Application of Life Cycle Assessment and Machine Learning for High-Throughput Screening of Green Chemical Substitutes

Xinzhe Zhu [†],Chi-Hung Ho [†], Xiaonan Wang*

Department of Chemical and Biomolecular Engineering, National University of Singapore,

4 Engineering Drive 4, 117585 Singapore

† These authors contribute equally to this work.

* Corresponding author:

Tel: +65 6601 6221

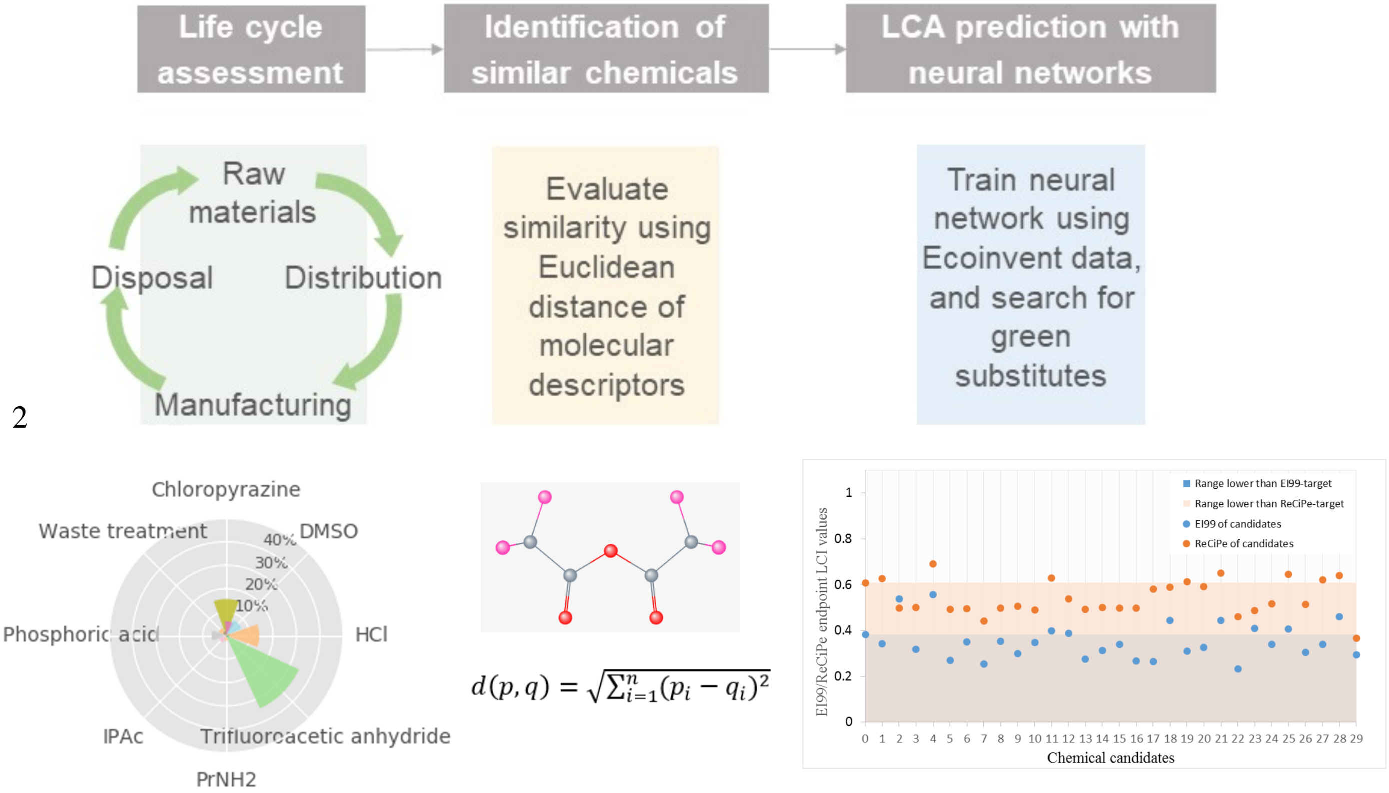
Email: chewxia@nus.edu.sg

ABSTRACT

The production process of many active pharmaceutical ingredients such as sitagliptin could cause severe environmental problems due to the use of toxic chemical materials and production infrastructure, energy consumption and wastes treatment. The environmental impacts of sitagliptin production process were estimated with life cycle assessment (LCA) method, which suggested that the use of chemical materials provided the major environmental impacts. Both methods of Eco-indicator 99 and ReCiPe endpoints confirmed that chemical feedstock accounted 83% and 70% of life-cycle impact, respectively. Among all the chemical materials used in the sitagliptin production process, trifluoroacetic anhydride was identified as the largest influential factor in most impact categories according to the results of ReCiPe midpoints method. Therefore, high-throughput screening was performed to seek for green chemical substitutes to replace the target chemical (i.e. trifluoroacetic anhydride) by the following three steps. Firstly, thirty most similar chemicals were obtained from two million candidate alternatives in PubChem database based on their molecular descriptors. Thereafter, deep learning neural network models were developed to predict life-cycle impact according to the chemicals in Ecoinvent v3.5 database with known LCA values and corresponding molecular descriptors. Finally, 1,2-ethanediyl ester was proved to be one of the potential greener substitutes after the LCA data of these similar chemicals were predicted using the well-trained machine learning models. The case study demonstrated the applicability of the novel framework to screen green chemical substitutes and optimize the pharmaceutical manufacturing process.

Keywords: Machine learning; Life cycle assessment; Green chemistry; High-throughput screening; Pharmaceutical manufacturing process

Graphical Abstract


Life cycle assessment
Identification of similar chemicals
LCA prediction with neural networks
Raw materials
Distribution
Manufacturing
Disposal
Evaluate similarity using Euclidean distance of molecular descriptors
Train neural network using Ecoinvent data, and search for green substitutes
Chloropyrazine
Waste treatment
DMSO
Phosphoric acid
HCl
IPAc
Trifluoroacetic anhydride
PrNH2
40%
30%
20%
10%
$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$
EI99/ReCiPe endpoint LCI values
Chemical candidates
0
0.2
0.4
0.6
0.8
1
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29
Range lower than EI99-target
Range lower than ReCiPe-target
EI99 of candidates
ReCiPe of candidates

SYNOPSIS

Neural network models were trained using molecular descriptors and life-cycle impact of known chemicals, and used to search greener substitutes in a huge library of chemicals.

1. INTRODUCTION

Recently the concept of green chemistry has been widely acknowledged and applied in the field of chemical industry due to the rising environmental concerns.[1] The principle theory of green chemistry is to minimize human health and environmental risks, for example, using greener chemical substitutes in the manufacturing process or optimizing the production process to reduce energy consumption.[2-4] Pharmaceutical manufacturing has been found to bear more severe environmental impacts than basic chemicals production because of the complex molecular structures of chemical feedstocks, intricate synthesis and separation reactions, and high-standard purifications in the production process.[5] In order to quantify and evaluate the anthropogenic environmental impacts in pharmaceutical process, life cycle assessment (LCA) has received much attention by virtue of its broad applicable scope and outstanding holism.[5, 6] Meanwhile, the implementation of LCA also significantly supports the development of greener concepts according to the relationships between the manufacturing process and resulting environmental impacts.[6,7]

LCA is a methodological framework that has been developed for several decades, from simple energy analysis, environmental burden analysis to present “compilation and evaluation of the inputs, outputs, and potential environmental impacts of a product system throughout its life cycle”.[8,9] The general procedures of LCA include the definition of goal and scope, the life cycle inventory (LCI) analysis for the whole system, the life cycle impact assessment (LCIA) calculation, and the interpretation for impact assessments results.[9-11] The goal and scope definition of LCA is applied to describe the boundaries (e.g. “cradle to gate” and “cradle to grave”) and functional unit of production system. LCI as the foundation of LCA is used to summarize the total resources consumption, waste flows and emissions, while LCIA is performed to quantify the potential environmental consequences by multiplying LCI with corresponding impact indicators.[8] Life cycle interpretation always happens at every stage in LCA to better serve decision makers.[6]

Although LCA calculations of pharmaceutical processes have many difficulties such as information privacy in many pharmaceutical companies and their complex life cycle inventory, relevant researches have been recently performed from simple case studies to decision-making support by comparing the LCA results in different types of chemistry and technologies.[12-14] For example, the synthesis process of an active pharmaceutical ingredient (API) was analyzed from cradle to factory gate, which suggested that the resources consumption and emissions accounted for the major contributions to the environmental impacts.[5] Furthermore, LCA was applied as a decision-support tool on another real case from pharmaceutical industry to demonstrate the importance of continuous flow reactors on the reduction of overall resource consumption.[15] Around the same time, the continuous pharmaceutical supply chain in Janssen-Cilag SpA was also proved with more greenness and environmental sustainability than conventional batch manufacturing mode by 10.2% through LCA calculations.[16] Similarly, Ott et al. reported the holistic LCA results for different rufinamide production pathways and proposed the green chemistry optimization schemes such as solvent recycling or reagent replacement to decrease environmental risks, which proved the role of LCA in the development of green chemistry.[17] However, the case-by-case comparison of pharmaceutical manufacturing systems was a complex and time-consuming task.[18] Moreover, LCIA data for pharmaceutical processes were always rarely available due to the use of fine chemicals with complex molecular structures, which increased the challenges to carry out the LCA calculation of pharmaceuticals production and optimize chemicals production process.[14, 19] Recently a similarity-based link prediction approach has been developed to predict LCIA data in a given chemical process according to the similarity theory, that is, similar processes tend to share similar inputs (e.g. materials, energy, etc.) and output such as wastes.[20] Nevertheless, the framework was not suitable for the complex pharmaceuticals manufacturing process because of numerous unknown LCIA data.

In order to fill the gap of missing data for LCI and find green chemical substitutes with lower environmental impacts, correlation models between manufacturing process and resulting environmental impacts have been attempted. For example, Wernet et al. demonstrated the dependences of several environmental impacts categories such as Cumulative Energy Demand (CED), Global Warming Potential (GWP), and Eco-indicator 99 score on the molecular structure of chemicals.[21] But the prediction abilities of simple regression models were limited, the emerging machine learning method was also tried to gain insights into the complex relationships in recent years. The model performance of artificial neural network (ANN) has been proved superior to linear regression in predicting LCI data based on the molecular structures of chemicals.[22] Afterwards, Song et al. improved the performances of deep learning ANN model by increasing the model complexity and expanding the training data size.[23] These previous researches provided us good references, but so far there is still a lack of research to apply the machine learning method in searching for green chemical substitutes based on the LCA prediction and building the overall framework for the purpose of achieving green chemistry.

As an important active pharmaceutical ingredient leading antidiabetic drug, sitagliptin production line and their LCA calculation are focused to raise a framework to green the pharmaceutical manufacturing process with the aid of high-throughput screening from chemicals libraries and machine learning prediction. The accurate prediction for life-cycle impact data will furtherly increase the application of LCA in the optimization of pharmaceutical industry, while the framework developed in this case will provide a guidance for finding green chemical substitutes.

2. METHODOLOGY

2.1 Life cycle assessment for sitagliptin production

The flowsheet of continuous sitagliptin manufacturing and all chemicals materials used in the process could be found in our previous work,[24] which were also shown in the supplementary material (**Fig. S1 and Table S1**). The system boundary of LCA calculation for the process was set to be "cradle to gate", which meant that it traced back to any ingredient used in upstream to synthesize the building blocks and ended with sitagliptin production.[13] The environmental impacts of chemical materials and infrastructures, the demand for process energy and the wastes treatment were calculated, respectively. It should be noted that the environmental impact of enzyme production was not included due to the limitation of data availability and the main objective being the search for greener chemical substitutes. The global databases updated in Ecoinvent v3.5 were used, which provided a more holistic point of view.[25] The functional unit (FU) was defined as producing 1 kg sitagliptin monophosphate to study the corresponding environmental impacts. The methods of Eco-indicator 99 (EI99), ReCiPe endpoints and midpoints shown in **Table S2** were used to calculate the LCA in the process of sitagliptin manufacturing based on the hierarchism perspective.[26] Ecosystem quality, human health, and resources depletion were considered when using EI99 and ReCiPe endpoints methods.[27] While ten impact categories were chosen based on ReCiPe midpoints method, including Global Warming Potential (GWP), Fossil Fuel Depletion Potential (FDP), Freshwater Eutrophication Potential (FEP), Human Toxicity Potential (HTP), Metal Depletion Potential (MDP), Natural Land Transformation Potential (NLTP), Ozone Depletion Potential (ODP), Photochemical Oxidant Formation Potential (POFP), Terrestrial Acidification Potential (TAP), and Terrestrial Eco-toxicity Potential (TETP).

However, missing data of LCI, especially in pharmaceutical process was a common issue for LCA calculation. In order to complete data gaps, the following methods were performed. Firstly, for missing data of LCI in chemical materials, the retrosynthetic breakdown method was applied by summarizing the LCI data of reagents to the target chemical materials.[17] Otherwise, the data could be substituted with that of structurally similar substances or generic data.[17] Secondly, the energy consumption of a single process was hard to split from overall manufacture energy records. The approximation method proposed by Kim and Overcash was applied, in which the "gate-to-gate" process energy consumption was estimated according to 4 MJ kg^{-1} in manufacturing of organic chemicals.[28] Thirdly, the generic data "chemical factory construction, organics, Rest of the World (RoW)" in Ecoinvent database was used to model the LCI of all infrastructures in this study, while the data of "treatment of spent solvent mixture, hazardous waste incineration, RoW" was applied to model the LCI of wastes treatment.

2.2 Identification of similar chemicals

In order to reduce the environmental impact of chemical materials used in the pharmaceutical process, we tried to find the substituted green chemicals with lower life-cycle impact based on their similar molecular structures and compositions. The similarity between target and candidate chemicals was quantified according to their molecular characteristics, and the flowsheet for identification of similar chemicals was shown in **Fig. S2**. High-throughput methods were applied to find the similar chemicals, in which two million chemicals were collected from PubChem database (https://pubchem.ncbi.nlm.nih.gov/). A total of 125 molecular characteristics descriptors of these chemicals were generated by python package Rdkit. The data of each molecular descriptor was standardized to the same scale with the following equation (1):

$$x' = \frac{x-\mu}{\sigma} \qquad (1)$$

where x' and x represented the standardized data and original data, μ and σ were the mean and standard deviation of all data with respect to a given descriptor, respectively.

Thereafter, principal component analysis (PCA) was carried out so as to improve computational efficiency by reducing the dimensions of molecular descriptors and simultaneously ensure little information loss, in which orthogonal transformation was used to convert a set of correlated variables into a set of linearly uncorrelated variables.[29] After PCA, the obtained n principal components could represent the majority of information in molecular descriptors. The similarity identification was calculated based on the Euclidean distance[30] between each of two million candidate chemicals and the target chemical with following equation (2):

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \quad (2)$$

where p_i and q_i represented the molecular descriptors of target and candidate chemicals, respectively.

2.3 LCA prediction with deep neural networks

After the similar chemicals were identified, their corresponding LCA data would be predicted using well-trained deep neural network models based on their structural information. The predictive models were built based on available 224 nonionic organic chemicals with known LCA in Ecoinvent v3.5 database. Their corresponding molecular descriptors as the model inputs were obtained from Rdkit (https://www.rdkit.org/) and AlvaDesc1.0 (https://chm.kode-solutions.net/products_alvadesc.php), respectively. But what needs to be reminded is that names of chemicals could not be directly recognized by the above two systems, thus the Simplified Molecular Input Line Entry System (SMILES) structure[31] should be firstly obtained through ChemSpider (http://www.chemspider.com/). Meanwhile, the LCA data of ecosystem, human health, resources and the total impact obtained by EI99 and ReCipe endpoint methods were applied as outputs, respectively. The data of molecular descriptors from Rdkit and AlvaDesc were also pre-processed with PCA approach, and the LCI data with extremely high values were excluded from the dataset as outliers according to their boxplot distribution.

Artificial neural network was used to build and train machine learning models, in which the model architecture was composed by an input layer, multiple hidden layers, and an output layer [32]. The collected data was randomly divided into three parts, including training group (70%), validation group (20%) and test group (10%). The data in training group was used to build models with ReLU activation functions and train the models through adjusting the weights of connections between neurons in different layers with back-propagation algorithm.[33,34] Thereafter, the validation dataset was introduced into the fitted model to perform an unbiased evaluation and meanwhile adjust the accuracy of models by tuning the hyper-parameters such as the number of hidden layers and the number of neurons in each hidden layer.[32] The data in test group was used to evaluate the final model as external validation. All the ANN models were developed with Tensorflow framework [35] in Python.

After the best model for each impact category was identified, the molecular descriptors of these similar chemicals were introduced into the models to predict their corresponding LCI values. The chemicals with higher similarities and lower environmental impacts could be considered as potential greener substitutes. The whole framework of the proposed high-throughput screening for green chemical substitutes based on LCA and ANN was shown in **Fig. 1**.

3. RESULTS AND DISCUSSION

3.1 Life cycle assessment in the sitagliptin manufacturing process

We used the manufacturing process of sitagliptin as a case study to demonstrate the developed methodology. The production of sitagliptin started with chloropyrazine through nine main steps, in which the holistic LCA results with EI99 and ReCiPe endpoints methods were shown in **Fig. 2.** Both of the two LCA calculation methods **(Figs. 2A and 2B)** suggested that the sitagliptin production process had largest impact on human health (i.e. 53% and 44% for EI99 and ReCiPe endpoint, respectively), followed by the impacts on resources depletion (i.e. 40% and 36% for EI99 and ReCiPe endpoint, respectively) and ecosystem quality (i.e. 7% and 20% for EI99 and ReCiPe endpoint, respectively). The influences of drug production on human health may be related to the consequent global warming and respiratory tract effect of chemical feedstocks [5]. From another perspective **(Figs. 2C and 2D)**, the use of chemical feedstock was the major contributor to the total environmental impacts with 83% and 70% in EI99 and ReCiPe endpoint, respectively (**Fig. 2**). The large proportion of environmental impacts caused by chemical feedstock was consistence with the LCA results of previous pharmaceutical synthesis process.[5] In detail, chemical materials provided 86%, 76% and 92% of the damage to ecosystem quality, human health and resources availability with EI99 method, respectively (**Fig. 2C**). The influences of process energy and waste treatments accounted for similar proportions as the second most important factors for all the impact categories, while the impacts of infrastructure were negligible due to their long-term use.[36] From the LCA results with ReCiPe endpoints method, we could also obtain similar conclusion that the influences of chemical materials were most significant for ecosystem (55%), human health (62%) and resources (87%) in the sitagliptin manufacturing process **(Fig. 2D)** in spite of some deviations with the absolute values from EI99 method.[37]

More detailed impact categories were also applied to identify the relative importance of chemical materials, infrastructure, process energy and waste treatment with ReCiPe midpoint method (**Fig. 3**). The use of chemical materials still accounted for the majorities of environmental impacts, while the infrastructure also had the lowest proportion in each of the ten categories [19]. The overall environmental impact caused by this process is as follows: 547.76 kg of CO2 equivalents per FU for GWP; 155.16 kg of oil equivalents per FU for FDP; 0.03 kg of P equivalents per FU for FEP; 55.26 kg of 1,4-dichlorobenzene per FU for HTP; 13.88 kg of Fe equivalents per FU for MDP; -0.02 m2 per FU for NLTP; 0.0002 kg of chlorofluorocarbon-11 per FU for ODP; 1.31 kg of non-methane volatile organic compounds per FU for POFP; 1.78 kg of SO2 equivalents per FU for TAP; 0.51 kg of 1,4-DCB per FU for TETP. Since chemical feedstock has been proved as the most significant factor for environmental impacts, the next step would be in-depth analysis to recognize the relative influence of each chemical material used in the sitagliptin manufacturing process and try to explore greener substitute. The life-cycle impacts of twenty-one chemicals in the sitagliptin production were calculated with ReCiPe midpoint method[38] and the relative proportion of each chemical at the midpoint level was investigated shown with radar chart in **Fig. 4**. Taking the radar chart of TETP as an example, hydrazine provided more than 50% of terrestrial eco-toxicity in the 21 chemicals because of the toxicity and its interactions with environmental medium.[39] Making a general survey of the ten impact categories, it could be found that trifluoroacetic anhydride showed the highest impact in all the impact categories other than GWP, ODP and TETP. Moreover, the environmental impact of trifluoroacetic anhydride came in second in GWP and ODP impact categories (**Fig. 3**). Therefore, it is necessary to search for greener substitutes for trifluoroacetic anhydride to reduce the environmental impact.

3.2 Identification of similar chemicals for trifluoroacetic anhydride

The quantified molecular descriptors of two million chemicals were preprocessed with PCA to reduce the dimensionality and improve the calculation efficiency. The first component explained almost 25% of variance in the original data as shown in the cumulative explained variance plot (**Fig. S3)**. In order to minimize information loss,[40] sixty-six principal components were finally selected by covering almost the entire data variability (99%).

The Euclidean distances were subsequently calculated between each collected chemical from PubChem databases and target chemical (i.e. trifluoroacetic anhydride) with the corresponding PCA-molecular descriptors. The similarity indices of chemicals were ranged from 0 (total similarity) to infinity (complete dissimilarity) and 30 chemicals that were most similar to the target chemical were shown in **Table S3** according to the Euclidean distances.[41] It could be found that the Euclidean distance between chemical 0 and the target was zero, because they are the same substances and the SMILES structure of trifluoroacetic anhydride is O=C(OC(=O)C(F)(F)F)C(F)(F)F. The excellent performance of Euclidean distance index has been confirmed as the useful similarity index.[42] The results suggested that the majority of the most similar 30 chemicals were comprised of carbon, oxygen and fluorine, which were the same with the atoms composition of the target chemical. However, the LCA values of the thirty chemicals were unknown, so the next step would be to train prediction models according to the chemicals with known LCA values and corresponding molecular descriptors.

3.3 LCA prediction models with deep learning neural networks

In order to improve the accuracy of machine learning models, the data used for building LCA prediction models were collected from nonionic organic chemicals, consistent with the target chemical such as petrochemicals, pharmaceuticals and industrial chemicals in Ecoinvent v3.5 database. The distributions of LCA values with eight impact categories (i.e. ecosystem quality, human health, resources and the above total LCA values obtained with EI99 and ReCiPe endpoints methods, respectively) were shown with boxplot in **Fig. S4**. The lines from bottom to top in the boxplot presented the minimum, the first quartile (Q1), the median, the third quartile (Q3), and the maximum of these statistical data.[32] Both of the two methods showed that the influences of these organic chemicals on ecosystem were lower than that on human health and resources (**Fig. S4)**, which was accordance with the chemical materials in our sitagliptin production process **(Fig. 2)**. However, the LCI values of chemicals with ReCiPe endpoint method were generally larger than that with EI99.[43] Especially for the ecosystem quality impact category, the median value of these chemicals obtained with ReCiPe endpoints was higher than that with EI99 by four folds (**Fig. S4)**. Although ReCiPe endpoint was developed based on the classical LCA method (EI99), there were still inherent differences such as their endpoints characterization factors.[37] The damage to ecosystem quality was calculated according to the potentially disappeared fraction of species in terrestrial ecosystem for EI99 method, while both of terrestrial and aquatic (including freshwater and marine water) damages were considered in ReCiPe endpoint.[44] The rhombus shape in the boxplot represented the outliers of statistical data, and the points far away from the normal values would be excluded in the next section.

The molecular descriptors generated by Rdkit (125 descriptors) and AlvaDesc (3,874 descriptors) were extracted by PCA, respectively, in which PCA-based a models had been proved to have the best performances in the previous LCA prediction models.[23] The PCA-molecular descriptors were thereafter used as the inputs of ANN models. However, the performances of molecular descriptors from Rdkit were very poor and not reported here, which may be related to the less input information from original molecular descriptors. It has been proved in our previous study that prediction ability improved when more relevant information was introduced into the developed models.[45] The performances of the best ANN model for each of the eight impact categories based on the molecular descriptors from AlvaDesc were shown in **Fig. 5**. The regression coefficient (R^2) and root mean square error (RMSE) of test group were applied to quantify the prediction ability of developed models. The results suggested that the chemicals with larger LCA values tended to have higher prediction errors due to less training data within the same range.[23] Meanwhile, both of total EI99 (R^2= 0.8356)and total ReCiPe endpoints (R^2= 0.883) showed the highest R^2 for the comprehensive evaluation compared to the corresponding individual prediction for ecosystem, human health or resources. As a whole, the total ReCiPe model was slightly better than total EI99 model, which may be because the ReCiPe method, as the successor of EI99, could reflect the environmental impacts more objectively. Among the eight models, the prediction models for ecosystem in both EI99 and ReCiPe endpoints methods showed lower performances, with R^2 values of 0.6454 and 0.6328 on the test group, respectively. A rational explanation would be that it is difficult to monitor the damage of these chemicals to ecosystem due to the heterogeneous and complex characteristics.[44]

3.4 Prediction of LCA characterized results for similar chemicals

According to the best prediction models based on total EI99 and ReCiPe endpoints methods, the LCIA data of these 30 most similar chemicals with the target (same with chemical 0) were evaluated (**Fig. 6**). The chemicals were considered as the greener substitute candidates when they had lower LCIA values obtained with both EI99 and ReCiPe endpoints than that of the target molecules. Ultimately, 17 chemicals were found to have lower environmental impacts than trifluoroacetic anhydride as shown in **Fig. 7**. In terms of EI99 method, chemical 22 (i.e. Methyl pentafluoropropionylacetate) had the lowest LCIA values and meanwhile the LCIA predictive values based on ReCiPe endpoints method were lower than that of target chemical. Likewise, chemical 29 (i.e. 1,2-ethanediyl ester) had the lowest LCA predictive values obtained from ReCiPe endpoints and was also lower than the target based on EI99 method. Finally, chemical 29 was assessed as the suitable substitute candidate for the two functional groups (i.e. C=O, CF3), because the two chemical groups –COCF3 of trifluoroacetic anhydride provided an important role in the sitagliptin production.[24] Therefore, considering structural similarity, desired functional groups, and lower environmental impacts, it is worthwhile to experimentally assess the feasibility of substituting chemical 29 for trifluoroacetic anhydride in the sitagliptin manufacturing process.

3.5 Outlook and improvement in the future

A feasible direction has been demonstrated to predict the life-cycle impact data of fine chemicals using the information of molecular structures without a priori knowledge of the production process. After all, molecular descriptors could reflect the physicochemical properties (e.g. solubility, molar refractivity, topological polar surface area, etc.) and molecular fingerprint (e.g. atom type, aromaticity, functional groups, the number of attached hydrogen atoms, connectivity, etc.). Nevertheless, the available LCIA data in Ecoinvent database was insufficient to furtherly improve the performances of machine learning models. Although a large number of other LCA databases have been developed such as ELCD database and GaBi Database, the differences of standards and methods among these databases limited the expanding the data size used to build LCA prediction models.[46] The integration of these separate databases in the future with united criteria may be a potential solution to better realize the LCA prediction.

Furthermore, the molecular structures of chemicals could not represent the information of overall production, especially for fine pharmaceutical production process. For example, stringent standard in separation and purification would require higher energy consumption and chemical feedstocks. Therefore, if we could consider the whole production process with the increase of related data in the future, including the molecular structures and process parameters, a generalizable machine learning model will be more meaningful for simplifying the application of LCA in the pharmaceutical manufacturing field.[18]

4. Conclusions

Taking the sitagliptin production as an example, the overall framework of greening the pharmaceutical manufacturing process was proposed in this article based on the holistic LCA calculation and emerging machine learning methods. Both of EI99 and ReCiPe endpoints LCA results suggested that the use of chemical feedstocks provided the major contribution to the total environmental impacts, while trifluoroacetic anhydride accounted for the majority of chemical materials in most impact categories according to the results of ReCiPe midpoints method. In order to reduce environmental footprint caused by the sitagliptin production, searching for greener and similar chemicals with the target chemical (i.e. trifluoroacetic anhydride) was subsequently performed. Thirty most similar chemicals to the target were firstly selected as candidate substitutes from PubChem database containing two million chemicals on the basis of Euclidean distance calculations. Meanwhile, machine learning models were built to explore the relationship between LCIA values of chemicals and their corresponding molecular structures. Herein, 224 nonionic organic chemicals with known LCI from Ecoinvent v3.5 database were used to build, train and test the predictive models with deep learning ANN algorithm. Thereafter, the molecular descriptors of the 30 similar chemicals were introduced into the well-trained ML models to calculate their LCIA values. Finally, the chemical 1,2-ethanediyl ester was reported as the potential greener substitute that is worth being experimentally validated according to the lower LCI data and similar molecular compositions and function groups. The overall screening framework provided a reference to search greener substitute in pharmaceutical process from large libraries of chemicals, which could decrease the experimental burden and costs.

Abbreviations

API- Active Pharmaceutical Ingredients

LCA- Life Cycle Assessment

LCI- Life Cycle Inventory

LCIA- Life Cycle Impact Assessment

ANN- Artificial Neural Network

FU- Functional Unit

EI99- Eco-indicator 99

CED- Cumulative Energy Demand

GWP- Global Warming Potential

FDP- Fossil Fuel Depletion Potential

FEP- Freshwater Eutrophication Potential

HTP- Human Toxicity Potential

MDP- Metal Depletion Potential

NLTP- Natural Land Transformation Potential

ODP- Ozone Depletion Potential

POFP- Photochemical Oxidant Formation Potential

TAP- Terrestrial Acidification Potential

TETP- Terrestrial Eco-toxicity Potential

SMILES- Simplified Molecular Input Line Entry System

R^2- Regression Coefficient

RMSE- Root Mean Square Error

ACKNOWLEDGMENT

The authors thank the MOE AcRF Grant in Singapore for financial support to the project (R-279-000-513-133 and R-279-000-541-114). The authors are grateful to Prof. S. A. Khan and Prof. Z. Li for the helpful discussions of the pharmaceutical manufacturing.

REFERENCES

(1) Poliakoff, M.; Licence, P. Sustainable technology: Green chemistry. *Nature* **2007,** *450* (7171), 810-812.

(2) Poliakoff, M.; Fitzpatrick, J. M.; Farren, T. R.; Anastas, P. T. Green chemistry: science and politics of change. *Science* **2002,** *297* (5582), 807-810.

(3) Phillips, K. A.; Wambaugh, J. F.; Grulke, C. M.; Dionisio, K. L.; Isaacs, K. K. High-throughput screening of chemicals as functional substitutes using structure-based classification models. *Green Chem.* **2017,** *19* (4), 1063-1074.

(4) Alfonsi, K.; Colberg, J.; Dunn, P. J.; Fevig, T.; Jennings, S.; Johnson, T. A.; Kleine, H. P.; Knight, C.; Nagy, M. A.; Perry, D. A. Green chemistry tools to influence a medicinal chemistry and research chemistry based organisation. *Green Chem.* **2008,** *10* (1), 31-36.

(5) Wernet, G.; Conradt, S.; Isenring, H. P. D.; Jimenezgonzalez, C.; Hungerbuhler, K. Life cycle assessment of fine chemical production: a case study of pharmaceutical synthesis. *Int. J. Life Cycle Ass.* **2010,** *15* (3), 294-303.

(6) Kralisch, D.; Ott, D.; Gericke, D. Rules and benefits of life cycle assessment in green chemical process and synthesis design: a tutorial review. *Green Chem.* **2015,** *17* (1), 123-145.

(7) Dunn, P. J. The importance of Green Chemistry in Process Research and Development. *Chem. Soc. Rev.* **2012,** *41* (4), 1452-1461.

(8) Guinee, J. B.; Heijungs, R.; Huppes, G.; Zamagni, A.; Masoni, P.; Buonamici, R.; Ekvall, T.; Rydberg, T. Life Cycle Assessment: Past, Present, and Future. *Environ. Sci. Technol.* **2011,** *45* (1), 90-96.

(9) Hellweg, S.; Canals, L. M. I. Emerging approaches, challenges and opportunities in life cycle assessment. *Science* **2014,** *344* (6188), 1109-1113.

(10) Pennington, D.; Potting, J.; Finnveden, G.; Lindeijer, E.; Jolliet, O.; Rydberg, T.; Rebitzer, G. Life cycle assessment Part 2: Current impact assessment practice. *Environ. Int.* **2004,** *30* (5), 721-739.

(11) Rebitzer, G.; Ekvall, T.; Frischknecht, R.; Hunkeler, D.; Norris, G. A.; Rydberg, T.; Schmidt, W. P.; Suh, S.; Weidema, B. P.; Pennington, D. Life cycle assessment part 1: framework, goal and scope definition, inventory analysis, and applications. *Environ. Int.* **2004,** *30* (5), 701-720.

(12) Diab, S.; Gerogiorgis, D. I. Process modeling, simulation, and technoeconomic evaluation of separation solvents for the continuous pharmaceutical manufacturing (CPM) of diphenhydramine. *Org. Process Res. Dev.* **2017,** *21* (7), 924-946.

(13) Mata, T. M.; Martins, A. A.; Neto, B.; Martins, M. L.; Romualdo, S.; Costa, C. LCA tool for sustainability evaluations in the pharmaceutical industry. *Chem. Eng. Trans.* **2012,** *26*.

(14) Jimenezgonzalez, C.; Overcash, M. The evolution of life cycle assessment in pharmaceutical and chemical applications-a perspective. *Green Chem.* **2014,** *16* (7), 3392-3400.

(15) Der Vorst, G. V.; Aelterman, W.; De Witte, B.; Heirman, B.; Van Langenhove, H.; Dewulf, J. Reduced resource consumption through three generations of Galantamine·HBr synthesis. *Green Chem.* **2013,** *15* (3), 744-748.

(16) De Soete, W.; Dewulf, J.; Cappuyns, P.; Der Vorst, G. V.; Heirman, B.; Aelterman, W.; Schoeters, K.; Van Langenhove, H. Exergetic sustainability assessment of batch versus continuous wet granulation based pharmaceutical tablet manufacturing: a cohesive analysis at three different levels. *Green Chem.* **2013,** *15* (11), 3039-3048.

(17) Ott, D.; Borukhova, S. S.; Hessel, V. Life cycle assessment of multi-step rufinamide synthesis – from isolated reactions in batch to continuous microreactor networks. *Green Chem.* **2016,** *18* (4), 1096-1116.

(18) De Soete, W.; Debaveye, S.; De Meester, S.; Der Vorst, G. V.; Aelterman, W.; Heirman, B.; Cappuyns, P.; Dewulf, J. Environmental Sustainability Assessments of Pharmaceuticals: An Emerging Need for Simplification in Life Cycle Assessments. *Environ. Sci. Technol.* **2014,** *48* (20), 12247-12255.

(19) Ott, D.; Kralisch, D.; Dencic, I.; Hessel, V.; Laribi, Y.; Perrichon, P.; Berguerand, C.; Kiwiminsker, L.; Loeb, P. Life cycle analysis within pharmaceutical process optimization and intensification: case study of active pharmaceutical ingredient production. *Chemsuschem* **2014,** *7* (12), 3521-3533.

(20) Hou, P.; Cai, J.; Qu, S.; Xu, M. Estimating Missing Unit Process Data in Life Cycle Assessment Using a Similarity-Based Approach. *Environ. Sci. Technol.* **2018,** *52* (9), 5259-5267.

(21) Wernet, G.; Papadokonstantakis, S.; Hellweg, S.; Hungerbuhler, K. Bridging data gaps in environmental assessments: Modeling impacts of fine and basic chemical production. *Green Chem.* **2009,** *11* (11), 1826-1831.

(22) Wernet, G.; Hellweg, S.; Fischer, U.; Papadokonstantakis, S.; Hungerbühler, K. Molecular-structure-based models of chemical inventories using neural networks. *Environ. Sci. Technol.* **2008,** *42* (17), 6717-6722.

(23) Song, R.; Keller, A. A.; Suh, S. Rapid life-cycle impact screening using artificial neural networks. *Environ. Sci. Technol.* **2017,** *51* (18), 10777-10785.

(24) Ho, C.; Yi, J.; Wang, X. Biocatalytic continuous manufacturing of diabetes drug: plantwide process modeling, optimization, environmental and economic analysis. *ACS Sustain. Chem. Eng.* **2018**, 7(1): 1038-1051.

(25) Wernet, G.; Bauer, C.; Steubing, B.; Reinhard, J.; Moreno-Ruiz, E.; Weidema, B. The ecoinvent database version 3 (part I): overview and methodology. *Int. J. Life Cycle Ass.* **2016,** *21* (9), 1218-1230.

(26) Dong, Y. H.; Ng, S. T. Comparing the midpoint and endpoint approaches based on ReCiPe—a study of commercial buildings in Hong Kong. *Int. J. Life Cycle Ass.* **2014,** *19* (7), 1409-1423.

(27) Boulay, A.M.; Bulle, C.; Bayart, J. B.; Deschênes, L.; Margni, M. Regional characterization of freshwater use in LCA: modeling direct impacts on human health. *Environ. Sci. Technol.* **2011,** *45* (20), 8948-8957.

(28) Kim, S.; Overcash, M. Energy in chemical manufacturing processes: gate-to-gate information for life cycle assessment. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environ. Clean Technol.* **2003,** *78* (9), 995-1005.

(29) Jolliffe, I. Principal component analysis. In *International encyclopedia of statistical science*, Springer: 2011; pp 1094-1096.

(30) Netzeva, T. I.; Worth, A. P.; Aldenberg, T.; Benigni, R.; Cronin, M. T.; Gramatica, P.; Jaworska, J. S.; Kahn, S.; Klopman, G.; Marchant, C. A. Current status of methods for defining the applicability domain of (quantitative) structure-activity relationships: The report and recommendations of ecvam workshop 52. *Altern. Lab. Anim.* **2005,** *33* (2), 155-173.

(31) Ozturk, H.; Ozkirimli, E.; Ozgur, A. A comparative study of SMILES-based compound similarity functions for drug-target interaction prediction. *BMC Bioinformatics* **2016,** *17* (1), 128-128.

(32) Zhu, X.; Wang, X.; Ok, Y. S. The application of machine learning methods for prediction of metal sorption onto biochars. *J. Hazard. Mater.* **2019**, 378, 120727.

(33) Zhu, X.; Wu, G.; Coulon, F.; Wu, L.; Chen, D. Correlating asphaltene dimerization with its molecular structure by potential of mean force calculation and data mining. *Energ. Fuel.* **2018,** *32* (5), 5779-5788.

(34) Schmidhuber, J. Deep learning in neural networks: An overview. *Neural networks* **2015,** *61*, 85-117.

(35) Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M. In *TensorFlow: a system for large-scale machine learning*, operating systems design and implementation, 2016; pp 265-283.

(36) Remy, D. I. C.; Ruhland, I. A. Ecological assessment of alternative sanitation concepts with Life Cycle Assessment. *Technical University Berlin, Berlin, Germany* **2006,** *55*.

(37) Lamnatou, C.; Baig, H.; Chemisana, D.; Mallick, T. K. Environmental assessment of a building-integrated linear dielectric-based concentrating photovoltaic according to multiple life-cycle indicators. *J. Clean. Prod.* **2016,** *131*, 773-784.

(38) Huijbregts, M. A.; Steinmann, Z. J.; Elshout, P. M.; Stam, G.; Verones, F.; Vieira, M.; Zijp, M.; Hollander, A.; van Zelm, R. ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level. *Int. J. Life Cycle Ass.* **2017,** *22* (2), 138-147.

(39) Garrod, S.; Bollard, M. E.; Nicholls, A. W.; Connor, S. C.; Connelly, J.; Nicholson, J. K.; Holmes, E. Integrated metabonomic analysis of the multiorgan effects of hydrazine toxicity in the rat. *Chem. Res. Toxicol.* **2005,** *18* (2), 115-122.

(40) Jolliffe, I. T.; Cadima, J. Principal component analysis: a review and recent developments. *Philos. T. R. Soc. A: Mathematical, Physical and Engineering Sciences* **2016,** *374* (2065), 20150202.

(41) Liberti, L.; Lavor, C.; Maculan, N.; Mucherino, A. Euclidean Distance Geometry and Applications. *Siam Rev.* **2014,** *56* (1), 3-69.

(42) Gallegos-Saliner, A.; Poater, A.; Jeliazkova, N.; Patlewicz, G.; Worth, A. P. Toxmatch-A chemical classification and activity prediction tool based on similarity measures. *Regul. Toxicol. Pharm.* **2008,** *52* (2), 77-84.

(43) Stavropoulos, P.; Giannoulis, C.; Papacharalampopoulos, A.; Foteinopoulos, P.; Chryssolouris, G. Life cycle analysis: Comparison between different methods and optimization challenges. *Procedia CIRP* **2016,** *41*, 626-631.

(44) Zelm, R. ReCiPe 2008: a life cycle impact assessment method which comprises harmonised category indicators at the midpoint and the endpoint level. *Den Haag, The NetherlandsGuinee JB, Gorree M, Heijungs R, Huppes G, Kleijn R, de Koning A, van Oers L, Sleeswijk AW, Suh S, Udo de Haes HA, de Bruijn H, van Duin R, Huijbregts MAJ (2002) Life cycle assessment an operational guide to the ISO standards, eco-efficiency in industry and science* **2009,** *7*, 445460.

(45) Zhu, X.; Li, Y.; Wang, X. Machine learning prediction of biochar yield and carbon contents in biochar based on biomass characteristics and pyrolysis conditions. *Bioresource technol.* **2019,** *288*, 121527.

(46) Martínez-Rocamora, A.; Solís-Guzmán, J.; Marrero, M. LCA databases focused on construction materials: A review. *Renew. Sust. Energ. Rev.* **2016,** *58*, 565-573.

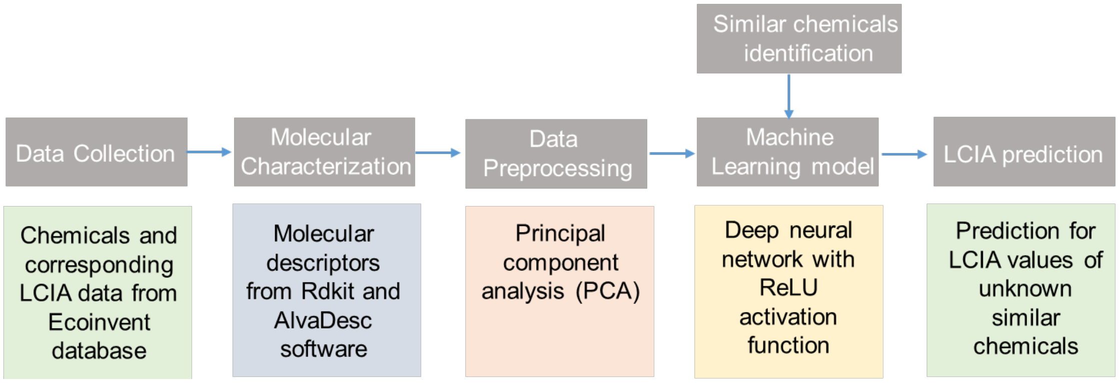

Similar chemicals identification
Data Collection
Molecular Characterization
Data Preprocessing
Machine Learning model
LCIA prediction
Chemicals and corresponding LCIA data from Ecoinvent database
Molecular descriptors from Rdkit and AlvaDesc software
Principal component analysis (PCA)
Deep neural network with ReLU activation function
Prediction for LCIA values of unknown similar chemicals

Fig. 1 High-throughput screening framework of green chemical substitutes based on life cycle assessment and machine learning method

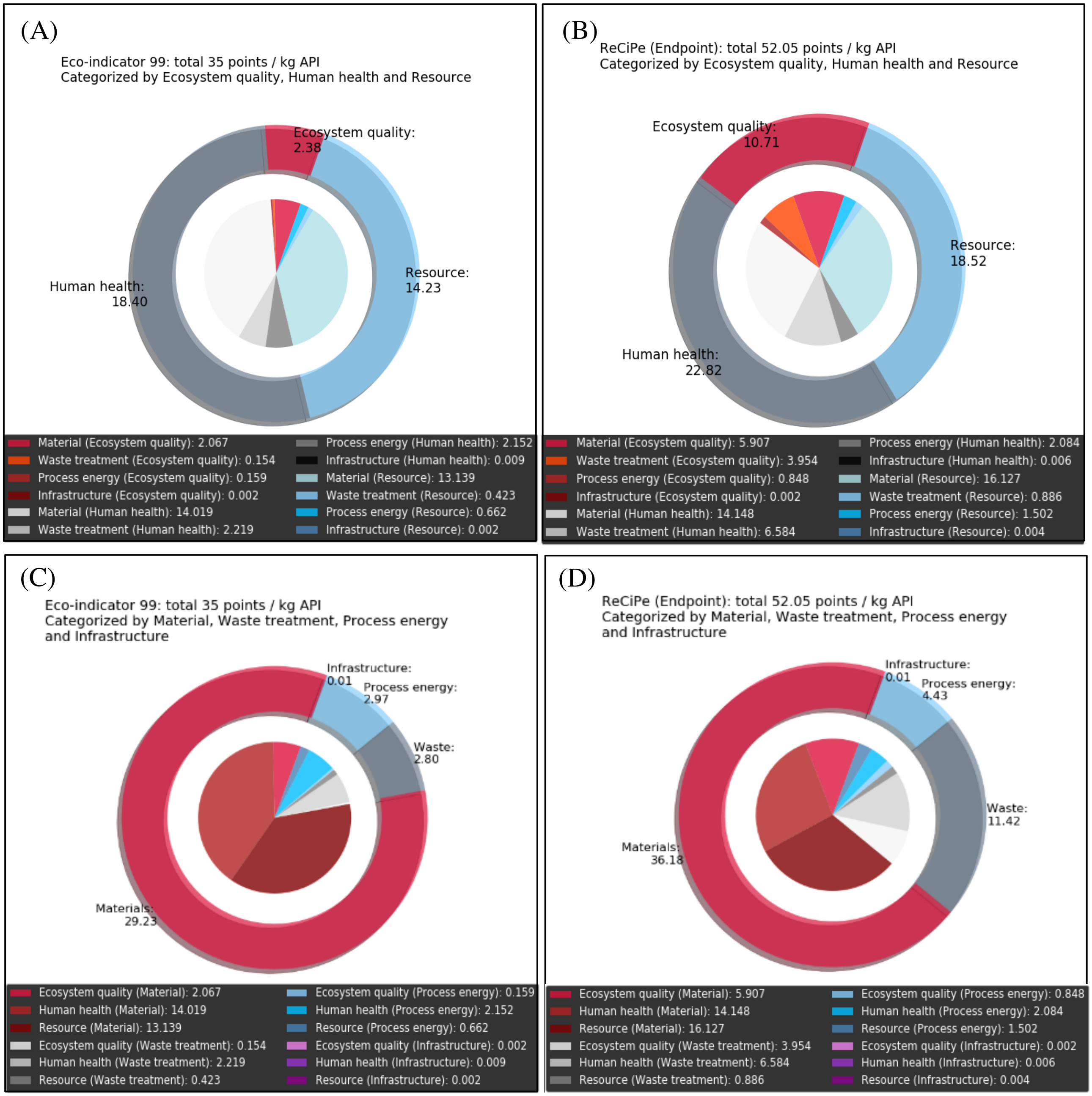


Fig. 2 LCA results of sitagliptin manufacturing process with EI99 (A, C) and ReCiPe endpoints (B, D). For A and B, the outside circle shows the total environmental threat to human health, ecosystem, and resource, while the inside circle shows the proportion of detailed influential factors, including chemical materials and infrastructure, the demand for process energy and the wastes treatment to human health, ecosystem and resource, respectively. They are reverse for C and D.

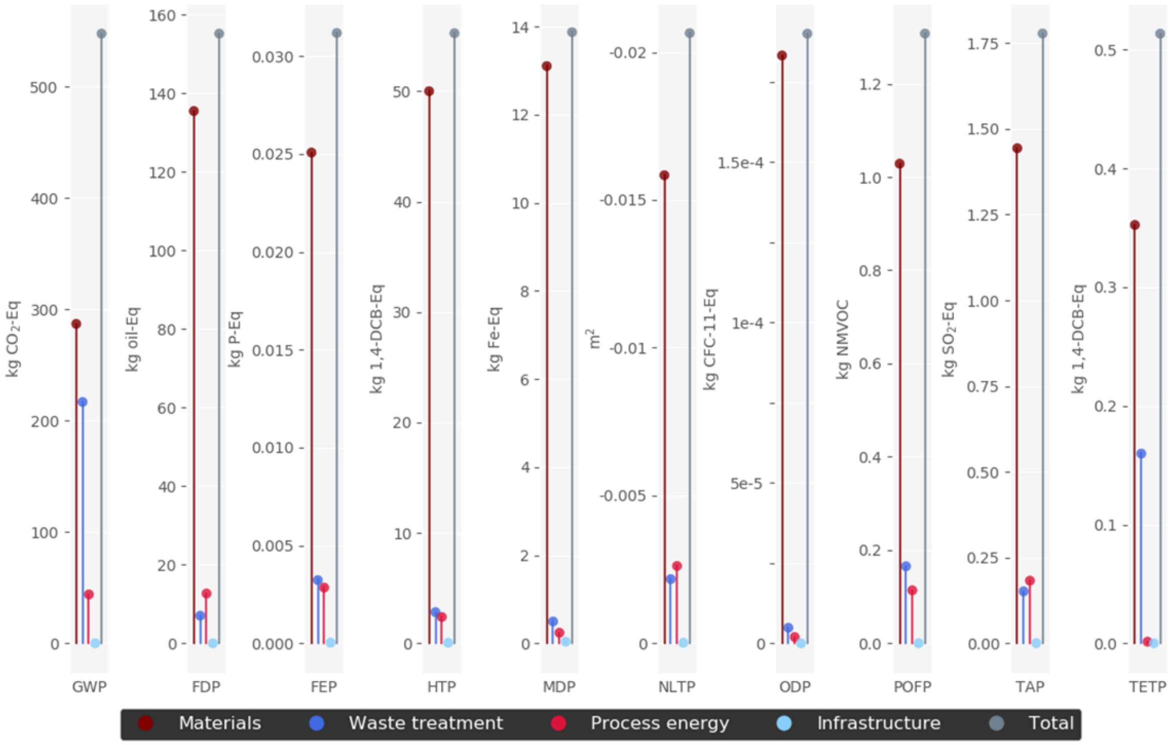


Fig. 3 LCIA results of sitagliptin manufacturing with ReCiPe midpoints according to the following detailed impact categories: Global Warming Potential (GWP), Fossil Fuel Depletion Potential (FDP), Freshwater Eutrophication Potential (FEP), Human Toxicity Potential (HTP), Metal Depletion Potential (MDP), Natural Land Transformation Potential (NLTP), Ozone Depletion Potential (ODP), Photochemical Oxidant Formation Potential (POFP), Terrestrial Acidification Potential (TAP), Terrestrial Eco-toxicity Potential (TETP).

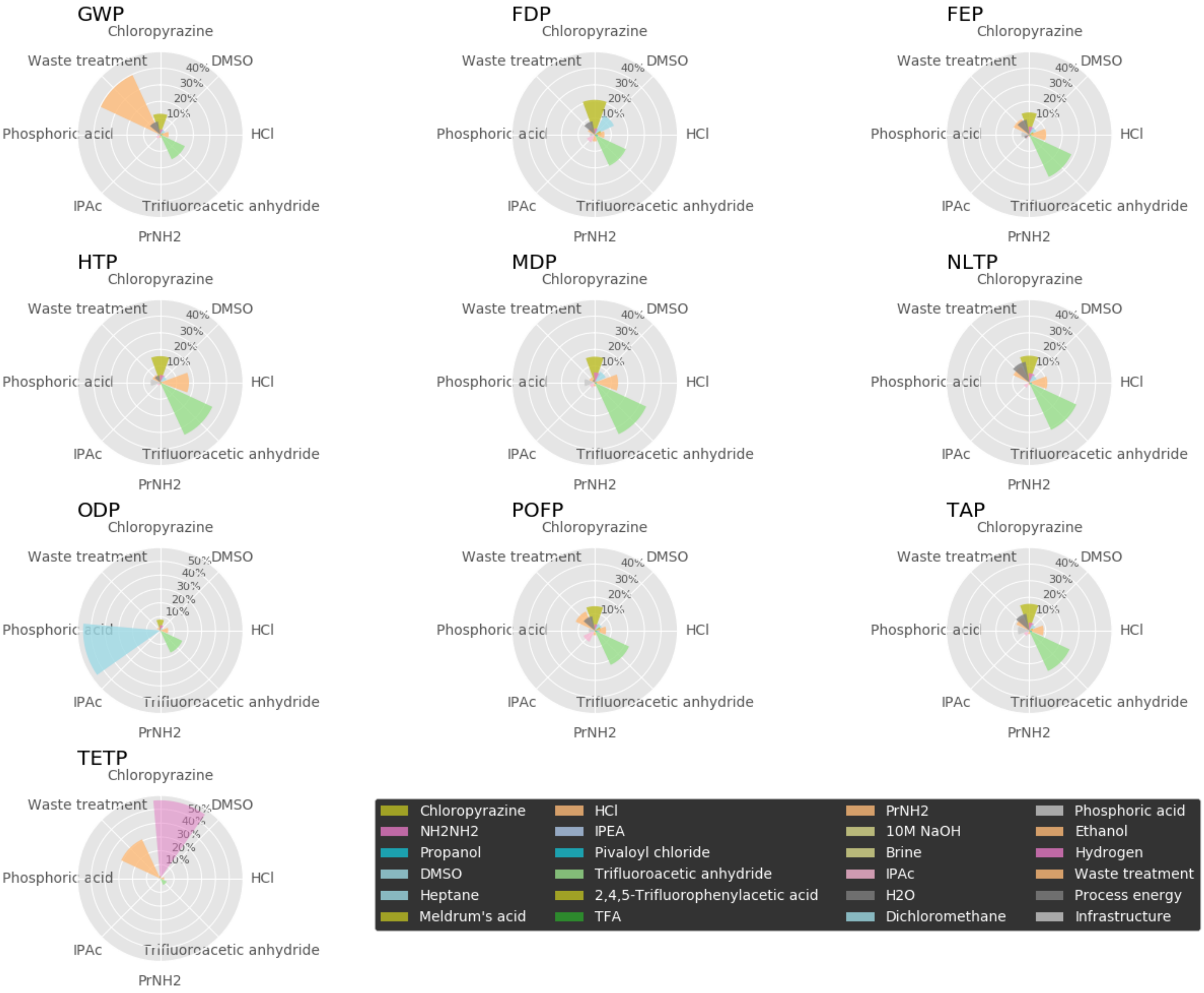


Fig. 4 Comparison of ReCiPe midpoints LCIA results of chemicals used in sitagliptin manufacturing. Following indictors were shown: Global Warming Potential (GWP); Fossil Fuel Depletion Potential (FDP); Freshwater Eutrophication Potential (FEP); Human Toxicity Potential (HTP); Metal Depletion Potential (MDP); Natural Land Transformation Potential (NLTP); Ozone Depletion Potential (ODP); Photochemical Oxidant Formation Potential (POFP); Terrestrial Acidification Potential (TAP); Terrestrial Eco-toxicity Potential (TETP).

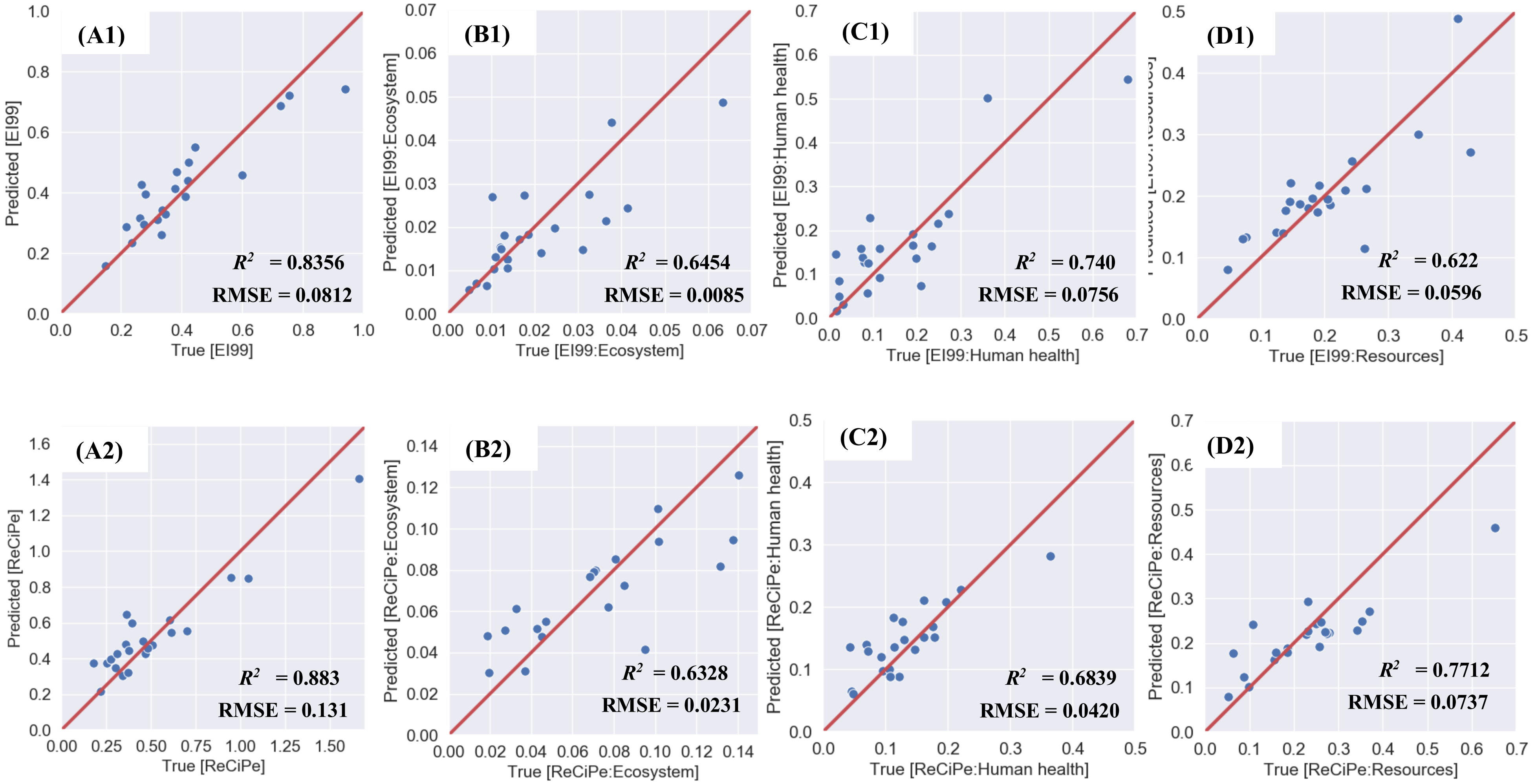


Fig. 5 Comparison of predicted LCA and actual values using test data for EI99 (A1), EI99-ecosystem (B1), EI99-Human health (C1), EI99-Resources (D1), and ReCiPe (A2), ReCiPe -ecosystem (B2), ReCiPe -Human health (C2), ReCiPe -Resources (D2), The red lines refer to the line y=x.

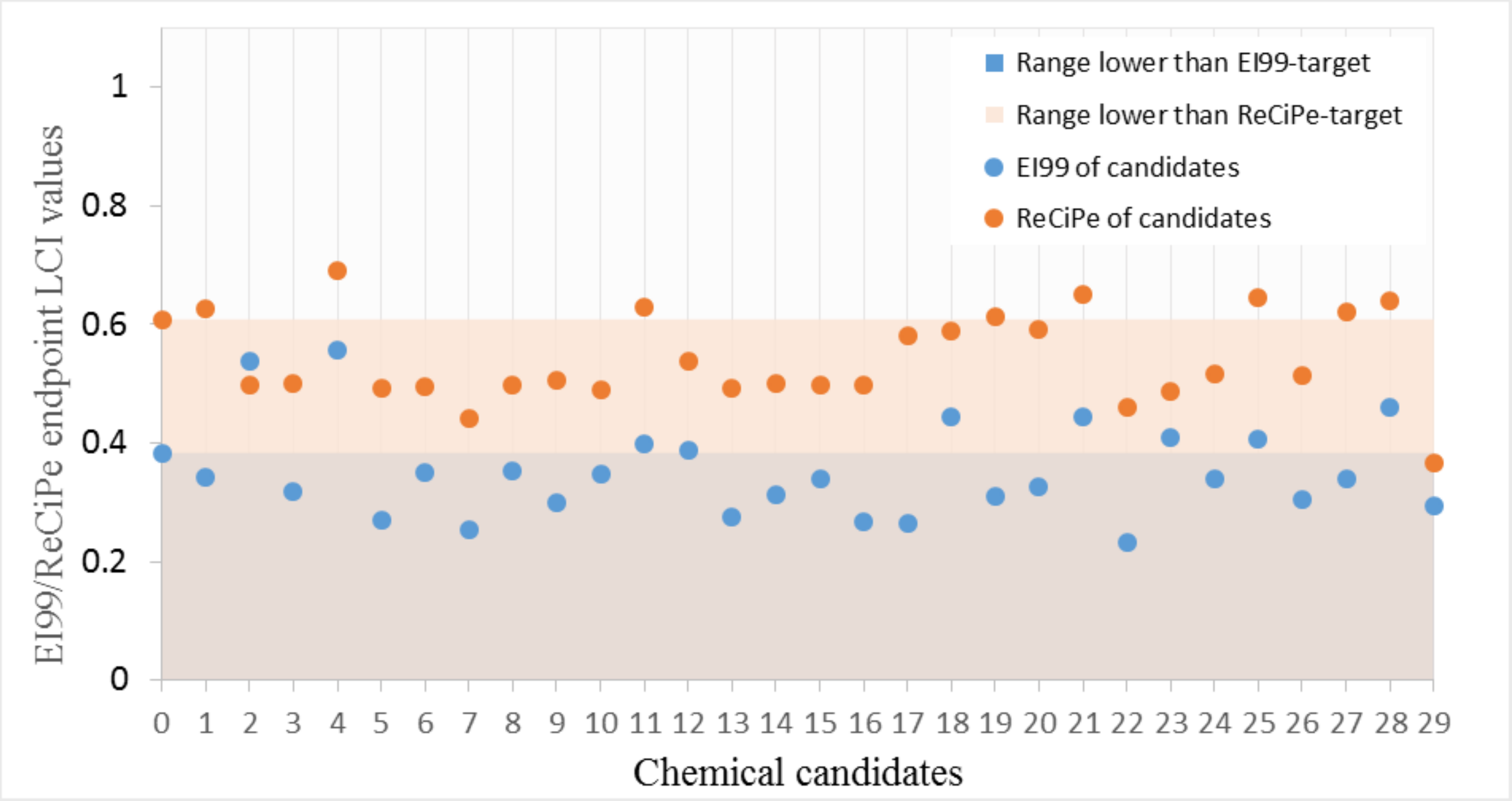


Fig. 6 LCIA prediction values of 30 most similar chemicals with trifluoroacetic anhydride

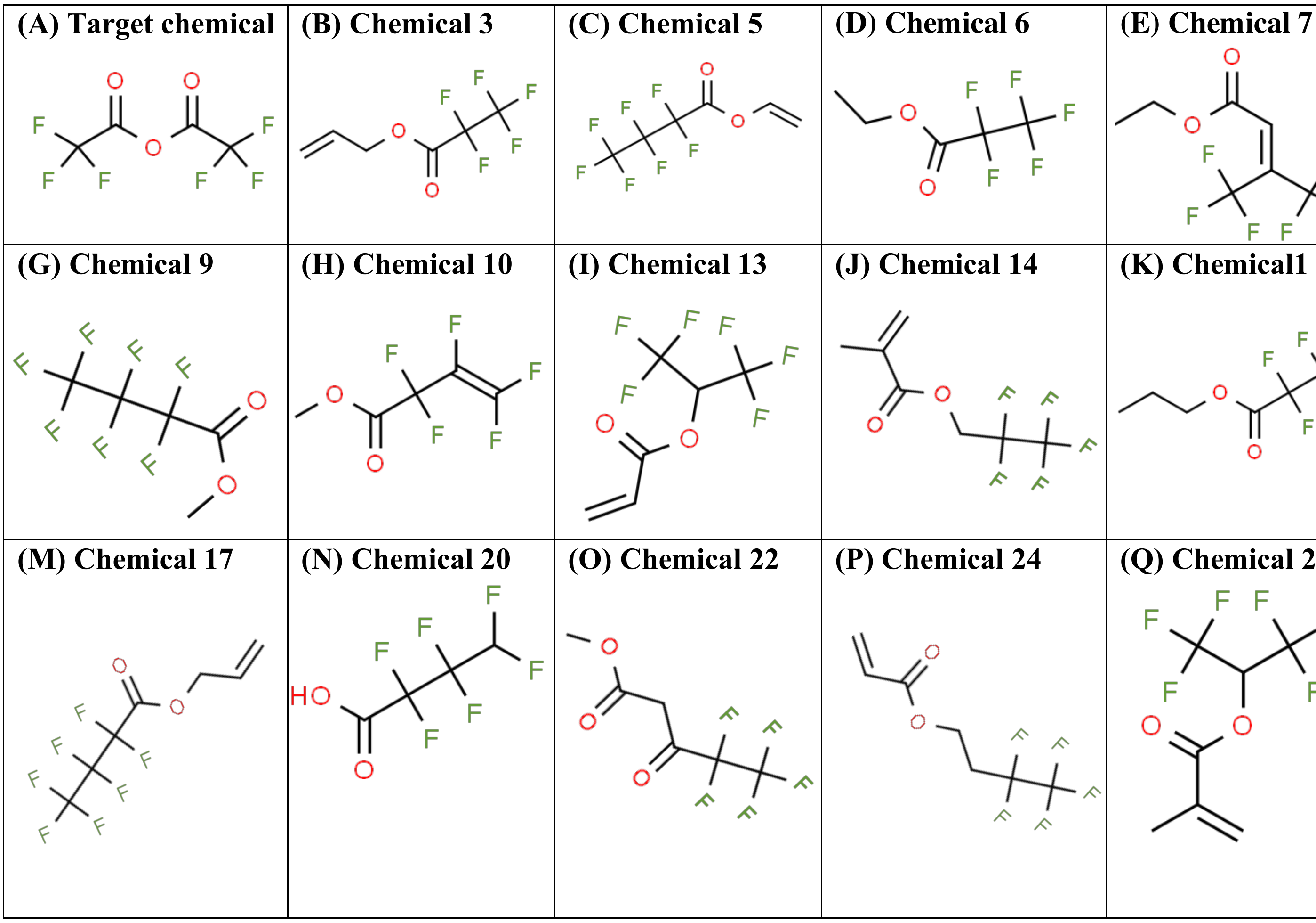


Fig. 7 The structures of target chemicals (A, Trifluoroacetic anhydride) and potential green chemical substitutes, including Allyl pentafluoropropanoate (B), Vinyl perfluoro butyrate (C), ethyl perfluoropropionate (D), Ethyl 4,4,4-trifluoro-3-(trifluoromethyl)-2-butenoate (E), 2,2,3,3,3-Pentafluoropropyl acrylate (F), Methyl heptafluorobutanoate (G), Methyl 2,2,3,4,4-pentafluoro-3-butenoate (H), 1,1,1,3,3,3-Hexafluoro-2-propanyl acrylate (I), 1H,1H-Pentafluoropropyl methacrylate (J), Propyl pentafluoropropanoate (K), Ethyl heptafluorobutanoate (L), Allyl heptafluorobutanoate (M), 2,2,3,3,4,4-hexafluorobutanoic acid (N), Methyl pentafluoropropiony-lacetate (O), 3,3,4,4,4-Pentafluorobutyl acrylate (P), 1,1,1,3,3,3-Hexafluor-2-propanylmethacrylat (Q), 1,2-ethanediyl ester (R).

Supporting Information

High-Throughput Screening of Green Chemical Substitutes with Life-Cycle Impact Using Machine Learning

Xinzhe Zhu [†], Chi-Hung Ho[†], Xiaonan Wang*

Department of Chemical and Biomolecular Engineering, Faculty of Engineering, National University of Singapore, Block E5, Engineering Drive 4, 117585 Singapore

† These authors contribute equally to this work.

* Corresponding author:

Tel: +65 6601 6221

Email: chewxia@nus.edu.sg

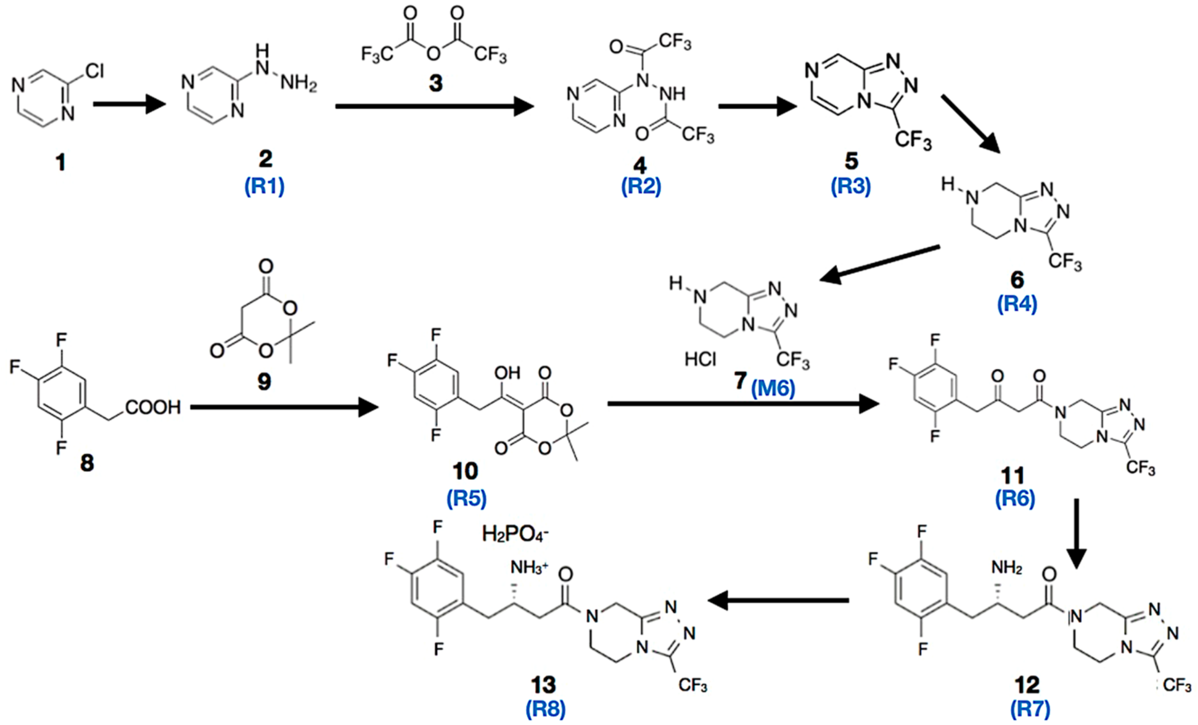


Fig. S1 Synthetic route of sitagliptin production used in the study [1]

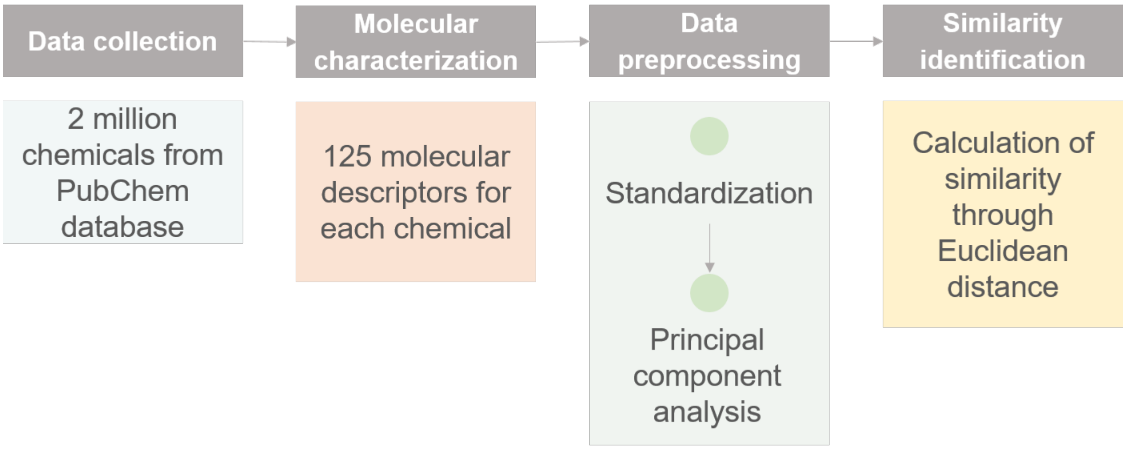

Data collection
Molecular characterization
Data preprocessing
Similarity identification
2 million chemicals from PubChem database
125 molecular descriptors for each chemical
Standardization
Principal component analysis
Calculation of similarity through Euclidean distance

Fig. S2 Flowsheet of the identification of similar chemicals from PubChem database

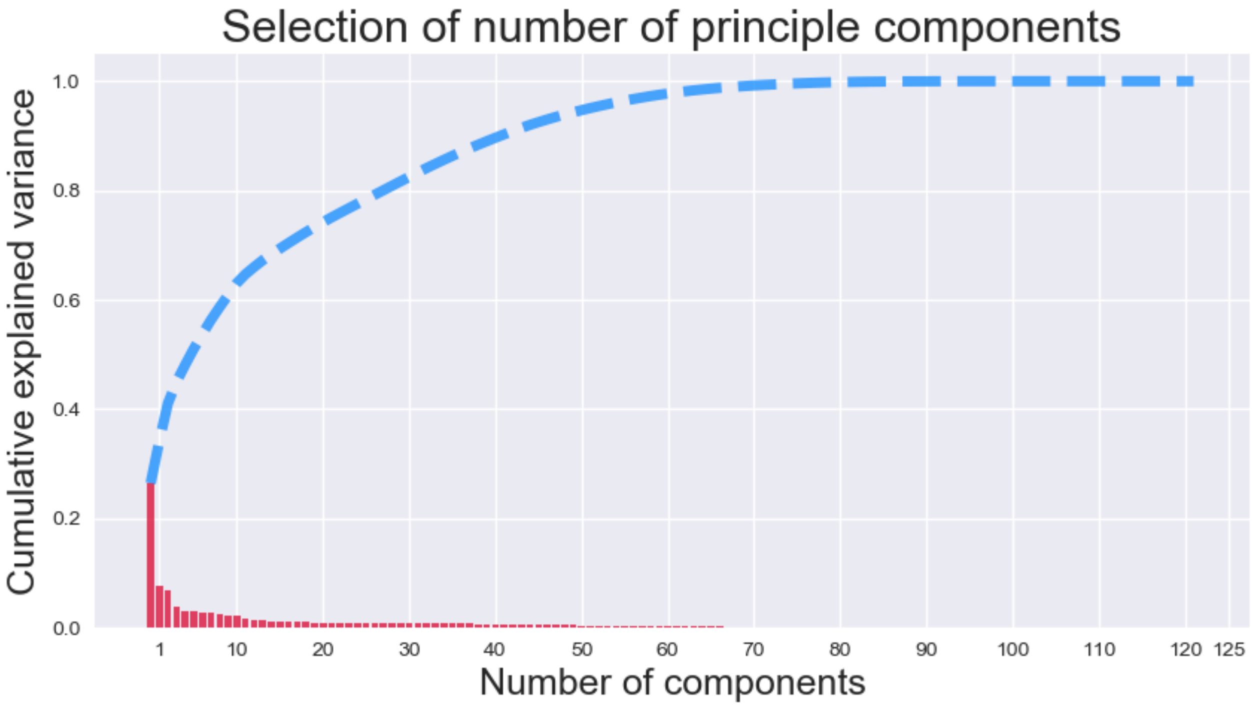


Fig. S3 Number of descriptors extracted by PCA against the cumulative variance preserved by the corresponding descriptors. The red referred to the information preserved of each principle component, while the blue one was the cumulative values of preserved principle components.

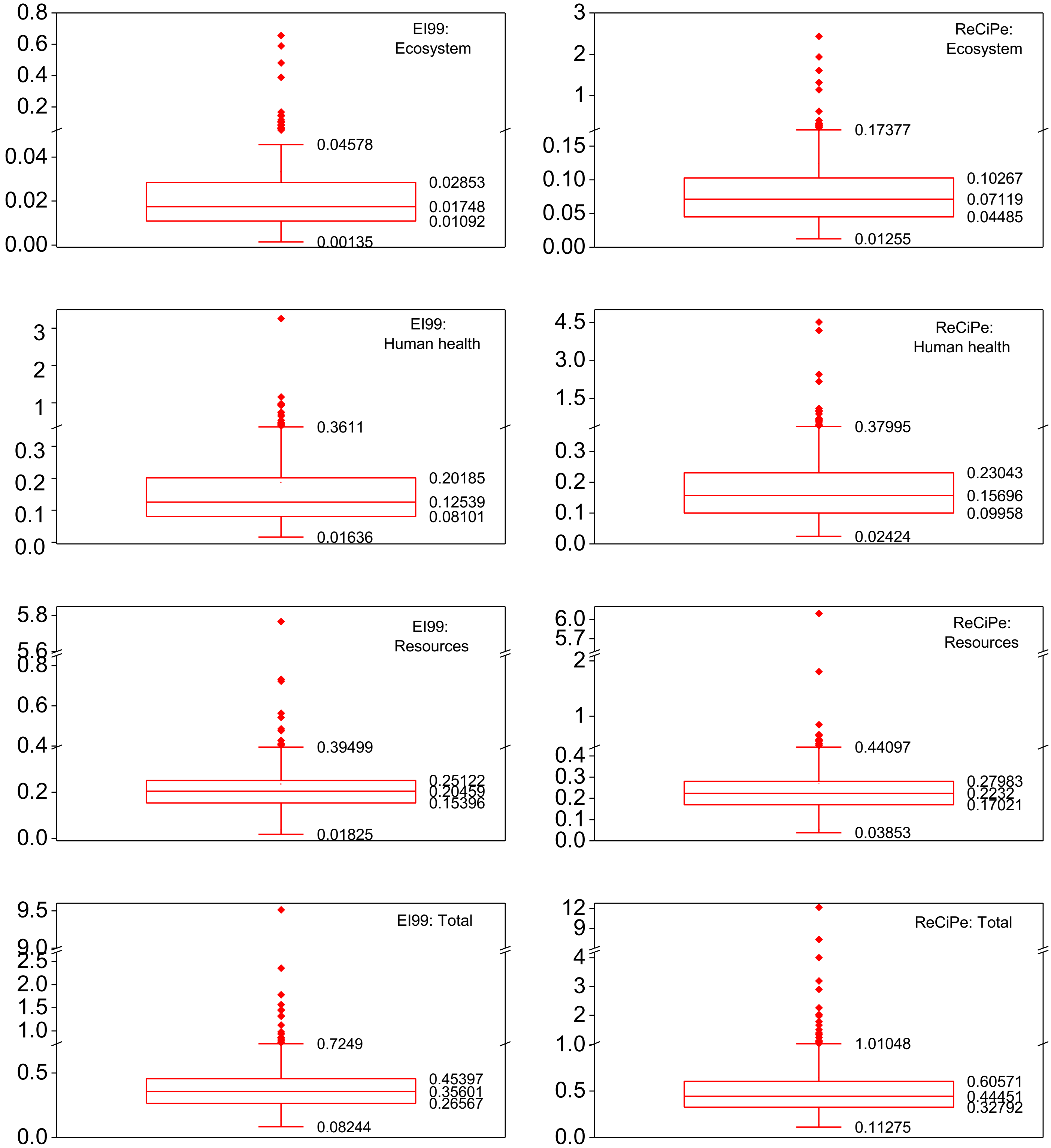


Fig. S4 LCA data distributions of nonionic organic chemicals from Ecoinvent v3.5 database

Table S1 Mass balance throughout the sitagliptin manufacturing process[1]

Mass flow rate (kg/hr)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18 (solid)	18 (solution)	19	20	21
1	0.03409	0.00000	0.03409	0.00341	0.00000	0.00000	0.00341	0.00068	0.00273	0.00000	0.00273	0.00273	0.00000	0.00273	0.00273	0.00000	0.00000	0.00000	0.00273	0.00000	0.00273	0.00000
H_2O	0.00000	0.02657	0.02657	0.02657	0.00000	0.00000	0.02657	0.02657	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NH_2NH_2	0.00000	0.01431	0.01431	0.00572	0.00000	0.00000	0.00572	0.00572	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.02950	0.00000	0.00000	0.02950	0.00361	0.02589	0.00000	0.02589	0.02589	0.00000	0.02589	0.00777	0.00000	0.00000	0.00000	0.00777	0.00000	0.00777	0.00000
HCl	0.00000	0.00000	0.00000	0.00977	0.00000	0.00000	0.00977	0.00977	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2-propanol	0.00000	0.00000	0.00000	0.00000	0.00147	0.02950	0.02950	0.00000	0.02950	0.02802	0.00147	0.01353	0.00000	0.01353	0.01353	0.00000	0.00000	0.00000	0.01353	0.00000	0.01353	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.01327	0.26548	0.26548	0.00000	0.26548	0.25220	0.01327	0.01327	0.00000	0.01327	0.01327	0.00000	0.00000	0.00000	0.01327	0.00000	0.01327	0.00000
IPAc	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10000	0.00000	0.10000	0.10000	0.00000	0.00000	0.00000	0.10000	0.00000	0.10000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.09877	0.09877	0.02963	0.00000	0.00000	0.00000	0.02963	0.00000	0.02963	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.04973	0.00000	0.00000	0.03481	0.01492	0.00000	0.01492	0.00000
Trifluoroacetic acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.03753	0.00000	0.00000	0.00000	0.03753	0.00000	0.03753	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10000	0.00500	0.00000	0.00000	0.00000	0.00000	0.09500
Superphosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01000	0.01000	0.00000
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
i-$PrNH_2$	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Acetone	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Phosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Mass flow rate (kg/hr)	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37 (Pd catalyst)	38	39	40	41	42	43
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00273	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2O	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.22857	0.00000	0.00000	0.00000	0.22857
NH_2NH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00777	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.09698	0.00000	0.00000	0.00000	0.10000
2-propanol	0.01218	0.00012	0.01206	0.00000	0.01206	0.00135	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.00000	0.01327	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPAc	0.09000	0.00090	0.08910	0.01090	0.10000	0.01000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.02963	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.01492	0.03481	0.00000	0.03481	0.00348	0.00348	0.00000	0.00000	0.00000	0.00348	0.00000	0.00348	0.00348	0.00000	0.00000	0.00000	0.00000
Trifluoroacetic acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.03753	0.00000	0.00000	0.00000	0.01182	0.01182	0.00000	0.00000	0.00000	0.01182	0.00000	0.01182	0.01182	0.00000	0.00000	0.00000	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00500	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Superphosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.01000	0.00000	0.05000	0.05000	0.05000	0.05000	0.00000	0.00000	0.00000	0.05000	0.00000	0.05000	0.05000	0.00000	0.00000	0.00000	0.00000
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01951	0.01951	0.00000	0.00000	0.00000	0.00390	0.00000	0.00390	0.00390	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01000	0.00967	0.00033	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01577	0.00000	0.01577	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01880	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.20000	0.01000	0.19000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
i-$PrNH_2$	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Acetone	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Phosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Mass flow rate (kg/hr)	44 (solid)	44 (solution)	45	46	47	48	49	50	51	52	53	54	55	56	57	58 (solid)	58 (solution)	59	60	61	62	63
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2O	0.00000	0.22857	0.22857	0.00000	0.22857	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.20000	0.00000	0.20000	0.30000	0.00000	0.10000	0.04086	0.04086
NH_2NH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HCl	0.00000	0.09698	0.09698	0.00000	0.09698	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00214	0.00000	0.00000	0.00214	0.00214	0.00000	0.00000	0.00000	0.00000
2-propanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPAc	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00348	0.00348	0.00000	0.00348	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Trifluoroacetic acid	0.00000	0.01182	0.01182	0.00000	0.01182	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Superphosphoric acid	0.00000	0.05000	0.05000	0.00000	0.05000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	0.00390	0.00390	0.00000	0.00390	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.01331	0.00548	0.00548	0.01331	0.00548	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01331	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.20000	0.20000	0.00000	0.01000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.01176	0.00000	0.00000	0.00000	0.01176	0.00035	0.00000	0.00035	0.00035	0.00000	0.00000	0.00035	0.00035	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00891	0.00000	0.00000	0.00891	0.00045	0.00000	0.00045	0.00045	0.00000	0.00000	0.00045	0.00045	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00746	0.00000	0.00746	0.00746	0.00000	0.00746	0.00746	0.00000	0.00000	0.00746	0.00746	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01483	0.01483	0.01483	0.00000	0.01483	0.01483	0.00000	0.00000	0.01483	0.01483	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.01858	0.00000	0.01858	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00129	0.00000	0.00129	0.00129	0.00000	0.00000	0.00129	0.00129	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00201	0.00201	0.00201	0.00000	0.00000	0.00201	0.00201	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.02387	0.00000	0.02111	0.00276	0.00276	0.02111	0.00000	0.00000	0.02111
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00588	0.00000	0.00000	0.00588	0.00588	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10000	0.00000	0.10000	0.04086	0.04086
i-$PrNH_2$	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Acetone	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Phosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Mass flow rate (kg/hr)	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2O	0.00000	0.04086	0.02724	0.06810	0.06810	0.04086	0.00070	0.04156	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.04156	0.04156	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NH_2NH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00030	0.00030	0.00000	0.00000	0.00000	0.00000	0.00000	0.00027	0.00003	0.00030	0.00027	0.00000	0.00000	0.00000	0.00000	0.00000
2-propanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPAc	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.20000	0.01190	0.18810	0.00190	0.19000	0.20000	0.00000	0.20000	0.01000	0.00400	0.36000	0.36400	0.20000	0.16400
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Trifluoroacetic acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Superphosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.02111	0.00001	0.02112	0.00003	0.00002	0.00000	0.00002	0.00000	0.00000	0.00000	0.00000	0.00000	0.00002	0.00000	0.00002	0.00002	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.04086	0.02724	0.06810	0.06810	0.04086	0.00000	0.04086	0.00000	0.00000	0.00000	0.00000	0.00000	0.03677	0.00409	0.04086	0.03677	0.00000	0.00000	0.00000	0.00000	0.00000
i-$PrNH_2$	0.03074	0.03074	0.01943	0.05017	0.04857	0.02914	0.00000	0.02914	0.00000	0.00000	0.16065	0.00162	0.16227	0.17081	0.01898	0.18979	0.00854	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00734	0.00734	0.01834	0.01101	0.00000	0.01101	0.00000	0.00000	0.00000	0.00000	0.00000	0.00099	0.01002	0.01101	0.00099	0.00000	0.00000	0.00000	0.00000	0.00000
Acetone	0.00000	0.00000	0.00105	0.00105	0.00262	0.00157	0.00000	0.00157	0.00000	0.00000	0.00000	0.00000	0.00000	0.00141	0.00016	0.00157	0.00141	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00405	0.00405	0.01011	0.00607	0.00000	0.00607	0.00000	0.00000	0.00000	0.00000	0.00000	0.00607	0.00000	0.00607	0.00607	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Phosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Mass flow rate (kg/hr)	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104 (anhydrous Na_2SO_4)	105	106	107
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2O	0.00000	0.00000	0.00144	0.04300	0.04300	0.00000	0.04300	0.04300	0.04300	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NH_2NH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2-propanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPAc	0.20000	0.03600	0.00000	0.00000	0.20000	0.20000	0.00000	0.20000	0.00000	0.20000	0.40000	0.00000	0.00000	0.00000	0.00000	0.00000	0.40000	0.40000	0.00000	0.40000	0.04000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Trifluoroacetic acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Superphosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.00000	0.00000	0.00409	0.00409	0.00368	0.00041	0.00041	0.00004	0.00037	0.40451	0.40451	0.00405	0.40047	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
i-$PrNH_2$	0.00000	0.00000	0.00000	0.01898	0.01898	0.01708	0.00190	0.00190	0.00019	0.00171	1.87895	1.87895	0.01879	1.86016	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00000	0.01002	0.01002	0.00781	0.00220	0.00220	0.00013	0.00207	0.01097	0.00110	0.00001	0.00109	0.00000	0.00000	0.00987	0.00987	0.00000	0.00987	0.00987	0.00000
Acetone	0.00000	0.00000	0.00000	0.00016	0.00016	0.00014	0.00002	0.00002	0.00000	0.00001	0.01554	0.01554	0.00016	0.01538	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00056	0.00054	0.00054	0.00048	0.00005	0.00005	0.00001	0.00005	0.05303	0.05303	0.00053	0.05250	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00005	0.00005	0.00004	0.00000	0.00000	0.00000	0.00000	0.00483	0.00483	0.00005	0.00478	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.10000	0.10000	0.00100	0.09900	0.00100	0.10000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.20000
Phosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000

Mass flow rate (kg/hr)	108	109	110	111	112	113	114 (solid)	114 (solution)	115	116	117	118	119	120	121	122	123
1	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2O	0.00000	0.00000	0.00000	0.00126	0.00126	0.00126	0.00000	0.00126	0.00126	0.00126	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NH_2NH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
HCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
2-propanol	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
dichloromethane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPAc	0.00000	0.00000	0.04000	0.00000	0.04000	0.04000	0.00000	0.04000	0.04000	0.00400	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
4	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Trifluoroacetic acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
heptane	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Superphosphoric acid	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
H_2 gas	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
8	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
9	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PivCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
IPEA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
10	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R5	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
TFA	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
11	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R6	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
DMSO	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
i-PrNH_2	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
12	0.00000	0.00000	0.00987	0.00000	0.00987	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Acetone	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
by-product from R7	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaOH	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
NaCl	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Brine	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
Ethanol	0.01000	0.19000	0.20000	0.00000	0.20000	0.20000	0.00000	0.20000	0.20000	0.01000	0.00000	0.00000	0.00010	0.01000	0.00010	0.00990	0.01000
Phosphoric acid	0.00000	0.00000	0.00000	0.00713	0.00713	0.00475	0.00000	0.00475	0.00475	0.00475	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
13	0.00000	0.00000	0.00000	0.00000	0.00000	0.01225	0.01102	0.00122	0.00122	0.00122	0.01102	0.01102	0.00000	0.00000	0.00000	0.00000	0.00000

Table S2 Summary of life cycle impact assessment methods used in this work

LCIA method	Impact category			
EI99	Ecosystem	Human health	Resources	Total
ReCiPe endpoints	Ecosystem	Human health	Resources	Total
ReCiPe midpoints	Global Warming Potential (GWP, in kg of CO$_2$ equivalents per FU)			
	Ozone Depletion Potential (ODP, in kg of chlorofluorocarbon-11 per FU)			
	Terrestrial Acidification Potential (TAP, in kg of SO$_2$ equivalents per FU)			
	Freshwater Eutrophication Potential (FEP, in kg of P equivalents per FU)			
	Human Toxicity Potential (HTP, in kg of 1,4-dichlorobenzene per FU)			
	Photochemical Oxidant Formation Potential (POFP, in kg of non-methane volatile organic compounds per FU)			
	Terrestrial Ecotoxicity Potential (TETP, in kg of 1,4-DCB per FU)			
	Natural Land Transformation Potential (NLTP, in m^2 per FU)			
	Metal Depletion Potential (MDP, in kg of Fe equivalents per FU)			
	Fossil Fuel Depletion Potential (FDP, in kg of oil equivalents per FU)			

Table S3 Thirty most similar chemicals based on Euclidean distance

	The SMILE structure of similar chemicals	Euclidean distance after PCA
0	O=C(OC(=O)C(F)(F)F)C(F)(F)F	0.000
1	O=C(OCC(F)(F)F)C(F)(F)F	2.534
2	O=C(NOC(=O)C(F)(F)F)C(F)(F)F	2.847
3	C=CCOC(=O)C(F)(F)C(F)(F)F	2.949
4	C=C(F)C(=O)OC(F)(F)C(C)(F)F	3.062
5	C=COC(=O)C(F)(F)C(F)(F)C(F)(F)F	3.063
6	CCOC(=O)C(F)(F)C(F)(F)F	3.121
7	CCOC(=O)C=C(C(F)(F)F)C(F)(F)F	3.155
8	C=CC(=O)OCC(F)(F)C(F)(F)F	3.167
9	COC(=O)C(F)(F)C(F)(F)C(F)(F)F	3.189
10	COC(=O)C(F)(F)C(F)=C(F)F	3.292
11	O=C([O-])C(F)(F)C(F)(F)C(F)F	3.303
12	COC(=O)C(F)(F)C(F)(F)F	3.323
13	C=CC(=O)OC(C(F)(F)F)C(F)(F)F	3.325

14	C=C(C)C(=O)OCC(F)(F)C(F)(F)F	3.328
15	CCCOC(=O)C(F)(F)C(F)(F)F	3.370
16	CCOC(=O)C(F)(F)C(F)(F)C(F)(F)F	3.473
17	C=CCOC(=O)C(F)(F)C(F)(F)C(F)(F)F	3.498
18	O=C(O)C(F)(F)C(F)(F)C(F)(F)F	3.500
19	C=CC(=O)OCC(F)(F)C(F)(F)C(F)(F)F	3.599
20	O=C(O)C(F)(F)C(F)(F)C(F)F	3.627
21	O=C(C=C(O)C(F)(F)F)C(F)(F)F	3.632
22	COC(=O)CC(=O)C(F)(F)C(F)(F)F	3.634
23	O=C(Cl)ON(C(F)(F)F)C(F)(F)F	3.693
24	C=CC(=O)OCCC(F)(F)C(F)(F)F	3.729
25	O=C(CC(=O)C(F)(F)F)C(F)(F)F	3.735
26	C=C(C)C(=O)OC(C(F)(F)F)C(F)(F)F	3.744
27	C=C(C)C(=O)OCC(F)(F)C(F)C(F)(F)F	3.770
28	CC(=CC(=O)C(F)(F)F)C(F)(F)F	3.806
29	O=C(OCCOC(=O)C(F)(F)F)C(F)(F)F	3.809

References

(1) Ho, C.; Yi, J.; Wang, X. Biocatalytic continuous manufacturing of diabetes drug: plantwide process modeling, optimization, environmental and economic analysis. *ACS Sustain. Chem. Eng.* **2018**, 7(1): 1038-1051.