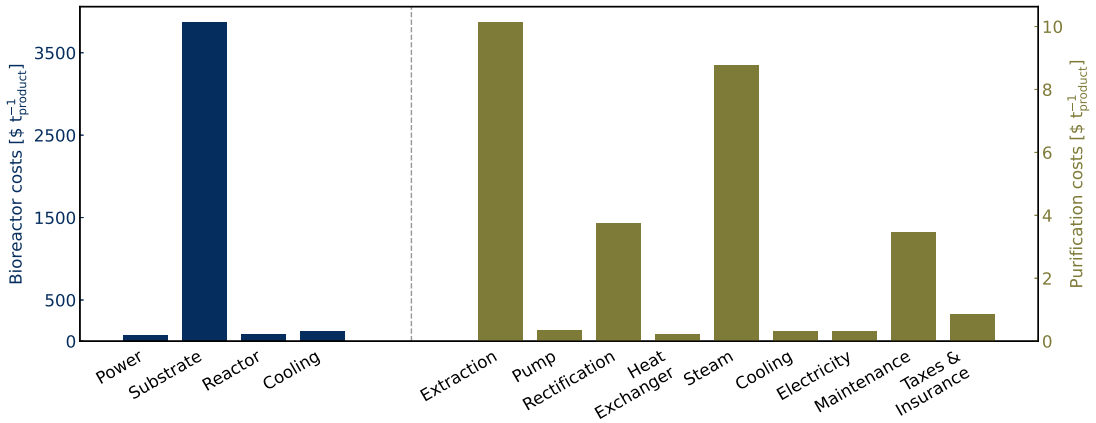


Fig. 6. Cost breakdown of the full bioprocess using the best performing solvent candidate, isopentyl isopropyl ether. The blue bars indicate the cost contributions of the bioreactor, whereas the green bars represent those of the purification process. The cost allocation for the pump reflects the investment costs for the plunger pump in the extraction column. The cost contribution of the heat exchanger describes the investment costs for preheater, condenser, and reboiler.

than the cost of the bioreactor. For the downstream process, OPEX (at $0.013\,75\,\mathrm{USD\,kg^{-1}}$) and CAPEX (at $0.014\,44\,\mathrm{USD\,kg^{-1}}$) contribute evenly to the total cost. In summary, we propose an optimized layout for a methyl ketone production in bulk. This design includes optimization of the organism and extraction solvent alongside the equipment design under economic consideration and environmental constraints. We achieve a production capacity of $1\times10^7\,\mathrm{kg\,a^{-1}}$ at a combined cost of $4.167\,\mathrm{USD\,kg^{-1}}$.

These combined production costs of around $4.2\,\mathrm{USD\,kg^{-1}}$ are subject to data uncertainty [75]. This means that the costs depend on process parameter values such as the glucose price, electricity price, solvent recovery ratio, and production scale, but also microbial parameter values such as the Monod constant, maximum growth rate, or maximum substrate uptake rate. Among the process values, especially the material prices (i.e., glucose) and the electricity price appear to have a high impact, since material costs and utility costs are the main cost drivers, upstream as well as downstream (see Fig. 6). Moreover, prices of raw materials and electricity may fluctuate significantly. For example, Panofen et al. [76] accounted for a deviation of ±30 % on utility and supply costs in their uncertainty quantification of a fuel production process. Among the microbial parameters, high impact is attributed to maximum substrate uptake rate. This parameter denotes the upper bound on the substrate uptake flux. Through the optimization, the substrate uptake flux always reaches its upper bound. Since substrate uptake influences the number of carbon atoms available for product

and biomass formation, this parameter can highly affect the final result.

The combined price of around $4.2\,\mathrm{USD\,kg^{-1}}$ is significantly lower than an available price of $100\,\mathrm{USD\,kg^{-1}}$ [77] we found in the literature. However, it is important to consider that the reference price is for a sample and we calculate the cost for bulk production. Therefore a direct comparison of the two prices is not necessarily accurate. Nevertheless, this raises some speculative points. Possible explanations for the large discrepancy in the prices may be that we did not consider cost of labor in our results, which, at least for the purification, would be a significant cost factor. Furthermore we only consider a single methyl ketone but the microorganism actually produces a mixture of medium-chained methyl ketones. To separate the individual methyl ketones in the product additional purification steps would be necessary which will in turn increase the cost.

4. Conclusion

We successfully performed sequential optimization of a methyl ketone production process from fermentation to product purification. Using our SimulKnockReactor [25], we proposed a strain and bioreactor design suitable for large-scale production. Using the extended features of the COSMO-CAMPED [42] framework, we achieved purification of methyl ketones in an extraction-distillation process using an environmentally-benign solvent candidate. The total production costs for fermentation and purification amount to $4.2\,\mathrm{USD\,kg^{-1}}$ for an annual methyl ketone production capacity of $1 \times 10^7\,\mathrm{kg\,a^{-1}}$. At this scale, the cost of the substrate, which accounts for 93 % of the total costs, is the largest cost factor. At the achievable production cost, the usage of methyl ketones as a biofuel is questionable, but it would still be valid for other applications, such as solvents for paints, coatings, or adhesives. Although fermentation costs dominate the overall process cost for optimized methyl ketone production, this may not be representative for other applications of the combined framework. Furthermore, the framework addresses dimensioning, strain design, and solvent design, all of which are critical for the configuration of any biotechnological process. For solvent and downstream process optimization, we applied the COSMO-CAMPED framework. The results revealed that process and equipment design and optimization, considering operating and investment costs, enable a comprehensive evaluation of the performance of solvent candidates. In the downstream process, both OPEX and CAPEX contribute evenly to the total cost, thus justifying consideration of investment cost as a necessary criterion. In addition, we showed that considering solvent properties beyond the established key performance indicators, such as phase equilibrium and heat of vaporization, also influences performance, underscoring the importance of extending the process model. However, purification costs remained minor compared to fermentation, the best performing downstream configurations of all 5 design runs were in a range from $0.028\,\mathrm{USD\,kg^{-1}}$ to $0.032\,\mathrm{USD\,kg^{-1}}$.

Predicting solvent bio-compatibility was implemented to enable raffinate recycle into the fermentation process, thereby limiting solvent loss and simultaneously recovering substrate and biomass to enhance fermentation. Solvent loss is further mitigated by high-purity product separation, which reduces losses in the distillation sequence; under these assumptions, the current process model can already quantify absolute solvent loss, though it does not yet capture all practically relevant loss mechanisms or the solvent cost effects of newly designed molecules. Validating these underlying assumptions and extending the framework accordingly is therefore required to ensure a consistent and comprehensive process description.

- Explicit incorporation of raffinate recycling, including substrate and biomass recovery, into the fermentation mass balances is required. This integration is essential for accurately representing the coupling between extraction and fermentation and for assessing the impact of the coupling on process performance.
- The proposed solvent needs to be benign to the bio-organism. Bio-compatibility evaluation, especially the estimation based on the solvent's Log(P), can provide a useful preliminary indicator, though experimental validation remains essential.
- The cost of the solvent should be considered in the overall economics. Predictive cost models exist in principle, but suitable methods for newly designed bulk chemicals still need to be evaluated and a suitable approach selected and implemented in future work.

In conclusion, the proposed process concept requires further validation to confirm its suitability for full raffinate recycling. Including a model for solvent recovery efficiency and solvent loss would also benefit the model in a more holistic consideration of tradeoffs. Nevertheless, the presented extensions to the COSMO-CAMPED and SimulKnockReactor frameworks provide a robust foundation for early-stage design and techno-economic evaluation of methyl ketone production processes.

To overcome limitations, we propose future development of the proposed method, that merit investigation but are beyond the scope of this paper.

Experimental validation of the predicted gene knockouts and extraction behavior is a logical next step. This can be pursued at an early stage of process design without the requirement for a pilot plant. Note that we conducted a COSMO-RS calculation for the solubility of the product in water. The calculated solubility of 2-undecanone in water is below the bioreactor maximum concentration. However, due to uncertainties of COSMO-RS, the miscibility prediction is likely not reliable. The miscibility of 2-undecanone and water should therefore be experimentally verified.

Considering the large-scale fermentation modeled with SimulKnockReactor, the methodology would benefit from some refinements. Namely, the air pressure should be adapted or at least refined depending on reactor height, further increasing power requirements with increasing reactor size. Furthermore, economies of scale should be considered for raw material costs and power. The reliability of the results would benefit from experimental verification of pH, as the pH results did not appear consistent. In addition, Monod kinetics are presumably not well suited to describe the growth behavior of *P. taiwanensis* due to the high substrate affinity of the strain. Currently, however, Monod is the only kinetics implemented in SimulKnockReactor; it is desirable to implement other options, e.g., Michaelis–Menten kinetics. Finally, co-feeding is an attractive option in order to decrease fermentation costs and could be considered in further studies. For example, co-feeding of ethanol has improved methyl ketone yield [10].

The COSMO-CAMPED method is already well-suited to model large-scale production. Nevertheless, due to the fully predictive nature of the model, it is prone to error propagation since the process model evaluation relies on predicted property values. Therefore, errors in property prediction methods propagate into the process model, e.g. a deviation of the predicted boiling point affects the necessary amount of rectification steps which in turn affects OPEX and CAPEX. We mitigated this by enforcing molecular constraints and excluding molecules with boiling points outside the COSMO-RS core range, within which property predictions are reasonably reliable. However, the confidence in the results could be further improved by including uncertainty statements for each predicted value to classify the results and estimate the error range, similar to as we have done with the EHS properties. Apart from that, not all of the designed molecules are commercially available and sometimes may not even be synthesizable. The field of synthesis route prediction and especially synthesis price prediction is rapidly evolving. To our knowledge methods are currently only available for fine chemicals [78], and yield inaccurate results for bulk chemical production. Thus, integrating tools to estimate the synthesizability and, in a subsequent step, the stability of the molecules during the process would enhance the methodology in the future. For the design of

separation sequences involving more complex mixtures, such as multi-component fermentation products, automated process synthesis tools could be incorporated [66].

Lastly, uncertainty quantification and sensitivity analysis are of interest. Therefore, structural, model, and data uncertainty could be identified and quantified [75,79]. Similar to Panofen et al. [76], a sampling method could be used to generate different scenarios. Since both the upstream and the downstream optimization steps are computationally intensive, advanced sampling methods or Bayesian optimization could be of interest.

Abbreviations

Abbreviation	Meaning
BCF	Bioconcentration factor
CAMD	Computer-Aided Molecular Design
CAMPD	Computer-Aided Molecular and Process Design
CAMPED	Computer-Aided Molecular, Process, and Equipment Design
CAPEX	Capital expenditures
COSMO	Conductor like Screening Model
COSMO-RS	Conductor like Screening Model for Real Solvents
EHS	Environmental, Health, and Safety
FBA	Flux balance analysis
FOB	Free-on-board
FUG	Fenske-Underwood-Gilliland
GA	Genetic Algorithm
HETS	Height equivalents to a theoretical stage
LC	Lethal concentration
LCA	Life Cycle Assessment
LD	Lethal dose
LEA	Ligand by Evolutionary Algorithm
MED	Minimum energy demand
MINLP	Mixed-Integer Nonlinear program
NRTL	non-random two-liquid
OPEX	Operating expense
QSAR	Quantitative Structure-Activity Relationship
REACH	Registration, Evaluation, Authorization and Restriction of Chemicals
RBM	Rectification body method
SMILES	Simplified Molecular Input Line Entry System
TC	Total costs
USD	US-Dollar
VLE	Vapor–liquid equilibrium

Greek symbols

σ	Interfacial tension	$N\,m^{-1}$

Latin symbols

c	concentration	$g\,L^{-1}$
D	diameter	m
h	enthalpy	$kJ\,kg^{-1}$
K	distribution coefficient	–
P	electrical power	MW
p	pressure	Pa
$\dot{Q}$	heat flow	MW
υ_{bio}	growth rate	h^{-1}
V	volume	m^3
$\dot{V}$	volume flow	$m^3\,h^{-1}$
x	mole fraction	–

Subscripts

A	atom
boil	boiling point
crit	critical point
F	flash point
min	minimum
v	vaporization

CRediT authorship contribution statement

Henry Hilker: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Lukas Raßpe-Lange:** Writing – original draft, Validation, Software, Methodology, Investigation, Conceptualization. **Lukas Polte:** Writing – original draft, Validation, Software, Methodology, Investigation, Conceptualization. **Anita L. Ziegler:** Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Clemens Kortmann:** Writing – review & editing, Validation, Software, Investigation. **Alexander Mitsos:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Andreas Jupke:** Writing – review & editing, Supervision, Resources, Funding acquisition. **Kai Leonhard:** Writing – review & editing, Supervision, Resources, Funding acquisition.

ALZ, LP, and LRL jointly developed the overall concept for the manuscript.

ALZ conceptualized the upstream investigations. HH, LP, and LRL conceptualized the downstream investigations. HH and LRL conceptualized the integration of QSAR models and the warm-start strategy in the COSMO-CAMPED framework. LP conceptualized the extension of the rectification model in the COSMO-CAMPED framework. ALZ and CK performed the adaptations of the SimulKnockReactor framework for methyl ketone production and performed the upstream case studies; ALZ with a focus on modeling, and CK with focus on implementation. HH performed the adaptations of the COSMO-CAMPED framework under the supervision of LP and LRL. HH and LRL performed the downstream case studies.

Data and software availability

The source code of the SimulKnockReactor framework is openly available under https://gitlab.git.nrw/rwth-avt-svt/public/simulknock.

The source code for the COSMO-CAMPED Framework including the screening workflow, the molecular optimization routine property and property predictions is available under https://git-ce.rwth-aachen.de/ltt/cosmo-camped.

The process model is not included in this git repository, for access to the process model please inquire with the corresponding author.

Several external software packages are necessary to use this code the programming language Perl 5.0. The programming language Perl 5.0 is open-source. All COSMO-related software in this work is proprietary commercial software distributed by COSMOlogic GmbH, Leverkusen, Germany, now part of Dassault Systèmes SE., and requires a separate license. However an open source distribution of COSMO-RS is available here: https://www.tuhh.de/v8/de/softwares/opencosmo-rs. The MATLAB software is also proprietary commercial software distributed by MathWorks, and requires a separate license. The QSPR models for the toxicological analysis are open source and can be found under https://github.com/kmansouri/OPERA and https://github.com/kmansouri/OPERA.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the authors used grammarly to correct grammar and spelling mistakes and improve wording. The

authors also used the DeepL software for translation. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article. No generative AI was used to write sections or paragraphs of this manuscript; nor did the authors generative AI to create the figures.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

L. Raßpe-Lange, A. L. Ziegler, A. Mitsos, and K. Leonhard gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – Cluster of Excellence 2186, The Fuel Science Center–ID: 390919832. A. L. Ziegler and A. Mitsos acknowledge support of the Werner Siemens Foundation in the frame of the WSS Research Center "catalaix". Simulations were performed with computing resources granted by RWTH Aachen University under project rwth0836 and thes1577. Finally, we thank Carolin Grütering, Tobias Alter, and Leon Poduschnick for a fruitful discussion on methyl ketone fermentation conditions.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cej.2026.180680.

References

[1] M. Gavrilescu, Y. Chisti, Biotechnology - A sustainable alternative for chemical industry, Biotech. Adv. 23 (2005) 471–499.
[2] G. Lopez, D. Keiner, M. Fasihi, T. Koiranen, C. Breyer, From fossil to green chemicals: sustainable pathways and new carbon feedstocks for the global chemical industry, Energy Environ. Sci. 16 (7) (2023) 2879–2909.
[3] O. Marie, A.V. Ignatchenko, M. Renz, Methyl ketones from carboxylic acids as valuable target molecules in the biorefinery, Catal. Today 367 (2021) 258–267.
[4] C. Grütering, et al., Methyl ketones: a comprehensive study of a novel biofuel, Sustain. Energy Fuels 8 (2024) 2059–2072.
[5] E.-B. Goh, E.E. Baidoo, J.D. Keasling, H.R. Beller, Engineering of bacterial methyl ketone synthesis for biofuels, Appl. Environ. Microbiol. 78 (1) (2012) 70–80.
[6] E.K. Hanko, C.M. Denby, V.S. i Nogué, W. Lin, K.J. Ramirez, C.A. Singer, G.T. Beckham, J.D. Keasling, Engineering β-oxidation in yarrowia lipolytica for methyl ketone production, Metab. Eng. 48 (2018) 52–62.
[7] Z. Zhu, Y.J. Zhou, A. Krivoruchko, M. Grininger, Z.K. Zhao, J. Nielsen, Expanding the product portfolio of fungal type I fatty acid synthases, Nat. Chem. Biol. 13 (4) (2017) 360–362.
[8] J. Dong, Y. Chen, V.T. Benites, E.E. Baidoo, C.J. Petzold, H.R. Beller, A. Eudes, H.V. Scheller, P.D. Adams, A. Mukhopadhyay, et al., Methyl ketone production by pseudomonas putida is enhanced by plant-derived amino acids, Biotechnol. Bioeng. 116 (8) (2019) 1909–1922.
[9] S.C. Nies, T.B. Alter, L.M. Blank, B.E. Ebert, High titer methyl ketone production with tailored pseudomonas taiwanensis VLB120, Chem. Ing. Tech. 92 (9) (2020) 1183–1184.
[10] A.L. Ziegler, C. Grütering, L. Poduschnick, A. Mitsos, L.M. Blank, Co-feeding enhances the yield of methyl ketones, J. Ind. Microbiol. Biotechnol. 50 (2023).
[11] A. Ait El Fakir, P. Du, B. Yang, N.H.M. Dostagir, J.W. Fischer, A. Anzai, K.-i. Shimizu, T. Toyao, A review on catalytic ethanol synthesis via hydrogenation of carbon dioxide, ChemSusChem 18 (14) (2025) e202500188.
[12] A. Knoll, B. Maier, H. Tscherrig, J. Büchs, The oxygen mass transfer, carbon dioxide inhibition, heat removal, and the energy and cost efficiencies of high pressure fermentation, in: U. Kragl (Ed.), in: Technology Transfer in Biotechnology, vol. 92, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 77–99.
[13] V.M. Cardoso, G. Campani, M.P. Santos, G.G. Silva, M.C. Pires, V.M. Gonçalves, R. de C Giordano, C.R. Sargo, A.C.L. Horta, T.C. Zangirolami, Cost analysis based on bioreactor cultivation conditions: Production of a soluble recombinant protein using escherichia coli BL21(DE3), Biotechnol. Rep. (Amsterdam Netherlands) 26 (2020) e00441.
[14] J.A. Gómez, Ó.J. Sánchez, L.F. Correa, Techno-economic and environmental evaluation of cheesemaking waste valorization through process simulation using SuperPro designer, Waste Biomass Valorization 11 (11) (2020) 6025–6045.
[15] J. Villadsen, J. Nielsen, G. Lidén, Bioreaction Engineering Principles, third ed., Springer Science+Business Media LLC, Boston, MA, 2011.
[16] L. Olsson, P. Rugbjerg, L. Torello Pianale, C. Trivellin, Robustness: linking strain design to viable bioprocesses, Trends Biotechnol. 40 (8) (2022) 918–931.
[17] A. Richelle, B. David, D. Demaegd, M. Dewerchin, R. Kinet, A. Morreale, R. Portela, Q. Zune, M. von Stosch, Towards a widespread adoption of metabolic modeling tools in biopharmaceutical industry: a process systems biology engineering perspective, NPJ Syst. Biol. Appl. 6 (1) (2020) 6.
[18] R. Mahadevan, J.S. Edwards, F.J. Doyle, Dynamic flux balance analysis of diauxic growth in escherichia coli, Biophys. J. 83 (3) (2002) 1331–1340.
[19] J.D. Orth, I. Thiele, B.Ø. Palsson, What is flux balance analysis? Nature Biotechnol. 28 (3) (2010) 245–248.
[20] J.L. Hjersted, M.A. Henson, Optimization of fed-batch saccharomyces cerevisiae fermentation using dynamic flux balance models, Biotechnol. Prog. 22 (5) (2006) 1239–1248.
[21] X. Zhao, S. Noack, W. Wiechert, E. von Lieres, Dynamic flux balance analysis with nonlinear objective function, J. Math. Biol. 75 (6–7) (2017) 1487–1515.
[22] L. Chang, X. Liu, M.A. Henson, Nonlinear model predictive control of fed-batch fermentations using dynamic flux balance models, J. Process Control 42 (2016) 137–149.
[23] T. Ploch, J. Deussen, U. Naumann, A. Mitsos, R. Hannemann-Tamás, Direct single shooting for dynamic optimization of differential-algebraic equation systems with optimization criteria embedded, Comput. Chem. Eng. 159 (2022) 107643.
[24] R.D. de Oliveira, M.N. Guedes, J. Matias, G.A.C. Le Roux, Nonlinear predictive control of a bioreactor by surrogate model approximation of flux balance analysis, Ind. Eng. Chem. Res. 60 (40) (2021) 14464–14475.
[25] A.L. Ziegler, M.-D. Stumm, T. Prömper, T. Steimann, J. Magnus, A. Mitsos, Simultaneous design of microbe and bioreactor, Comput. Chem. Eng. 204 (2026) 109388.
[26] A.P. Burgard, P. Pharkya, C.D. Maranas, Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization, Biotechnol. Bioeng. 84 (6) (2003) 647–657.
[27] B. Scharzec, K.F. Kruber, M. Skiborowski, Model-based evaluation of a membrane-assisted hybrid extraction-distillation process for energy and cost-efficient purification of diluted aqueous streams, Chem. Eng. Sci. 240 (2021) 116650.
[28] J. Kampwerth, B. Weber, J. Rußkamp, S. Kaminski, A. Jupke, Towards a holistic solvent screening: On the importance of fluid dynamics in a rate-based extraction model, Chem. Eng. Sci. 227 (2020).
[29] C. Gertig, K. Leonhard, A. Bardow, Computer-aided molecular and processes design based on quantum chemistry: current status and future prospects, Curr. Opin. Chem. Eng. 27 (2020) 89–97.
[30] C.S. Adjiman, A. Galindo, Challenges and opportunities for computer-aided molecular and process design approaches in advancing sustainable pharmaceutical manufacturing, Curr. Opin. Chem. Eng. 47 (2025).
[31] M. Folić, R. Gani, C. Jiménez-González, D.J. Constable, Systematic selection of green solvents for organic reacting systems, Chin. J. Chem. Eng. 16 (3) (2008) 376–383.
[32] J. Scheffczyk, C. Redepenning, C.M. Jens, B. Winter, K. Leonhard, W. Marquardt, A. Bardow, Massive, automated solvent screening for minimum energy demand in hybrid extraction–distillation using COSMO-RS, Chem. Eng. Res. Des. 115 (2016) 433–442.
[33] L. König-Mattern, A.O. Komarova, A. Ghosh, S. Linke, L.K. Rihko-Struckmann, J. Luterbacher, K. Sundmacher, High-throughput computational solvent screening for lignocellulosic biomass processing, Chem. Eng. J. 452 (2023).
[34] S. Linke, K. McBride, K. Sundmacher, Systematic green solvent selection for the hydroformylation of long-chain alkenes, ACS Sustain. Chem. Eng. 8 (2020) 10795–10811.
[35] S. Sethi, X. Zhang, K. Sundmacher, Process-driven solvent screening for efficient extractive distillation using interpolative rational functions, Chem. Eng. Sci. 301 (2025).
[36] C. Gertig, L. Fleitmann, J. Schilling, K. Leonhard, A. Bardow, Rx-COSMO-CAMPD: Enhancing reactions by integrated computer-aided design of solvents and processes based on quantum chemistry, Chem. Ing. Tech. 92 (2020) 1489–1500.
[37] L. Fleitmann, J. Kleinekorte, K. Leonhard, A. Bardow, COSMO-susCAMPD: Sustainable solvents from combining computer-aided molecular and process design with predictive life cycle assessment, Chem. Eng. Sci. 245 (2021).
[38] J. Schilling, M. Lampe, J. Gross, A. Bardow, 1-stage CoMT-CAMD: An approach for integrated design of ORC process and working fluid using PC-SAFT, Chem. Eng. Sci. 159 (2017) 217–230.
[39] L. Neumaier, D. Roskosch, J. Schilling, G. Bauer, J. Gross, A. Bardow, Refrigerant selection for heat pumps: The compressor makes the difference, Energy Technol. 11 (2023).
[40] M. Mersch, D. Tillmanns, P. Sapin, J. Schilling, A. Bardow, C.N. Markides, Integrated thermo-economic organic rankine cycle and working fluid design – on the accuracy of molecular-based computer-aided methodologies, Comput. Chem. Eng. 199 (2025).
[41] A. Bardow, K. Steur, J. Gross, Continuous-molecular targeting for integrated solvent and process design, Ind. Eng. Chem. Res. 49 (2010) 2834–2840.

[42] L. Polte, L. Raßpe-Lange, F. Latz, A. Jupke, K. Leonhard, COSMO-CAMPED – solvent design for an extraction distillation considering molecular, process, equipment, and economic optimization, Chem. Ing. Tech. 95 (2023) 416–426.
[43] C. Grütering, Producing Methyl Ketones in Integrated Bioprocesses (Ph.D. thesis), RWTH Aachen University, 2024.
[44] N.-C. Choi, J.-W. Choi, S.-B. Kim, D.-J. Kim, Modeling of growth kinetics for pseudomonas putida during toluene degradation, Appl. Microbiol. Biotechnol. 81 (1) (2008) 135–141.
[45] J. Eberz, L. Strehl, M. Mann, A. Jupke, J.B. Magnus, Phase separation behaviour during direct solvent extraction of corynebacterium glutamicum fermentation broth–systematic study of crud suppression, Curr. Res. Green Sustain. Chem. 10 (2025) 100448.
[46] A. Klamt, V. Jonas, T. Bürger, J.C. Lohrenz, Refinement and parametrization of COSMO-RS, J. Phys. Chem. A 102 (26) (1998) 5074–5085.
[47] Y. Wang, L. Fleitmann, L. Raßpe-Lange, N. von der Assen, A. Bardow, K. Leonhard, Fine-tuning a genetic algorithm for CAMD: A screening-guided warm start, J. Chem. Inf. Model. (2025).
[48] D. Douguet, H. Munier-Lehmann, G. Labesse, S. Pochet, LEA3D: A computer-aided ligand design for structure-based drug design, J. Med. Chem. 48 (2005) 2457–2468.
[49] European Parliament, Council of the European Union, Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), European Union, 2006, URL https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32006R1907. Consolidated version (see EUR-Lex for latest consolidated date). Official Journal of the European Union, L 396, 1–849.
[50] J. Rowley, D. Freeman, R. Rowley, J. Oscarson, N. Giles, W. Wilding, Flash point: evaluation, experimentation and estimation, Int. J. Thermophys. 31 (4) (2010) 875–887.
[51] X. Chen, A. Qadeer, M. Liu, L. Deng, P. Zhou, I.T. Mwizerwa, S. Liu, Z. Ajmal, Z. Xingru, X. Jiang, Bioaccumulation of emerging contaminants in aquatic biota: PFAS as a case study, in: Emerging Aquatic Contaminants, Elsevier, 2023, pp. 347–374.
[52] J. Reinisch, A. Klamt, Predicting flash points of pure compounds and mixtures with COSMO-RS, Ind. Eng. Chem. Res. 54 (51) (2015) 12974–12980.
[53] C. Loschen, J. Reinisch, A. Klamt, COSMO-RS based predictions for the SAMPL6 logP challenge, J. Comput. Aided Mol. Des. 34 (4) (2020) 385–392.
[54] E. Benfenati, A. Manganaro, G.C. Gini, et al., VEGA-QSAR: AI inside a platform for predictive toxicology, Pai@ Ai* IA 1107 (2013) 21–28.
[55] K. Mansouri, C.M. Grulke, R.S. Judson, A.J. Williams, OPERA models for predicting physicochemical properties and environmental fate endpoints, J. Cheminform. 10 (1) (2018) 10.
[56] K. Mansouri, A.L. Karmaus, J. Fitzpatrick, G. Patlewicz, P. Pradeep, D. Alberga, N. Alepee, T.E. Allen, D. Allen, V.M. Alves, et al., CATMoS: collaborative acute toxicity modeling suite, Environ. Health Perspect. 129 (4) (2021) 047013.
[57] K. Mansouri, J.T. Moreira-Filho, C.N. Lowe, N. Charest, T. Martin, V. Tkachenko, R. Judson, M. Conway, N.C. Kleinstreuer, A.J. Williams, Free and open-source QSAR-ready workflow for automated standardization of chemical structures in support of QSAR modeling, J. Cheminform. 16 (1) (2024) 19.
[58] C. Toma, C.I. Cappelli, A. Manganaro, A. Lombardo, J. Arning, E. Benfenati, New models to predict the acute and chronic toxicities of representative species of the main trophic levels of aquatic environments, Molecules 26 (22) (2021) 6983.
[59] C. Zhao, E. Boriani, A. Chana, A. Roncaglioni, E. Benfenati, A new hybrid system of QSAR models for predicting bioconcentration factors (BCF), Chemosphere 73 (11) (2008) 1701–1707.
[60] A. Gissi, A. Lombardo, A. Roncaglioni, D. Gadaleta, G.F. Mangiatordi, O. Nicolotti, E. Benfenati, Evaluation and comparison of benchmark QSAR models to predict a relevant REACH endpoint: The bioconcentration factor (BCF), Environ. Res. 137 (2015) 398–409.
[61] F. Pizzo, A. Lombardo, A. Manganaro, C.I. Cappelli, M.I. Petoumenou, F. Albanese, A. Roncaglioni, M. Brandt, E. Benfenati, Integrated in silico strategy for PBT assessment and prioritization under REACH, Environ. Res. 151 (2016) 478–492.
[62] M. Andersson, M. Bennetzen, A. Klamt, S.S. Stipp, First-principles prediction of liquid/liquid interfacial tension, J. Chem. Theory Comput. 10 (8) (2014) 3401–3408.
[63] J. Bausa, R.v. Watzdorf, W. Marquardt, Shortcut methods for nonideal multicomponent distillation: I. Simple columns, AIChE J. 44 (10) (1998) 2181–2198.
[64] C.D. Holland, Fundamentals of multicomponent distillation, in: McGraw-Hill Chemical Engineering Series, McGraw-Hill, New York, 1981, p. 626.
[65] AVT.SVT, RWTH Aachen University, EE-toolbox: Energy-efficiency toolbox for separation processes, 2017, https://www.avt.rwth-aachen.de/cms/AVT/Forschung/Software/~iptv/Shortcut-Toolbox/?lidx=1. (Accessed 18 April 2025).
[66] T. Sasi, M. Skiborowski, Automatic synthesis of distillation processes for the separation of homogeneous azeotropic multicomponent systems, Ind. Eng. Chem. Res. 59 (2020) 20816–20835.
[67] R. Agrawal, Synthesis of distillation column configurations for a multicomponent separation, Ind. Eng. Chem. Res. 35 (4) (1996) 1059–1071.
[68] R.T. Gooty, R. Agrawal, M. Tawarmalani, An MINLP formulation for the optimization of multicomponent distillation configurations, Comput. Chem. Eng. 125 (2019) 13–30.
[69] M. Fenske, Fractionation of straight-run Pennsylvania gasoline, Ind. Eng. Chem. 24 (5) (1932) 482–485.
[70] E. Gilliland, Multicomponent rectification, Ind. Eng. Chem. 32 (8) (1940) 1101–1106.
[71] W.D. Seider, D.R. Lewin, J.D. Seader, S. Widagdo, R. Gani, K.M. Ng, Product and Process Design Principles: Synthesis, Analysis and Design, fourth ed., Wiley, 2017.
[72] H.-P. Meyer, W. Minas, D. Schmidhalter, Industrial-scale fermentation, in: C. Wittmann, J.C. Liao (Eds.), Industrial Biotechnology, in: Advanced Biotechnology, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2016.
[73] P.M. Doran, Reactor engineering, in: P.M. Doran (Ed.), Bioprocess Engineering Principles, Academic Press, London, 1995, pp. 333–391.
[74] P.F. Stanbury, A. Whitaker, S.J. Hall, Principles of Fermentation Technology, third ed., Butterworth-Heinemann an imprint of Elsevier, Amsterdam, 2017.
[75] E. Hofer, The Uncertainty Analysis of Model Results, Springer International Publishing, Cham, 2018.
[76] M. Panofen, P. Ackermann, J. Viell, A. Mitsos, M. Dahmen, Uncertainty quantification in integrated fuel and process design, Energy Fuels 38 (15) (2024) 14743–14756.
[77] Sigma-Aldrich, 2-undecanone, 2025, URL https://www.sigmaaldrich.com/US/en/product/aldrich/w309303?srsltid=AfmBOooMyFyySSeaDcvoXLKv2OAwVB9FKMTawOTN8EmKhFwETKeS9ZKz. (Last Accessed 03 December 2025).
[78] F. Hastedt, K. Hellgardt, S. Yaliraki, A. del Rio Chanona, D. Zhang, Computational assessment of molecular synthetic accessibility using economic indicators, Syst. Control. Trans. (2025) 1744–1749.
[79] P.F. Pelz, P. Groche, M.E. Pfetsch, M. Schaeffner, Mastering Uncertainty in Mechanical Engineering, Springer International Publishing, Cham, 2021.