J Am Oil Chem Soc (2020) **97:** 817–837
DOI 10.1002/aocs.12361

ORIGINAL ARTICLE

The Application of Quantitative Structure–Property Relationship Modeling and Exploratory Analysis to Screen Catalysts for the Synthesis of Oleochemical Carbonates from CO_2 and Bio-Based Epoxides

Victor Hugo Jacks Mendes dos Santos[1,2,3] · **Darlan Pontin**[1] · **Raoní Scheibler Rambo**[3] · **Marcus Seferin**[1,2,3]

Received: 26 November 2019 / Revised: 1 April 2020 / Accepted: 3 April 2020

Abstract Screening catalysts for the synthesis of cyclic carbonates from CO_2 and epoxides presents a challenge for the research community. Thus, we propose the application of quantitative structure–property relationships (QSPR) modeling and exploratory analysis to assist in the selection of catalysts to produce oleochemical carbonates. QSPR modeling was developed by applying 2D-descriptors to evaluate the relationship between the molecular structure of organocatalysts and their activity in the production of bio-based organic carbonates. From the virtual screening, 122 potential catalysts were selected, their catalytic activities were estimated, and the best molecular targets highlighted. Already from the data mining and exploratory analysis, the catalysts' key structural features (e.g. organic structure, molecular arrangement, carbon chain size, and substituent type) were identified. Thus, it was possible to evaluate the similarity between the catalysts and to relate the 2D-descriptors to their activity. Then, based on QSPR modeling results, cetyltrimethylammonium bromide (CTAB) was proposed as a new catalyst to produce oleochemical carbonates. From the CTAB application, conversions greater than 98% of epoxide were observed in the cycloaddition of CO_2 to epoxidized vegetable oil (rice bran, canola, and soybean). Thus, it was concluded that QSPR modeling and exploratory analysis show potential for screening catalysts for oleochemical carbonate synthesis.

Keywords Vegetable oil · Cyclic carbonate · Quantitative structure–activity relationship · QSAR · Ionic liquid · Organocatalyst

J Am Oil Chem Soc (2020) **97:** 817–837.

Abbreviations

ALogP	Ghose-Crippen-Viswanadhan octanol–water partition coefficient
apol	sum of the atomic polarizabilities (including implicit hydrogens)
ATS2e	Broto-Moreau autocorrelation—lag 2/weighted by Sanderson electronegativities
bpol	sum of the absolute value of the difference between atomic polarizabilities of all bonded atoms in the molecule (including implicit hydrogens)
C2SP3	singly bound carbon bound to two other carbons
ETA Shape Y	extended topochemical atom shape index Y
GATS6i	geary autocorrelation—lag 6/weighted by first ionization potential

---

Supporting information Additional supporting information may be found online in the Supporting Information section at the end of the article.

✉ Victor Hugo Jacks Mendes dos Santos
victor.santos@pucrs.br

[1] School of Technology, PUCRS—Pontifical Catholic University of Rio Grande do Sul, 6681 Ipiranga Avenue—Building 12, Porto Alegre, 90619-900, Brazil

[2] Engineering and Materials Technology Graduate Program, Pontifical Catholic University of Rio Grande do Sul, 6681 Ipiranga Avenue—Building 32, Porto Alegre, 90619-900, Brazil

[3] Institute of Petroleum and Natural Resources, Pontifical Catholic University of Rio Grande do Sul, 6681 Ipiranga Avenue—Building 96J, Porto Alegre, 90619-900, Brazil

Lipoaffinity Index	atom type electrotopological state lipoaffinity index
MATS4m	moran autocorrelation—lag 4/weighted by mass
nAtom	number of atoms
nAtomLAC	number of atoms in the longest aliphatic chain
nBonds2	total number of bonds (including bonds to hydrogens)
nBr	number of bromine atoms
nCl	number of chlorine atoms
nI	number of iodine atoms
nRotBt	number of rotatable bonds, including terminal bonds
$SssCH_2$	sum of atom-type E-state: $–CH_2–$
VABC	Van der Waals volume

Introduction

The replacement of the petrochemical production base and the development of low carbon technologies became one of the main concerns of humanity at the beginning of the century. Carbon dioxide (CO_2) is an important anthropogenic greenhouse gas, with increasing atmospheric concentration and a significant role in global climate change (Alves et al., 2017). Thus, it is necessary to propose ways to avoid CO_2 emissions or find a safe destination for its surplus, which are issues widely addressed by Life Cycle Assessment (LCA) and Carbon Capture, Utilization, and Storage (CCUS) studies (Cuéllar-Franca and Azapagic, 2015; North and Styring, 2015).

From gaseous waste to the valorization of CO_2 as a raw material for chemicals, the perception of the role of CO_2 in a low-carbon economy scenario has been changing significantly. Thus, the utilization of CO_2 as a C1 building block will play an important role in the low-carbon-based chemical industry (Büttner et al., 2017a; Sternberg et al., 2017). From the CO_2 utilization, several valuable chemicals such as methanol, hydrocarbons, carboxylic acids, carbamates, urea, polymers (polyurethanes and polycarbonates), and organic carbonates can be produced (Gomes et al., 2012; Liu and Wang, 2017; Mustafa et al., 2020; Yu et al., 2018). However, CO_2 is a notoriously unreactive chemical whose activation presents significant energy barriers and thermodynamic drawbacks, which must be overcome by chemical and physical processes (Cai et al., 2017; Poliakoff et al., 2015). Therefore, the use of catalysts is essential for carbon dioxide-based processes to be economically viable and with a minimum energy penalty (Alves et al., 2017; Poliakoff et al., 2015).

One of the most important alternatives for using carbon dioxide in chemical production is by cycloaddition of CO_2 to epoxides to produce cyclic organic carbonates (Alves et al., 2015; Aquino et al., 2015; Vieira et al., 2018). In this process, carbon dioxide coupling has 100% atomic efficiency and great industrial potential, since there is an established industry of epoxidized derivatives (Alves et al., 2017; Cokoja et al., 2015). In a cycloaddition reaction, CO_2 and the epoxy group are activated by catalysts such as transition metals, phase transfer catalysts with alkaline halides, and organocatalysts (Alves et al., 2015; Appel et al., 2013; Büttner et al., 2017a; Desens and Werner, 2016). Ideally, these processes should involve few steps and use catalysts that enable chemical routes with positive carbon balance and energy savings (Gomes et al., 2012). Therefore, screening catalysts for the synthesis of cyclic carbonates from CO_2 and epoxides is a challenge to be addressed.

To reduce the time and costs involved in scientific research and to increase the chemical and mechanistic understanding of the carbonation process, the cheminformatics methods: Molecular Dynamics (MD), Quantum Mechanics, Quantitative Structure–Activity Relationship (QSAR), and Quantitative Structure–Property Relationships (QSPR) could be applied (Blay et al., 2016). In general, QSPR is based on the concept that it is possible to describe the properties of a compound from its molecular structure (Katritzky and Lobanov, 1995; Stec et al., 2015). Then, the QSPR aims to establish a cause-effect relationship between molecular characteristics and observed properties through mathematical and statistical tools (Begam and Kumar, 2016; Roy et al., 2012a; Roy et al., 2017). In the modeling process, three steps are involved: (1) structure representation, (2) descriptor analysis, and (3) model calibration and validation. Thus, QSPR can be used to reduce the time and costs involved in the development process by filling data gaps, predicting material properties, and guiding the choice of new molecular targets (Karelson et al., 1996; Roy et al., 2018).

Vegetable oils and their derivatives are renewable, abundant, inexpensive, biodegradable, and nontoxic (Karmakar et al., 2017; Miao et al., 2014; Samanta et al., 2016). Recently, there has been a growing interest in the production and application of cyclic carbonates derived from oleochemicals (fatty acids, methyl esters, and triacylglycerols). Among the main reasons for the great potential of bio-based carbonates is the high availability of CO_2 and the existence of a consolidated industry for the production of epoxidized oleochemicals, such as the epoxidized soybean oil (ESBO), a plasticizers widely applied in the polyvinyl chloride (PVC) industry (Assen et al., 2016; Danov et al., 2017).

The use of a catalyst is a key component of the chemical process to increase efficiency and reduce production costs. From the first report on the production of oleochemical carbonates by Tamami et al. (2004), to recent publications (Cai et al., 2019; Peña Carrodeguas et al., 2017; Zheng et al., 2018), most results are obtained using tetrabutylammonium halides as a catalyst for the carbonation process (Büttner et al., 2016, 2017b; Longwitz et al., 2018; Peña Carrodeguas et al., 2017; Schäffner et al., 2014; Tamami et al., 2004; Tenhumberg et al., 2016; Wang et al., 2012b; Zheng et al., 2018). Only recently, screening studies of catalysts for the production of oleochemical carbonates have been reported (Alves et al., 2015; Büttner et al., 2016, 2017b; Longwitz et al., 2018; Schäffner et al., 2014; Tenhumberg et al., 2016; Wang et al., 2012b); however, the description of new catalysts for the production of cyclic carbonates from carbon dioxide and epoxidized derivatives is still limited. Thus, QSPR modeling can be applied to screen catalysts and guide the development of new systems (Achary et al., 2016; Cruz et al., 2007; Fayet et al., 2009; Maldonado and Rothenberg, 2010; Martínez et al., 2012; Ratanasak et al., 2015; Rothenberg, 2008; Yao et al., 1999).

To date, a small number of catalysts have been applied to produce oleochemical carbonates from CO_2 and epoxidized vegetable oil, and only a few results were generated under comparable conditions to allow QSPR data modeling. Another relevant aspect is that many organocatalysts applied in scientific reports do not have chemical structures described in public databases such as PubChem. Therefore, the present work proposes the application of QSPR modeling and exploratory analysis to screen catalysts for the synthesis of oleochemical carbonates from CO_2 and bio-based epoxides. A representative number of catalysts were compiled and the first study on the application of QSPR modeling and exploratory analysis tools to select a new catalyst for the synthesis of organic carbonate derived from vegetable oils (rice bran, canola, and soybean) was presented. Then, based on the QSPR results, cetyltrimethylammonium bromide (CTAB) was proposed as a new catalyst to produce bio-based organic carbonates.

Materials and Methods

Catalyst selection is a key step in processes involving the cycloaddition of CO_2 to epoxidized oleochemicals. Thus, we propose the application of QSPR modeling and exploratory analysis to assist in the selection of catalysts to produce bio-based organic carbonates. The scope of the work comprises catalysts derived from organic halide salts and oleochemical carbonates derived from vegetable oil (triacylglycerols, fatty acids, and fatty acid alkyl esters) presented in Fig. 1. Only catalysts with a structure registered in public databases such as PubChem were considered.

Materials

The chemicals used were hydrogen peroxide (35%), glacial acetic acid (>99%), sulfuric acid (>95%), n-butanol (99% purity), CTAB (98%), high-purity carbon dioxide (99.995%), and three vegetable oils (rice bran, canola, and soybean) obtained from local suppliers. All reagents were used without further purification, while information on the fatty acid composition of vegetable oils and the average ethylenic unsaturation per triacylglycerol molecules is presented in Table S10.

QSPR Modeling and Exploratory Analysis Data Sets

This paper presents three data sets applied for the QSPR modeling (Data Set 01 and 02) and exploratory analysis (Data Set 03) of catalysts based on 2D-molecular descriptors. Data Set 01 (Table 1), from Alves et al. (2015), comprises 12 catalysts with structures registered in a public database, and was applied for variable selection, calibration, and validation of QSPR and to predict the activity of potential new catalysts. The application domain of this set considers the synthesis of organic cyclic carbonate from CO_2 and epoxidized triacylglycerols at 100 °C, 10 MPa CO_2, 5 hours, 1 mol% catalyst load, and constant stirring.

Already Data Set 02 (Table S4), from Büttner et al. (2017b), comprises nine catalysts with structures registered in a public database, and was applied to evaluate the transferability of the QSPR model and was developed with the same descriptors selected for Data Set 01 (pages S8–S16, Fig. S2–S4, and Tables S4–S8). The application domain of this set considers the synthesis of organic cyclic carbonate from CO_2 and epoxidized methyl oleate at 100 °C, 5 MPa CO_2, 16 hours, and 2 mol% catalyst load.

Furthermore, Data Set 03 (Table S9) comprises a representative number of catalysts (29 catalysts) that have been applied to produce oleochemical carbonates from CO_2 and epoxide. Exploratory analysis was applied to this data set based on unsupervised multivariate method and 2D-molecular descriptors to evaluate the key structural features of the catalysts. In the exploratory analysis, the same molecular descriptors previously selected were applied and the results are presented in the Supporting Information (Pages S16–S20, Fig. S5–S6, Table S9, and Table S17).

Descriptor Calculation

Molecular descriptors transcribe the chemical, physical, and biological characteristics of the chemical structure in mathematical terms, which need to be processed with statistical tools to develop pattern recognition or predictive models (Golbraikh and Tropsha, 2002; Katritzky et al., 1997). The molecular structures of catalysts were primarily

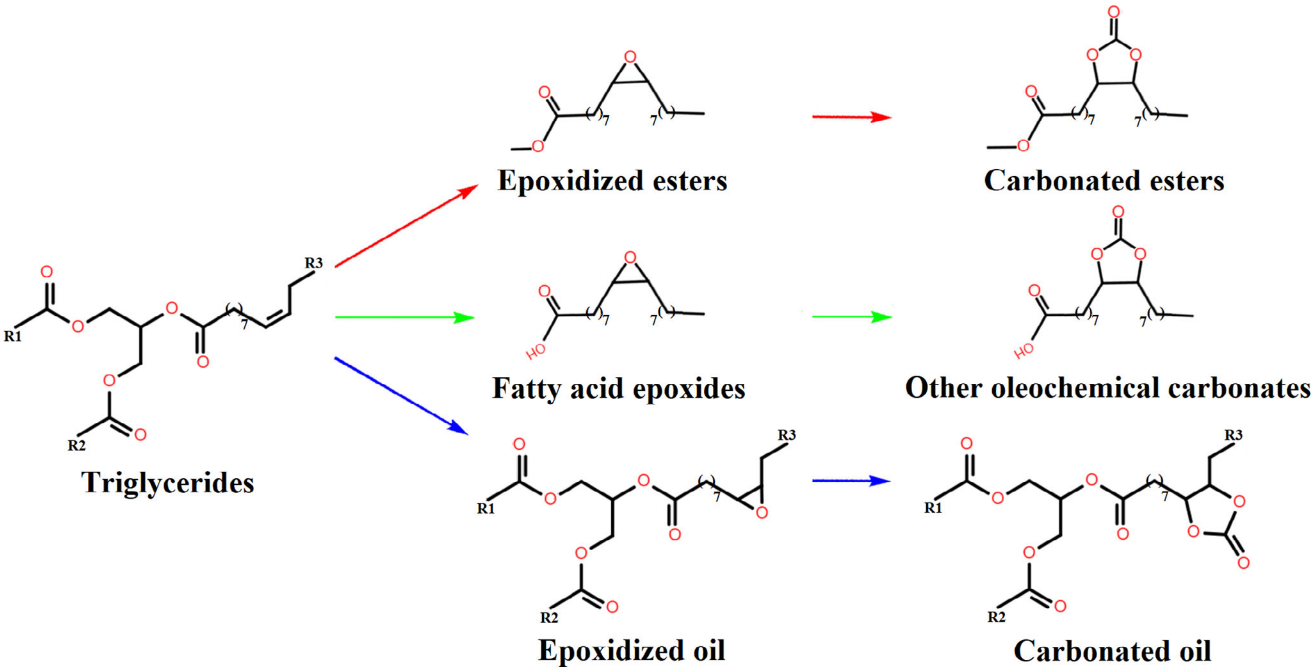

Fig. 1 Oleochemical carbonates derived from triacylglycerols

obtained from the PubChem database and stored in SDF files (Structured Data Format) (Kim, 2016). Then, 2D-molecular descriptors of the optimized structures are obtained using the PaDEL-Descriptor software (http://www.yapcwsoft.com/dd/padeldescriptor), comprising an initial data set of 1444 descriptors (Yap, 2011).

Variable Selection

Variable selection is an essential step in the QSAR/QSPR study to reduce the initial number of descriptors for a representative set of variables and enable the development of interpretable models. For QSPR modeling, molecular descriptors are applied as predictive variables (X), while epoxide conversion to carbonates (Data Set 01) was used as the response variable (Y). Then, the variable correlation analysis was applied and the linear correlation of the 1444 molecular descriptors with the conversion of epoxide to carbonates was evaluated. Thus, only the variables that present a fair correlation with the response variable ($R^2 > 0.3$ or $R^2 < -0.3$) were kept in the data set for subsequent variable selection steps (Kiralj and Ferreira, 2010).

Table 1 Catalyst Data Set 01, applied for QSPR modeling

Catalyst	PubChem CID	CAS	Conversion (%)
Tetrabutylammonium iodide	67553	311-28-4	26
Tetrabutylammonium bromide	74236	1643-19-2	30
Tetrabutylammonium chloride	70681	1112-67-0	17
Tetrabutylphosphonium iodide	201022	3115-66-0	21
Tetrabutylphosphonium bromide	76564	3115-68-2	28
Tetrabutylphosphonium chloride	75311	2304-30-5	19
1-Methyl-3-octylimidazolium iodide	71353115	188589-28-8	25
1-Methyl-3-octylimidazolium bromide	10849985	61545-99-1	30
1-Methyl-3-octylimidazolium chloride	2734223	64697-40-1	20
Triethylsulfonium iodide	74589	1829-92-1	0
1-Butyl-1-methylpyrrolidinium iodide	11076461	56511-17-2	19
1-Butylpyridinium iodide	14007922	874-81-7	12

Data from Alves et al. (2015).

After correlation analysis, the selection of variables proceeds by the stepwise method, based on the combination of forward selection and backward elimination procedures, applied within the PLS regression by using the leave-one-out (LOO) internal validation method. The selection stepwise variable is a time-consuming procedure that interactively included / excluded (step by step) the predictor variables (X) until the best model is obtained (Gonzalez et al., 2008; Mitra et al., 2010; Xu and Zhang, 2001). Considering the number of catalysts from Data Set 01, and to obtain a reliable model with a reasonable number of molecular descriptors, the variable selection procedure was repeated several times until the stepwise method uses the interval size of one. This strategy was previously presented as an alternative for making QSAR/QSPR models based on a small data set as robust as possible (Mitra et al., 2010).

Molecular Descriptors

From the variable selection step, the number of 1444 variables was reduced to only 18 molecular descriptors. This set of descriptors was selected by the stepwise method algorithm and are the variables that showed the greatest potential for predicting catalytic activity for the carbonation reaction, resulting in multivariate models with the best adjustments. Thus, the 18 molecular descriptors applied in QSPR modeling are: nCl^-, nBr^-, nI^-, ALogP, apol, ATS2e, bpol, C2SP3, ETA Shape Y, GATS6i, Lipoaffinity Index, MATS4m, nAtom, nAtomLAC, nBonds2, nRotBt, $SssCH_2$, and VABC. The details of the selected variables can be found in the "Abbreviations" and its definition can be found in the literature (Ghose and Crippen, 1986, 1987; Liu et al., 2001; Roy and Ghosh, 2004; Todeschini and Consonni, 2009; Zhao et al., 2003).

Data Analysis

Data analysis was performed using the Solo + MIA software (Eigenvector Research) and the statistical tools applied for QSPR modeling and exploratory analysis are Correlation Analysis (CA), Principal Component Analysis (PCA), Partial Least Squares Regression (PLS), and Support Vector Machine Regression (SVM).

QSPR Modeling

For QSPR modeling, molecular descriptors are applied as predictive variables (X), while epoxide conversion to carbonates was used as the response variable (Y). After the exhaustive variable selection step, 18 molecular descriptors are applied to perform the multivariate regression with the data autoscaled and mean centered (Rothenberg, 2008). Then, PLS was performed using the SIMPLS algorithm, while SVM was developed using the Linear Kernel Function.

QSPR Validation

Model validation is a crucial step in QSPR modeling and several criteria and threshold values have been presented as minimum requirements to ensure the robustness and transferability of QSAR/QSPR models (Alexander et al., 2015; Gramatica and Sangion, 2016; Pratim Roy et al., 2009; Todeschini et al., 2016; Tropsha et al., 2003). Thus, the validation of the QSPR models was performed according to the criteria of Golbraikh and Tropsha (Golbraikh and Tropsha, 2002; Tropsha, 2010; Tropsha et al., 2003), and r_m^2 metrics of Roy and coworkers (Mitra et al., 2010; Pratim Roy et al., 2009; Roy et al., 2012b; Roy and Mitra, 2012). A summary of the validation parameters and their limit values are given in Table 2.

Considering the size of Data Set 01, model stability was assessed through internal validation by leave-one-out (Q^2-LOO) and leave-many-out (Q^2-LMO), while predictability was evaluated using the $r_m^2{}_{(LOO)}$ and $r_m^2{}_{(LMO)}$ parameters by replacing the $R^2{}_{(test\ set)}$ with the cross-validation Q^2 (Gramatica and Sangion, 2016; Mitra et al., 2010; Pratim Roy et al., 2009; Roy and Mitra, 2012). In addition to internal validation, the data for the 12 catalysts are divided into independent training (9 samples) and external test sets (3 samples). The catalysts of each of the external test sets (03 catalysts) do not compose the calibration set. Thus, to avoid underestimation or overestimation of the QSPR model, all combinations of independent training and external test sets (220 PLS models) were modeled, and $R^2{}_{(cal)}$, $Q^2{}_{(LOO)}$, $R^2{}_{(test\ set)}$, and root-mean-square error of

Table 2 Parameters used for QSPR model validation

Parameter	Threshold value
R^2	>0.6
Q^2	>0.5
$\frac{\left(R^2 - R'^2_o\right)}{R^2}$	<0.1
$\frac{\left(R^2 - R'^2_o\right)}{R^2}$	<0.1
k	$0.85 \le k \le 1.15$
k'	$0.85 \le k \le 1.15$
$\lvert R^2 - R^2_o \rvert$	<0.3
$\lvert R^2 - R'^2_o \rvert$	<0.3
R^2_m	>0.5
R'^2_m	>0.5
$\lvert R^2_m - R'^2_m \rvert$	<0.2
$\frac{\lvert R^2_m - R'^2_m \rvert}{2}$	>0.5

R^2—correlation coefficient between the predicted and observed activities for a test set.

prediction (RMSEP) parameters were evaluated as distribution histograms.

In the present work, the occurrence of overfitting problem was evaluated by reducing the number of molecular descriptors applied to develop the QSPR model (Data Set 01) and the results are presented in the Supporting Information (Pages S5–S7, Fig. S1, and Tables S1–S3), while the transferability of the QSPR model was evaluated based on Data Set 02. Thus, QSPR modeling of Data Set 02 was performed with the same molecular descriptors, and validation was evaluated based on the same criteria (Table 2). All QSPR modeling procedures (calibration, validation, and discussion) developed based on Data Set 02 are presented in the Supporting Information (Pages S8–S16, Figs S2–S4, and Tables S4–S8).

Virtual Screening

Virtual screening is a computational method, based on structure or property, that guides the search for new active compounds in large chemical libraries (Roy and Mitra, 2011; Scior et al., 2012). So, PubChem search tools were applied to perform virtual screening of possible molecular targets with chemical structures similar to those used for calibration models (Kim, 2016). Then, 2D-molecular descriptors were obtained using the PaDEL-Descriptor software and the data evaluated to remove molecules with missing data and outliers (using PCA). Thus, 122 potential catalysts were compiled and are presented in the Table S13, while a summary of the data analysis procedures performed in the present work are depicted in Fig. 2.

Synthetic Procedures

The synthetic procedures were performed in two stages: epoxidation of raw vegetable oil and carbonation of epoxidized oil.

Epoxidation Reactions

In situ epoxidation of vegetable oils was performed with glacial acetic acid, hydrogen peroxide (35%), and sulfuric acid. Reactions were performed at 75 °C for 6 hours under mechanical stirring and by using hydrogen peroxide to ethylenic unsaturation molar ratio (2.0), acetic acid to ethylenic unsaturation molar ratio (0.5), and sulfuric acid catalyst (2 wt% based on aqueous fraction) (Dinda et al., 2008; Goud et al., 2007). Then, after the reactions, the products were dissolved in ethyl ether and washed with water to neutral pH, followed by removal of the solvent under vacuum. The specific dosage of chemicals for each of the reactions was estimated from the ethylenic unsaturation of each vegetable oil given in Table S10.

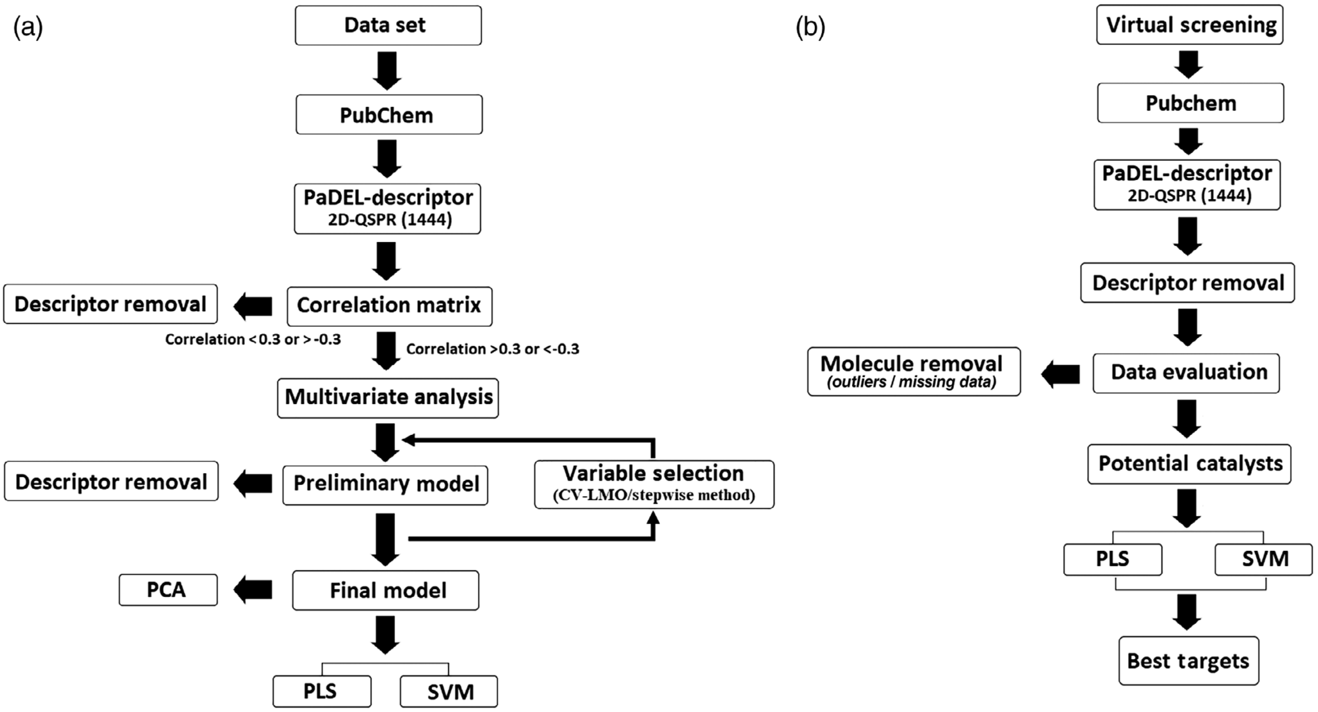

Fig. 2 Summary of QSPR modeling procedures: (a) QSPR calibration and validation and (b) virtual screening workflow

Carbonation of Epoxidized Oil

The present work used its own protocol adapted to the available resources, and the carbonation reactions of epoxidized oils were performed using CTAB catalyst, high-purity carbon dioxide, and n-butanol as solvent. Reactions were performed in a 50 cm^3 stainless steel autoclave at 120 °C for 48 hours and without stirring by using 2 g of epoxidized oil, 5 MPa CO_2, 4 mL of butanol, and 5 wt% of CTAB. Then, the butanol was removed under vacuum and the product was redissolved in ethyl acetate and washed twice with water and once with brine. Finally, the oleochemical carbonate was dried over anhydrous sodium sulfate and the solvent was removed under vacuum. Due to the technical limitations of our reactors, we cannot reproduce the pressure (10 MPa CO_2 at supercritical state) and stirring conditions of the QSPR model application domain (Data Set 01).

Characterization Methods

All vegetable oils, epoxidized oils, and carbonated products are characterized by Fourier Transform Infrared Spectroscopy (FTIR) and Nuclear Magnetic Resonance Spectroscopy (NMR) detailed below.

Infrared Analysis (FTIR)

Infrared spectra are obtained using Spectrum One (PerkinElmer) with HATR accessory, using the spectral range from 4000 to 650 cm^{-1}, 4 cm^{-1} resolution, and 16 scans per spectrum.

^{1}H NMR

All ^{1}H NMR spectra were recorded on a Bruker Avance 400 running at 400 MHz. Chemical shifts (δ) are reported in parts per million (ppm) relative to TMS signal (0 ppm) for ^{1}H NMR and using deuterated chloroform ($CDCl_3$) as solvent. The following abbreviations are used to indicate the multiplicity in NMR spectra: s—singlet; bs—broad singlet; d—doublet; t—triplet; q—quartet; m—multiplet; dd—double doublet.

Three parameters, (1) number of ethylenic unsaturation, (2) number of epoxy groups, and (3) epoxy group conversion in the carbonation reaction, were estimated based on the ^{1}H NMR spectra, and Fig. S7 shows the structure and NMR peaks of the oleochemical derivatives. For the calculations, it was necessary to use a normalization factor (NF) using Eq. 1. The signals of the four methyl protons of the glycerol portion (B) were used, since this relationship remains constant in all derivatives of triacylglycerols (raw, epoxidized, and carbonated) (Mazo and Rios, 2012).

$$NF = \frac{B}{4} \tag{1}$$

Thus, the number of ethylenic unsaturation of vegetable oils was estimated from Eq. 2, based on olefinic hydrogen signal (C) and NF (Kumar et al., 2012), while the number of epoxy groups of epoxidized vegetable oils was estimated from Eq. 3, based on oxirane ring hydrogens (D) and NF (Mazo and Rios, 2012).

$$\text{Ethylenic unsaturation}\;(\mathrm{C=C}) = \frac{C}{2NF} \tag{2}$$

$$\text{Epoxy group}\;(Ep) = \frac{D}{2NF} \tag{3}$$

Furthermore, the carbonation reaction conversion was estimated from Eq. 4, based on the consumption of epoxy groups in the cycloaddition of CO_2 to epoxidized vegetable oil (Aerts and Jacobs, 2004; Mazo and Rios, 2012; Xia et al., 2016).

$$\text{Conversion}\;(\eta\%) = 100 \times \left[\frac{Ep_{(\text{initial})} - Ep_{(\text{final})}}{Ep_{(\text{initial})}}\right] \tag{4}$$

Results and Discussion

The results and discussions were divided into five parts: (1) QSPR modeling based on 2D-molecular descriptors, (2) virtual screening of potential catalysts for the synthesis process, (3) data mining for pattern recognition of the QSPR model, (4) exploratory analysis of catalysts with known activity for the production of oleochemical carbonates (presented in the Supporting Information), and (5) synthesis of oleochemical carbonate from CTAB.

QSPR Modeling

The QSPR model was developed based on data from Alves et al. (2015), comprising 12 catalysts with structures registered in a public database (Table 1). After the variable selection step, 18 molecular descriptors were applied in the PLS and SVM regression and internal validations of the models were performed by the LOO and LMO cross-validation methods. Furthermore, leave-many-out (LMO) cross-validation was evaluated in two scenarios to assess the models' sensitivity, leaving out 16.7% and 25% of data in each calibration cycle. Thus, the QSPR model results are presented in Table 3, while the respective molecular descriptor values are available in Table S11.

Table 3 QSPR modeling to estimate catalyst activity in oleochemical carbonate synthesis

Data set 1	LOO		LMO[a]		LMO[b]	
	PLS	SVM	PLS	SVM	PLS	SVM
$R^2_{(cal)}$	0.9762	0.9747	0.9741	0.9715	0.9762	0.9769
$Q^2_{(CV)}$	0.9040	0.9142	0.9118	0.8373	0.8868	0.8896
RMSEC	1.25	1.30	1.31	1.58	1.26	1.26
RMSECV	2.56	2.44	2.44	3.30	3.57	2.94
F/SV	4	12	3	8	4	12

F, factor; SV, support vectors; RMSEC, root-mean-square error of calibration; RMSECV, root-mean-square error of cross-validation.

[a] 16.7% of the sample left out in the leave-many-out cross-validation.

[b] 25% of the sample left out in the leave-many-out cross-validation. The K-fold values of both LMO cross-validations are 6.

Table 3 shows that the QSPR model meets the minimum criteria of a reliable model, with high calibration coefficient of determination ($R^2 > 0.97$), good cross-validation coefficient of determination ($Q^2 > 0.83$), and acceptable root-mean-square error of cross-validation (RMSECV) values. Moreover, both multivariate regression models (PLS and SVM) showed similar results; however, PLS regression allows an interpretation of the results from the loadings and variable influence on projection (VIP scores) graphs. Thus, the validation of the QSPR model is evaluated and the results presented in Table 4, while the estimated conversion values are shown in Table S12.

Table 4 shows that all developed QSPR models are validated based on the criteria shown in Table 2, demonstrating potential to estimate the activity of new catalysts. In addition to internal validation, external validation procedures were performed by dividing catalyst data into training (nine samples) and external test (three samples) sets. Thus, to avoid underestimation or overestimation of the QSPR model, all combinations of independent training and test sets (220 PLS models) were modeled, and $R^2_{(cal)}$, $Q^2_{(LOO)}$, $R^2_{(test\ set)}$, and RMSEP parameters are presented in Fig. 3 as distribution histograms and summarized in Table 5.

Fig. 3 shows that most PLS models have high regression coefficient ($R^2_{(cal)}$, $Q^2_{(LOO)}$, $R^2_{(test\ set)}$) and low prediction errors (RMSEP), while Table 5 shows that all parameters have average values that meet the minimum requirements for QSPR model validation. Moreover, from the cumulative distribution function, it is observed that 95% of the PLS models present $Q^2_{(LOO)}$ (>0.5) and $R^2_{(test\ set)}$ (>0.6) above

Table 4 Validation of QSPR models based on PLS and SVM regressions

Data Set 01	PLS			SVM			Threshold value
	LOO	LMO[a]	LMO[b]	LOO	LMO[a]	LMO[b]	
R^2	0.98	0.97	0.98	0.97	0.97	0.98	>0.6
Q^2	0.90	0.91	0.89	0.91	0.84	0.89	>0.5
$\frac{(Q^2-Q^2_{o})}{Q^2}$	0.00	0.00	0.01	0.00	0.00	0.02	<0.1
$\frac{(Q^2-Q'^2_{o})}{Q^2}$	0.02	0.02	0.06	0.02	0.04	0.09	<0.1
k	0.99	0.99	0.99	0.99	0.99	0.99	$0.85 \leq k \leq 1.15$
k'	0.99	1.00	0.99	1.00	0.99	0.99	$0.85 \leq k \leq 1.15$
$\mid Q^2-Q^2_{o} \mid$	0.00	0.00	0.01	0.00	0.00	0.02	<0.3
$\mid Q^2-Q'^2_{o} \mid$	0.02	0.01	0.05	0.02	0.03	0.08	<0.3
Q^2_{m}	0.87	0.89	0.80	0.86	0.83	0.77	>0.5
Q'^2_{m}	0.78	0.80	0.68	0.78	0.69	0.63	>0.5
$\mid Q^2_{m}-Q'^2_{m} \mid$	0.09	0.08	0.12	0.09	0.14	0.13	<0.2
$\frac{\lvert \mathrm{Q2}_{m}+Q'^2_{m} \rvert}{2}$	0.82	0.85	0.74	0.82	0.76	0.70	>0.5
Validation	V	V	V	V	V	V	All criteria met

V, validated.

[a] 16.7% of the sample left out in the leave-many-out cross-validation.

[b] 25% of the sample left out in the leave-many-out cross-validation. The *K*-fold values of both LMO cross-validations are 6.

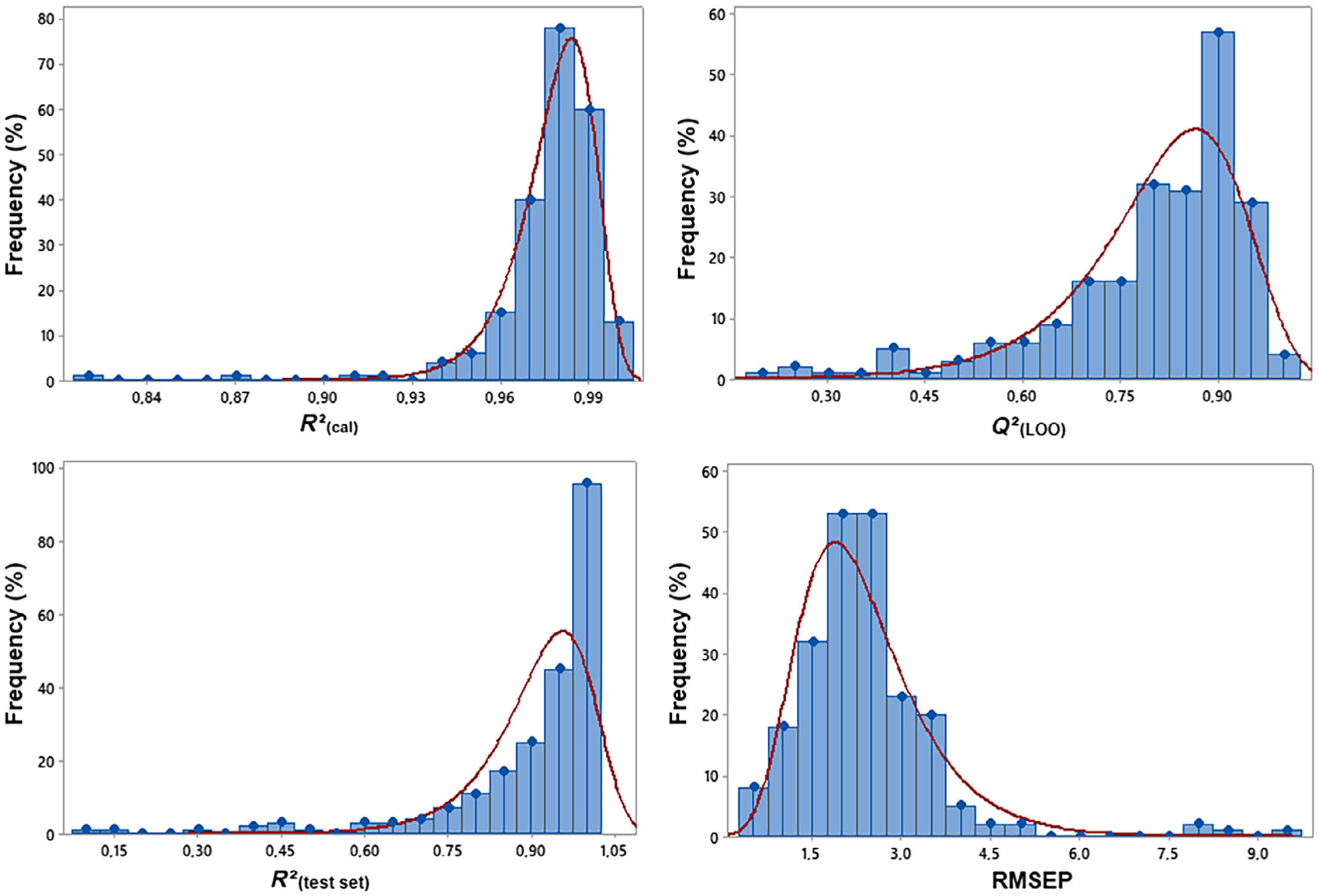

Fig. 3 Distribution histograms: (a) $R^2_{\text{(cal)}}$, (b) $Q^2_{\text{(LOO)}}$, (c) $R^2_{\text{(test set)}}$, and (d) RMSEP

the threshold values and acceptable prediction errors, with 80% of the models showing relative errors below 10% (RMSEP < 3.14) and 95% of them below 15% (RMSEP < 4.37). Thus, it is concluded that the QSPR model meets the internal and external validation criteria and presents good sensitivity for screening new catalysts. As the PLS and SVM regression models have been validated, the QSPR models can be applied in subsequent steps and Fig. 4 shows the predicted versus reference plot obtained from the LOO cross-validation of the PLS and SVM models.

From Fig. 4 and Table 3, it is observed that both regression models (PLS and SVM) result in models with good fit ($R^2 > 0.97$ and $Q^2 > 0.83$). Then, the interpretation of the relationship between the molecular descriptors (X) and the estimated conversion response (Y) was performed by the PLS regression coefficient (Das et al., 2017; Wold et al., 2001), shown in Fig. 5.

Table 5 QSPR model external validation results

Parameter	$R^2_{\text{(cal)}}$ [a]	$Q^2_{\text{(LOO)}}$ [a]	$R^2_{\text{(test set)}}$ [a]	RMSEP[a]
Mean	0.9776	0.8011	0.9036	2.360
SD	0.0189	0.1532	0.1483	1.170
Minimum	0.8172	0.1800	0.0880	0.410
Maximum	0.9994	0.9916	0.9999	9.570
Cumulative distribution (80%)	>0.9683	>0.7167	>0.8469	<3.14
Cumulative distribution (95%)	>0.9526	>0.5718	>0.7396	<4.37

RMSEP, root-mean-square error of prediction.

[a] Results of 220 PLS models performed with 75% of data for calibration and 25% as an external test set.

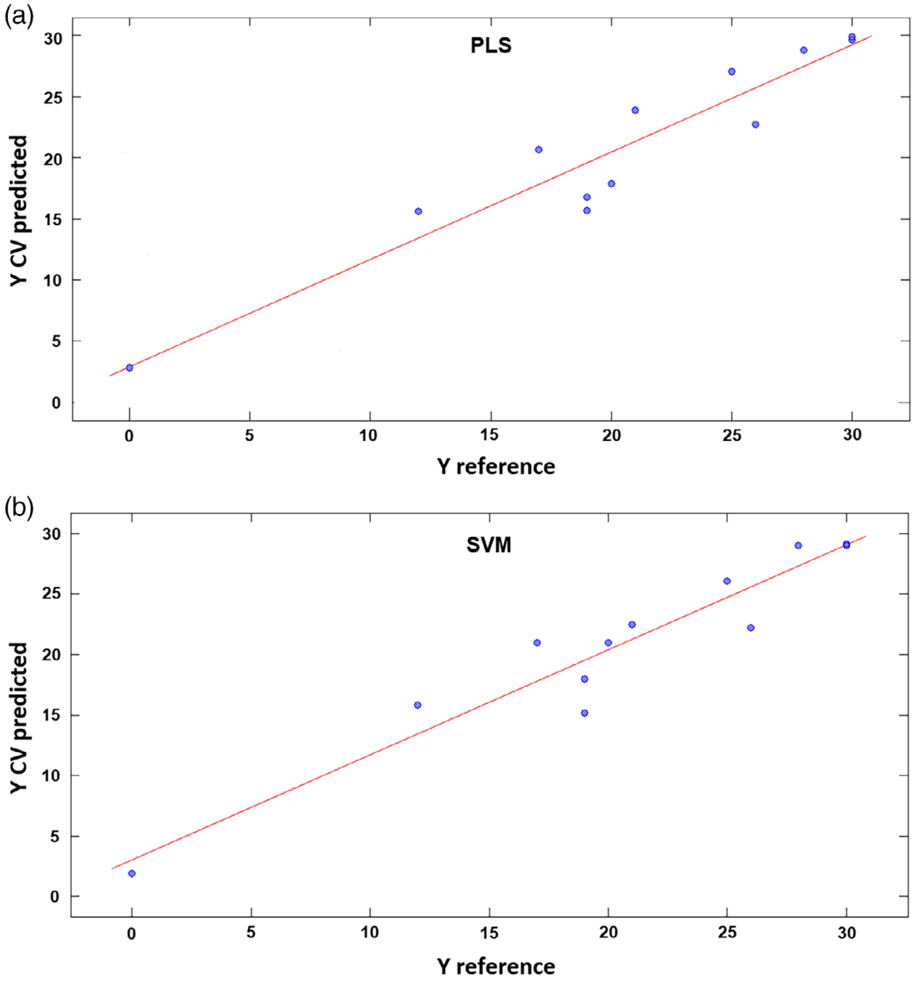

Fig. 4 Predicted versus reference plot to estimate the conversion of epoxide to cyclic carbonate: (a) PLS model and (b) SVM model

From the regression coefficient (Fig. 5), it is possible to observe the influence of halide on catalyst efficiency, with the order of anion activity identified as $Br^- > I^- > Cl^-$. The same profile was found by Langanke et al. (2013), relating this order as a result of a balance between the nucleophilicity and leaving character of the chemical species. Moreover, the solvent effect of supercritical CO_2 and its influence on mass transfer phenomena can play an important role in this order (Cai et al., 2017; Yang et al., 2017; Zheng et al., 2015).

The size of the organic structure and polarizability of the molecule are other important features of the catalysts. These characteristics are considered by the QSPR model through the molecular descriptors apol, bpol, C2SP3, nAtom, nAtomLAC, nBonds2, $SssCH_2$, and VABC, all presenting a positive regression coefficient with respect to the conversion of epoxide to cyclic carbonate. The same pattern has been described in the literature, which relates an increase in the bulkiness of the catalyst to a weakening in electrostatic interactions between cation and anion and an increase in halide nucleophilic character (Dharman et al., 2009; Han et al., 2012; Narang et al., 2016; Wang et al., 2016; Wei et al., 2013a). Thus, it is concluded that catalysts with higher molecular volume and longer saturated carbon chains are more efficient in the production of oleochemical carbonates.

Catalyst solubility plays an important role in the homogeneous phase reaction, but information on catalyst solubility in epoxidized derivatives is limited in the literature (Alves et al., 2017). From the molecular descriptors ALogP and Lipoaffinity Index, we identified that catalyst efficiency increases with its lipophilicity. As the application domain of the QSPR model comprises the synthesis of oleochemical carbonates derived from epoxidized triacylglycerols, the catalyst lipophilicity character may be related to its solubility. Thus, it can be concluded that the solvent effects of the oily matrix and supercritical CO_2 stabilize the charge of bulky organic cations and increase the

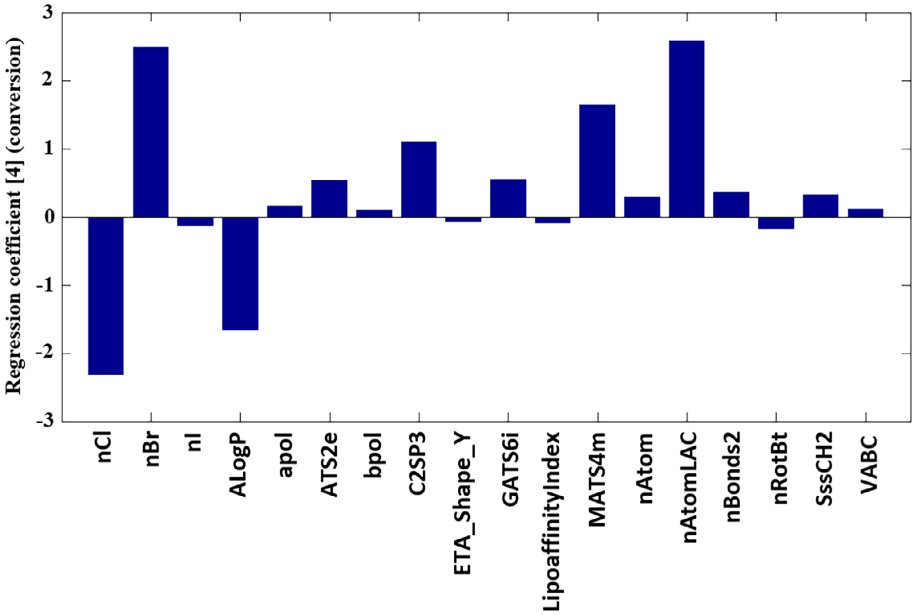

Fig. 5 PLS regression coefficient to estimate the conversion of epoxide to cyclic carbonate

halide anion nucleophilicity, reducing the energy barrier of CO_2 cycloaddition to epoxide (Narang et al., 2016; Sun and Zhang, 2007).

The autocorrelation descriptors (ATS2e, GATS6i, and MATS4m) have a positive PLS regression coefficient and are related to the distribution of properties along the molecular structure. Due to the complexity of these indices, no clear interpretation is possible. Then, after observing the PLS regression coefficients, the variable influence on projection (VIP scores) was analyzed to rank the relative importance of the molecular descriptors for the QSPR model and the projection is presented in Fig. 6 (Das et al., 2017; Wold et al., 2001).

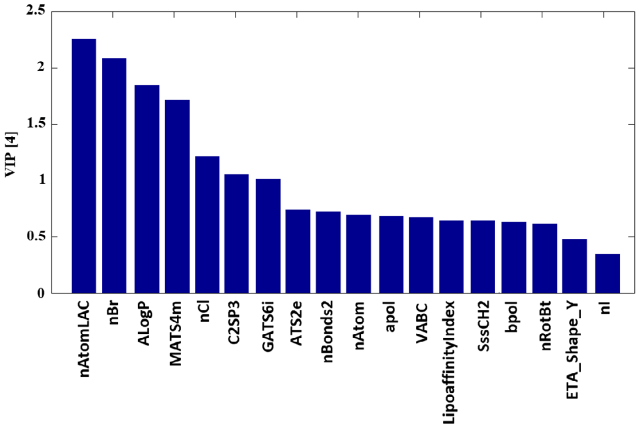

Fig. 6 Variable influence on projection for PLS model

From Fig. 6, it is possible to rank the relative importance of all molecular descriptors applied in the regression, following the order: carbon chain size > halide species > catalyst lipophilicity > distribution of properties along the molecular structure. Thus, within the application domain of the QSPR model, it is concluded that it is possible to establish a structure–property relationship between the characteristics of catalysts and their activity to produce oleochemical carbonates from CO_2 and epoxides. Moreover, the overfitting problem was not identified, being possible to obtain validated QSPR models from 06 molecular descriptors (nCl^-, nBr^-, ALogP, nAtomLAC, SssCH2, and VABC), while the transferability of the QSPR model was confirmed from the modeling based on Data Set 02. The results of the overfitting and transferability assessment are described in the Supporting Information (Pages S5–S16, Figs S1–S4, and Tables S1–S8).

Virtual Screening of Potential Catalysts

Virtual screening is a computational method that guides the search for new active compounds in large chemical libraries (Roy and Mitra, 2011; Scior et al., 2012). In the present work, virtual screening was performed using PubChem search tools, being restricted to the classes of compounds that compose the application domain of the QSPR model (i.e. organic halides derived from: ammonium, phosphonium, imidazolium, pyrrolidinium, and pyridinium).

From virtual screening, a total of 122 potential catalysts were retrieved from the virtual library and their

Table 6 Best molecular targets highlighted by QSPR models to produce oleochemical carbonates

Catalyst	PubChem CID	CAS	Conversion (%)[a]
Hexacosyl(trimethyl)ammonium bromide	23196158	—	71.9
Tetrakis(decyl)ammonium bromide	3014876	14937-42-9	63.0
Docosyl(trimethyl)ammonium bromide	10216960	21396-56-5	63.0
1-Docosyl-3-methylimidazolium bromide	86647477	943834-80-8	61.7
Eicosyltrimethylammonium bromide	23767	7342-61-2	58.5
Tributyl(hexadecyl)ammonium bromide	11420451	6439-67-4	56.4
Tributyl(hexadecyl)phosphonium bromide	84716	14937-45-2	54.9
Octadecyltrimethylammonium bromide	70708	1120-02-1	54.4
Heptadecyl(trimethyl)ammonium bromide	10045219	21424-24-8	51.9
Didodecyl(dimethyl)ammonium bromide	18669	3282-73-3	51.2
1-Butyl-3-hexadecylimidazolium bromide	90220325	937716-18-2	50.6
Hexadecyl-(2-hydroxyethyl)-dimethylammonium bromide	10960220	20317-32-2	50.1
Cetyltrimethylammonium bromide	5974	57-09-0	49.7
Tetraoctylphosphonium bromide	3015167	23906-97-0	47.7
Trimethyl(pentadecyl)ammonium bromide	14611710	21424-22-6	47.5
1-Hexadecyl-3-methylimidazolium bromide	2846928	132361-22-9	47.4
Ethyl-hexadecyl-dimethylammonium bromide	31280	124-03-8	47.2
Trioctyl(propyl)ammonium bromide	90449	24298-17-7	46.1
Tetraheptylammonium bromide	78073	4368-51-8	46.1
1-Methyl-3-pentadecylimidazolium bromide	45045358	349148-74-9	45.3

[a] Mean predicted results by the PLS and SVM models.

identification (Table S13) and respective calculated molecular descriptors (Table S14) can be found in the Supporting Information. To estimate catalyst activity, the descriptors (Table S14) were applied to both multivariate regression models (PLS and SVM) and the mean results predicted by the models (PLS + SVM) are considered as the catalyst output. Thus, the activity of 122 catalysts to promote the carbonation reaction of epoxidized triacylglycerols was estimated within the application domain of the QSPR model (100 °C, 10 MPa CO_2, 5 hours, 1 mol% catalyst load and constant agitation). Table 6 shows the 20 best molecular targets predicted by the QSPR model, while Tables S15 and S16 specify the conversion values estimated by each of the multivariate regression methods (PLS and SVM).

From Table 6, it was observed that only tetraoctylphosphonium bromide (Büttner et al., 2016) and tetraheptylammonium bromide (Langanke et al., 2013) have already been applied to produce oleochemical carbonates, both with high catalytic activity as estimated by QSPR models. Moreover, corroborating our results, other catalysts with molecular structures similar to those listed in Table 6 have been reported to be highly active, such as tetraoctylammonium chloride and tetradodecylammonium chloride for styrene carbonate synthesis and 1-tetradecyl-3-methylimidazolium bromide and trihexyltetradecylphosphonium bromide to produce oleochemical carbonates (Dharman et al., 2009; Langanke et al., 2013; Schäffner et al., 2014).

Comparing the results of Table 6 with the literature, it is observed that the QSPR models reproduce other important features of the catalytic system for cycloaddition of CO_2 to epoxide. Carvalho Rocha et al. (2016) applied CTAB and hexadecyl-(2-hydroxyethyl)-dimethylammonium bromide (HEA16Br) to produce styrene carbonate and, as estimated by the QSPR model, the bifunctional catalyst (HEA16Br) was described as more active. Furthermore, the predicted activity of HEA16Br was found to be higher than that of ethyl-hexadecyl-dimethylammonium bromide. These two compounds differ only in the presence of a hydroxyl substituent at the end of the ethyl chain and, therefore, it is concluded that the QSPR model was able to identify the advantages of a bifunctional catalyst, similar to that reported by several studies (Anthofer et al., 2015; Büttner et al., 2015a, b; Wang et al., 2015).

From virtual screening and the QSPR model application (Table S15), it is possible to compare the results obtained for tetrabutylammonium bromide (standard catalyst to produce oleochemical carbonates) with the highlighted molecular targets. The order of activity found (TBAB < trimethyl (decyl)ammonium bromide < trimethyl(dodecyl)ammo nium bromide < CTAB) is similar to that found by the QSPR model (for details, see Table S15) and follows the increase in halide nucleophilicity, which is induced by carbon chain elongation (Dharman et al., 2009; Han et al., 2012; Narang et al., 2016; Wang et al., 2016; Wei et al., 2013a). From the work of Wei et al. (2013a), conventional TBAB has lower efficiency for producing cyclic carbonate compared to CTAB and other longer carbon chain catalysts (both used as co-catalyst of the zinc-cobalt and double metal cyanide complex). Moreover, there are a few reports that apply CTAB as a catalyst for cyclic carbonate production, most of them using CTAB as a co-catalyst (Bu et al., 2010; Carvalho Rocha et al., 2016; Guo et al., 2014; Ion et al., 2009; Liu et al., 2016; Narang et al., 2016, 2017; Tharun et al., 2012; Wei et al., 2013a, b) and none applied to produce oleochemical carbonate. Thus, considering the availability, cost, and estimated activity of CTAB (Table 6), it can be concluded that CTAB is a good molecular target for the synthesis of oleochemical carbonates.

Data Mining of the QSPR Model

After estimating the activity of the virtual screening catalyst set, data mining was performed to evaluate the structure–property relationship of the catalysts. In this assessment, the methyl group was established as the standard substituent and only one molecular feature was changed at a time (e.g. organic structure, molecular arrangement, carbon chain size, and substituent type) and the results are shown in Fig. 7.

From Fig. 7a, it is possible to identify which organic structure, among the organic structures included in the QSPR model (application domain), presents a higher estimated conversion, and the order of activity found is: pyridinium < imidazolium < pyrrolidinium < phosphonium < ammonium. Identifying that ammonium salts are the most active compounds, the influence of the type of substituent on the conversion was evaluated by simulating the inclusion of the hexyl, cyclohexyl, and aryl groups, and the results are shown in Fig. 7b. Thus, we find that, keeping the same carbon number (06) of the substituent, the catalyst activity increases in the following order: aromatic < cycloaliphatic < linear aliphatic chain. Changes in the catalyst molecular structure (organic structure and type of substituent) result in modifications of all molecular descriptors; however, it can be inferred that catalysts with higher lipophilicity and halide nucleophilic character have higher estimated activity.

In addition to the molecular arrangement in the catalyst structure, the number of carbons of the catalyst is kept constant while its distribution is changed. From Fig. 7c, two profiles are identified: (1) the conversion increases with the lengthening of the carbon chain and these results are related to the higher regression coefficient of the nAtomLAC descriptor and (2) branched substituents results in a slight increase in conversion and this is explained by modifying a set of descriptors (ALogP,

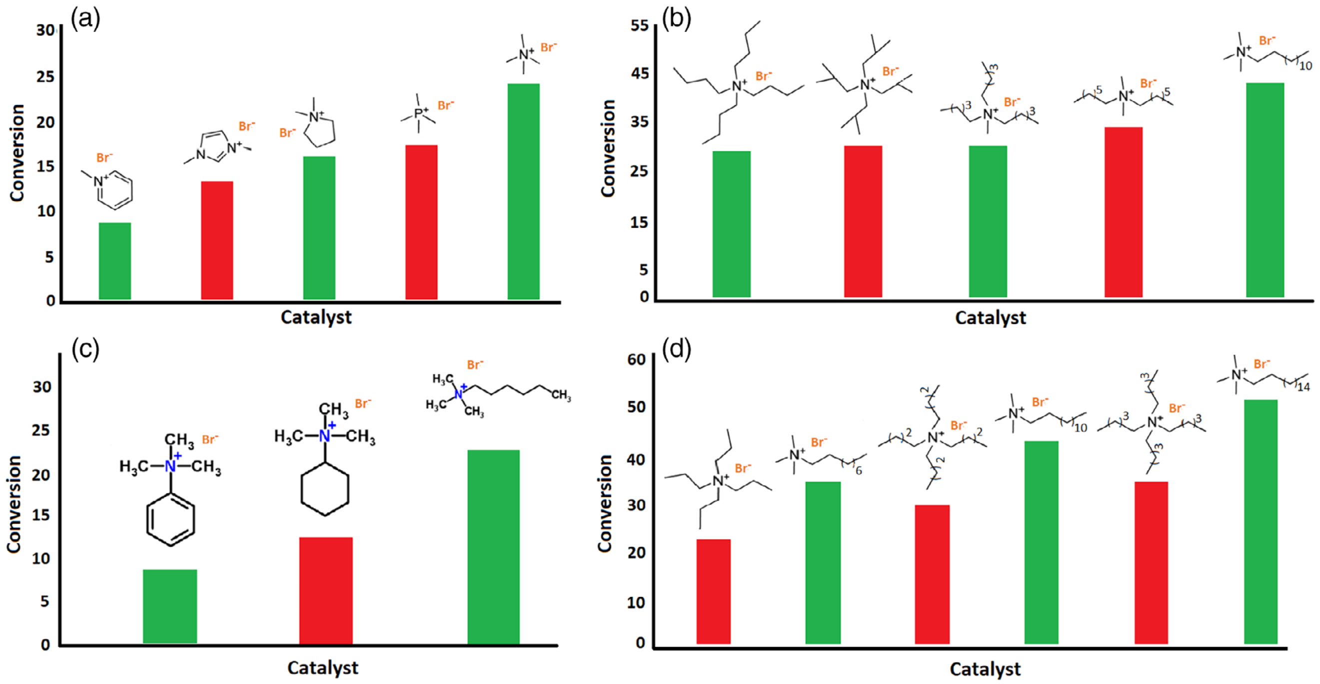

Fig. 7 Structure–property relationship of catalyst to produce oleochemical carbonates: (a) organic structure, (b) substituent type, (c) molecular arrangement, and (d) carbon chain size

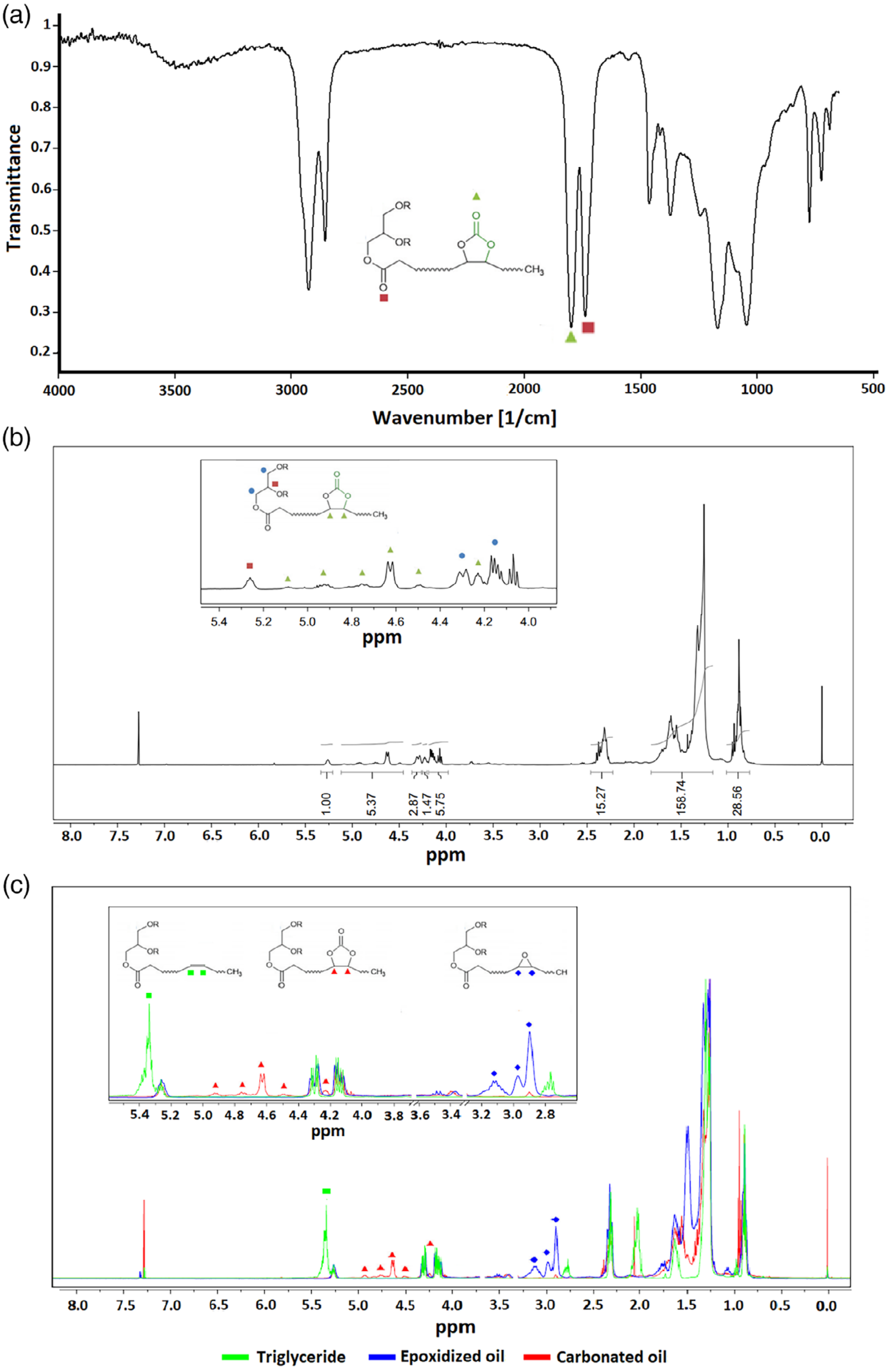

Fig. 8 Infrared and ^{1}H NMR spectra of oleochemical carbonates: (a) FTIR spectra, (b) ^{1}H NMR spectra, and (c) ^{1}H NMR of vegetable oil, epoxidized oil, and carbonated oil (overlapping spectra)

C2SP3, GATS6i, and SssCH$_2$). This difference can be explained mainly by the change in molecular volume, linear carbon chain length, molecular polarizability, and catalyst lipophilicity, indicated by the descriptors (ALogP, apol, bpol, C2SP3, nAtomLAC, SssCH$_2$, and VABC). Thus, it is concluded that both the catalyst solubility and the halide nucleophilic character are strongly influenced by the structure of the substituent. Moreover, from Fig. 7d, it is observed that carbon chain elongation results in an increase in the estimated conversion value, regardless of the carbon arrangement profile in the catalyst. Therefore, the increase in the bulkiness of the catalyst results in the

Table 7 ^{1}H NMR estimated parameters for triacylglycerol derivatives (raw, epoxidized, and carbonated)

Base oil	C=C[a]	Epoxy group[b]	Conversion (η %)[c]
Rice bran oil	4.00	2.29	98.4%
Canola oil	4.40	3.18	>99%
Soybean oil	4.95	3.55	>99%

[a] Mean ethylenic unsaturation per triacylglycerol unit,

[b] Mean epoxy group per triacylglycerol unit after epoxidation,

[c] Conversion estimated based on the initial epoxy value.

increase of the nucleophilic character of the halide (Dharman et al., 2009; Han et al., 2012; Narang et al., 2016; Wang et al., 2016; Wei et al., 2013a).

Oleochemical Carbonate Synthesis

After highlighting the best molecular targets of the virtual screening set and interpreting the catalyst structure–property relationship through QSPR modeling and exploratory analysis (Supporting Information), CTAB was applied as a catalyst to produce cyclic carbonate from CO_2 and bio-based epoxides. Although not the focus of the present work, the synthetic step was included to conduct the proof of concept on the application of computational methods to screen new catalysts for the synthesis of cyclic carbonates from CO_2 and epoxides. Thus, three epoxidized vegetable oils derived from rice bran oil, canola oil, and soybean oil were used to produce oleochemical carbonates. All the vegetable oils (rice, canola, and soybeans) are

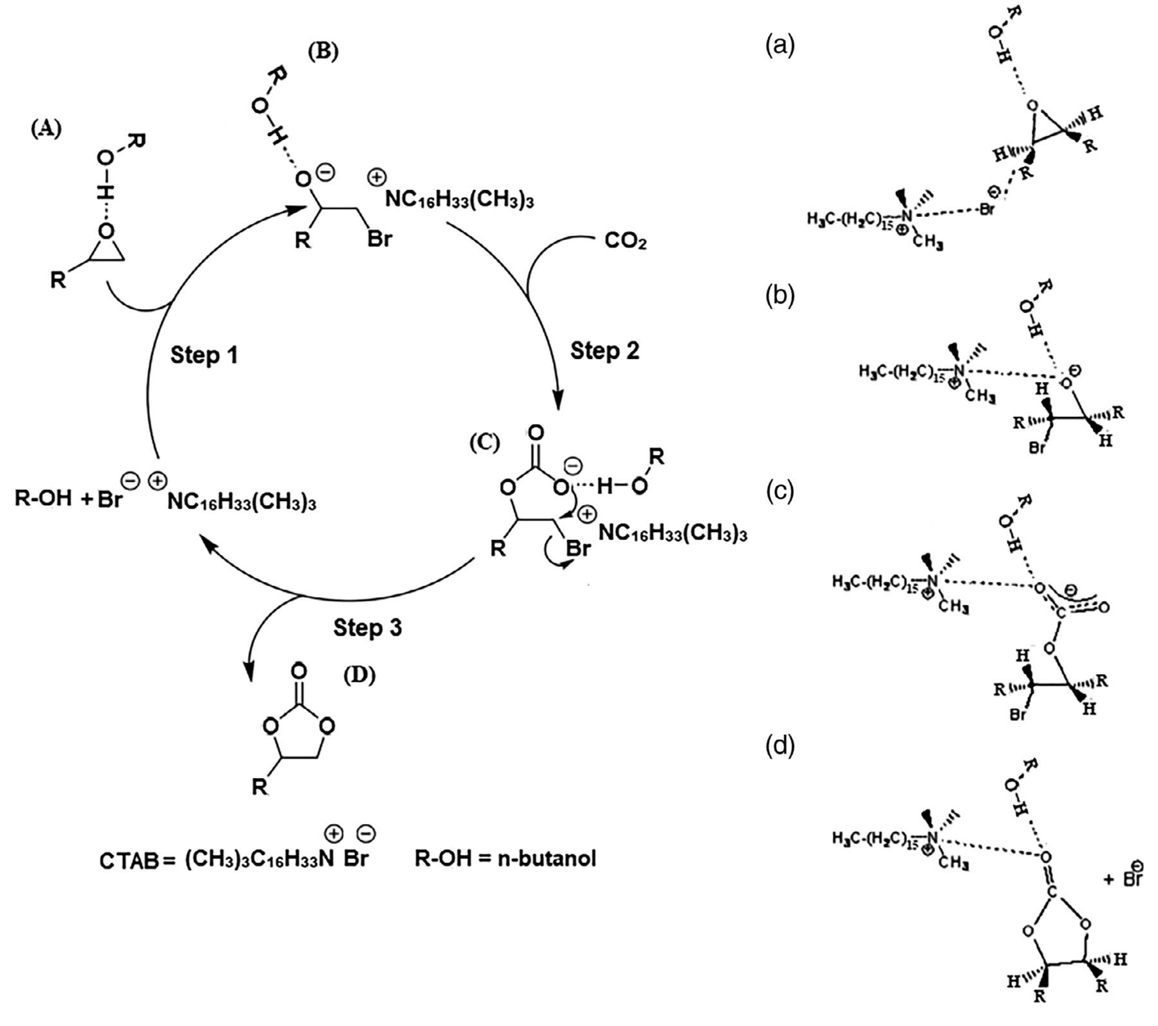

Fig. 9 Proposed reaction mechanism for CTAB-based catalytic system: (a) epoxide ring opening by the nucleophilic attack, (b) stabilization of the oxyanion by CTAB and protic solvent, (c) CO_2 insertion and stabilization of the carbonate ion, and (d) intramolecular formation of the 5-member cyclic carbonate

widely available and were used to evaluate whether fatty acid composition influences conversion efficiency (Longwitz et al., 2018; Peña Carrodeguas et al., 2017).

Initially, the solubility test indicated that CTAB is poorly soluble in epoxidized oil at room temperature, a result similar to those reported by Tamami et al., 2004 for benzyltrimentylammonium bromide, a catalyst with a structure similar to CTAB. Unlike TBAB, which is readily solubilized in oil due to the weaker pair of ions between the bromide and the farthest nitrogen center (Büttner et al., 2017a; Caló et al., 2002), the effective interaction between the CTAB polar head and the bromide ion, together with the low polarity of the medium, hinders its initial solubilization and ion stabilization in the oil matrix. Thus, we suggest that the application of CTAB must be assisted using polar solvent, supercritical CO_2, or phase transfer catalysts.

In the present work, n-butanol (protic solvent) was used to overcome CTAB solubility limitation due to its known miscibility with triacylglycerols and little influence on the conversion of epoxides to cyclic carbonates (Büttner et al., 2015a, b). Moreover, considering the n-butanol structure, it can also act as a hydrogen bond donor (HBD) activator to facilitate epoxy ring opening and modify the mass transfer phenomena involved, reducing medium viscosity and changing CO_2 solubility and diffusion rates (Alves et al., 2016; Cai et al., 2017; Wei et al., 2013b; Zheng et al., 2015). Thus, to conduct the proof of concept and confirm CTAB activity, the synthetic protocol was performed as described in the Materials and Methods section and the products are then characterized by infrared (Fig. 8a) and ^{1}H NMR (Fig. 8b, c) spectroscopy, representatively illustrated in Fig. 8. FTIR (Figs S8–S16) and ^{1}H NMR spectra (Figs S17–S25) with detailed attribution are provided in the Supporting Information.

Infrared characterization (Fig. 8a) was performed to identify the presence of cyclic carbonate in the product. The disappearance of the oxirane band between 842 and 823 cm^{-1} indicates epoxide consumption, while a new intensive carbonyl band (C=O) at 1795 cm^{-1} indicates the formation of 5-membered cyclic carbonate. Then, the ^{1}H NMR analysis (Fig. 8b, c) confirms the initial consumption of ethylenic unsaturation of vegetable oil, multiplet between 5.40 and 5.30 ppm (–**CH**=**CH**–), to produce the epoxy group (–**CH**O**CH**–), two multiplets at 2.9 and 3.1 ppm (Aerts and Jacobs, 2004; Xia et al., 2016). In the second reaction step, the conversion of epoxide to cyclic carbonate is observed from the disappearance of epoxy group signals (2.9 to 3.1 ppm) and the appearance of new signals related to cyclic carbonate protons of 4.19–4.24 ppm and 4.45–-5.12 ppm (Mazo and Rios, 2012; Tenhumberg et al., 2016). After characterization of the products, the conversion of epoxide to cyclic carbonate is calculated based on ^{1}H NMR spectra and the results are presented in Table 7.

From ^{1}H NMR analysis, over 98% conversion of epoxide to cyclic carbonate was estimated for all vegetable oils (rice, canola, and soybean), respectively, 98.4% for rice bran oil and (>99%) for canola oil and soybean oil. These results are above those obtained by the QSPR models, which estimated an average of 49.7% (Table 6) conversion of epoxide to cyclic carbonate within the application domain (100 °C, 10 MPa CO_2, 5 hours, 1 mol% catalyst load, and constant agitation). This difference can be explained by the nonsimilar experimental conditions applied in the present work (120 °C, 5 MPa CO_2, 48 hours, 5 wt% catalyst load, solvent, and static reaction medium) in relation to the application domain of the QSPR model (Data Set 01). Therefore, by increasing reaction time, temperature, catalyst load, and using n-butanol, it was possible to compensate for the static medium, lower gas pressure, and absence of supercritical CO_2 solvent effect and achieve high epoxy group conversions. Although the experimental procedures performed in the present work are not comparable to the application domain of the QSPR model, due to the technical limitations of our reactors, it was possible to conduct a proof of concept on the potential of multivariate models (QSPR and exploratory analysis) to screen catalysts, demonstrate CTAB activity to produce oleochemical carbonates, and meet the proposed objectives. Thus, Table 7 shows that CTAB has high activity to produce oleochemical carbonates regardless of the base raw material, while Fig. 9 shows the proposed mechanism for the CTAB catalytic system.

The proposed catalytic system is composed of three steps and four transition states, which are: *Step 1*: the hydrogen bond-activated epoxy ring is opened by the bromide nucleophilic attack (Fig. 9a), resulting in an oxyanion stabilized by both CTAB polar head and hydrogen bonding of the protic solvent (Fig. 9b) (Carvalho Rocha et al., 2016); *Step 2*: the insertion of CO_2 by oxyanion attack leads to the formation of a carbonate ion (Fig. 9c), also stabilized by hydrogen bonding and electrostatic interaction with the CTAB polar head; *Step 3*: the last step proceeds with the disruption of the C—Br chemical bond and the intramolecular formation of the 5-member cyclic carbonate (Fig. 9d).

Unlike the TBAB catalytic system, in which intermediate species stabilization occurs through weak Van Der Waals interaction between the anions and the alkyl chain (Alves et al., 2016), CTAB could promote better stabilization of intermediate species as a result of more effective electrostatic interaction between the anions and the accessible nitrogen center of the catalyst polar head (Carvalho Rocha et al., 2016). Thus, based on the QSPR regression coefficient and the proposed reaction mechanism, it can be assumed that a good stabilization of intermediates contributes to the efficiency of CTAB as a catalyst to produce oleochemical carbonates.

Conclusions

In the present work, we propose the application of QSPR modeling and exploratory analysis to assist in the selection of catalysts to produce oleochemical carbonates. From our results, it was identified that it is possible to establish a structure–property relationship between the catalyst characteristics and their activity to produce oleochemical carbonates from CO_2 and epoxides. All developed QSPR models were validated, no overfitting problem was identified, and the transferability of the model was confirmed. Therefore, we concluded that the QSPR method have the potential to reduce costs and time to screen new catalysts for the synthesis of oleochemical carbonates from CO_2 and epoxides.

Based on the QSPR model, we found that catalysts with higher bulkiness, lipophilicity, and halide nucleophilic character have higher estimated activity and, within the application domain of the QSPR model, the catalysts with the most potential are ammonium and phosphonium derivatives with long carbon chains. It should also be noted that the results of the present work are expected within the application domain of the respective QSPR models. Thus, changing the substrate (epoxidized oleochemical derivative) and the reaction conditions (e.g. time, temperature, pressure, and catalyst load) must result in different conversions in the carbonation process.

From the synthetic results, we provide the first report on the application of CTAB as a new active catalyst to produce oleochemical carbonates. From experimental procedures, over 98% conversion of epoxide to cyclic carbonate was estimated for all vegetable oils (rice, canola, and soybean), respectively, a conversion of 98.4% for rice bran oil and (>99%) for canola oil and soybean oil. Thus, the high conversion recommends further investigation of the CTAB-based catalytic system to achieve optimization of experimental procedures and conditions. As the experimental procedures of the present work were not optimized, the influence of the experimental factors (pressure, time, temperature, stirring speed, catalyst load, and solvent) on the reaction system should be evaluated in the future.

In the present work, the QSPR models were developed based on small but representative sets of catalysts; however, some relevant aspects may not have been obtained due to a literature restriction. Another limitation is that the application domain of the QSPR models also encompasses the vegetable oil feedstock, so it must be necessary to develop specific models for vegetable oils with very different fatty acid composition. Thus, to avoid pitfalls and misunderstandings of the catalysis process for the production of oleochemical carbonates, we suggest that future works: (1) increase the number and diversity of catalysts included in the QSPR model to increase their robustness and application domain, (2) develop specific models for vegetable oils with very different characteristics, (3) increase the complexity of molecular descriptors (e.g. 3D-descriptors and quantum-chemical descriptors) and evaluate software that can better represent molecular structures not fully connected, such as organic salts and ionic liquids, (4) experimentally confirm the order of catalyst activities suggested by our results (virtual screening and data mining steps), (5) compare the results of the new CTAB catalytic system with those obtained by the TBAB reference catalyst, and (6) assess whether directional electrostatic interaction between the anions (reactive intermediates) and the CTAB polar head can result in stereoselective products. All these studies should be supported by additional experimental procedures and other cheminformatics tools (e.g. Molecular Dynamics and Quantum Mechanics).

Acknowledgments The authors would like to thank the Pontifical Catholic University of Rio Grande do Sul for the infrastructure, Hewlett Packard—HP (PUCRS) for the research scholarships, the Engineering and Materials Technology Graduate Program for the support, and the Laboratory of Molecular Catalysis (UFRGS) for the ^1H NMR analysis.

Conflict of Interest The authors declare that they have no conflict of interest.

References

Achary, P. G. R., Begum, S., Toropova, A. P., & Toropov, A. A. (2016) A quasi-SMILES based QSPR approach towards the prediction of adsorption energy of Ziegler−Natta catalysts for propylene polymerization. *Materials Discovery*, **5**:22–28. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S2352924516300370

Aerts, H. A. J., & Jacobs, P. A. (2004) Epoxide yield determination of oils and fatty acid methyl esters using 1H NMR. *Journal of the American Oil Chemists' Society*, **81**:841–846. Retrieved from. http://link.springer.com/10.1007/s11746-004-0989-1

Alexander, D. L. J., Tropsha, A., & Winkler, D. A. (2015) Beware of R^2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of Chemical Information and Modeling*, **55**:1316–1322. Retrieved from. http://pubs.acs.org/doi/10.1021/acs.jcim.5b00206

Alves, M., Grignard, B., Gennen, S., Detrembleur, C., Jerome, C., & Tassaing, T. (2015) Organocatalytic synthesis of bio-based cyclic carbonates from CO_2 and vegetable oils. *RSC Advances*, **5**: 53629–53636. Retrieved from. http://xlink.rsc.org/?DOI=C5RA10190E

Alves, M., Grignard, B., Mereau, R., Jerome, C., Tassaing, T., & Detrembleur, C. (2017) Organocatalyzed coupling of carbon dioxide with epoxides for the synthesis of cyclic carbonates: Catalyst design and mechanistic studies. *Catalysis Science & Technology*, **7**: 2651–2684. Retrieved from. http://xlink.rsc.org/?DOI=C7CY00438A

Alves, M., Mereau, R., Grignard, B., Detrembleur, C., Jerome, C., & Tassaing, T. (2016) A comprehensive density functional theory study of the key role of fluorination and dual hydrogen bonding in the activation of the epoxide/CO_2 coupling by fluorinated alcohols.

RSC Advances, **6**:36327–36335. Retrieved from. http://xlink.rsc.org/?DOI=C6RA03427F

Anthofer, M. H., Wilhelm, M. E., Cokoja, M., Drees, M., Herrmann, W. A., & Kühn, F. E. (2015) Hydroxy-functionalized imidazolium bromides as catalysts for the cycloaddition of CO_2 and epoxides to cyclic carbonates. *ChemCatChem*, **7**:94–98. Retrieved from. http://doi.wiley.com/10.1002/cctc.201402754

Appel, A. M., Bercaw, J. E., Bocarsly, A. B., Dobbek, H., DuBois, D. L., Dupuis, M., … Waldrop, G. L. (2013) Frontiers, opportunities, and challenges in biochemical and chemical catalysis of CO_2 fixation. *Chemical Reviews*, **113**:6621–6658. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/cr300463y

Aquino, A. S., Bernard, F. L., Borges, J. V., Mafra, L., Vecchia, F. D., Vieira, M. O., … Einloft, S. (2015) Rationalizing the role of the anion in CO_2 capture and conversion using imidazolium-based ionic liquid modified mesoporous silica. *RSC Advances*, **5**:64220–64227. Retrieved from. http://pubs.rsc.org/en/Content/ArticleLanding/2015/RA/C5RA07561K

Assen, V. D. N., Müller, L. J., Steingrube, A., Voll, P., & Bardow, A. (2016) Selecting CO_2 sources for CO_2 utilization by environmental-merit-order curves. *Environmental Science & Technology*, **50**:1093–1101. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/acs.est.5b03474

Begam, B. F., & Kumar, J. S. (2016) Computer assisted QSAR/QSPR approaches—A review. *Indian Journal of Science and Technology*, **9**:1–8. Retrieved from. http://www.indjst.org/index.php/indjst/article/view/87901

Blay, V., Gullón-Soleto, J., Gálvez-Llompart, M., Gálvez, J., & García-Domenech, R. (2016) Biodegradability prediction of fragrant molecules by molecular topology. *ACS Sustainable Chemistry & Engineering*, **4**:4224–4231. Retrieved from. http://pubs.acs.org/doi/10.1021/acssuschemeng.6b00717

Bu, Z., Wang, Z., Yang, L., & Cao, S. (2010) Synthesis of propylene carbonate from carbon dioxide using trans-dichlorotetrapyridineruthenium(II) as catalyst. *Applied Organometallic Chemistry*, **24**:813–816. Retrieved from. http://doi.wiley.com/10.1002/aoc.1708

Büttner, H., Grimmer, C., Steinbauer, J., & Werner, T. (2016) Iron-based binary catalytic system for the valorization of CO_2 into biobased cyclic carbonates. *ACS Sustainable Chemistry & Engineering*, **4**:4805–4814. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/acssuschemeng.6b01092

Büttner, H., Lau, K., Spannenberg, A., & Werner, T. (2015a) Bifunctional one-component catalysts for the addition of carbon dioxide to epoxides. *ChemCatChem*, **7**:459–467. Retrieved from. http://doi.wiley.com/10.1002/cctc.201402816

Büttner, H., Longwitz, L., Steinbauer, J., Wulf, C., Werner, T., Buttner, H., … Werner, T. (2017a) Recent developments in the synthesis of cyclic carbonates from epoxides and CO_2. *Topics in Current Chemistry*, **375**:50. Retrieved from. http://link.springer.com/10.1007/s41061-017-0136-5

Büttner, H., Steinbauer, J., & Werner, T. (2015b) Synthesis of cyclic carbonates from epoxides and carbon dioxide by using bifunctional one-component phosphorus-based organocatalysts. *ChemSusChem*, **8**:2655–2669. Retrieved from. http://doi.wiley.com/10.1002/cssc.201500612

Büttner, H., Steinbauer, J., Wulf, C., Dindaroglu, M., Schmalz, H.-G., & Werner, T. (2017b) Organocatalyzed synthesis of oleochemical carbonates from CO_2 and renewables. *ChemSusChem*, **10**:1076–1079. Retrieved from. http://doi.wiley.com/10.1002/cssc.201601163

Cai, X., Matos, M., & Leveneur, S. (2019) Structure–reactivity: Comparison between the carbonation of epoxidized vegetable oils and the corresponding epoxidized fatty acid methyl Ester. *Industrial and Engineering Chemistry Research*, **58**:1548–1560. Retrieved from. https://pubs.acs.org/doi/10.1021/acs.iecr.8b05510

Cai, X., Zheng, J. L., Wärnå, J., Salmi, T., Taouk, B., & Leveneur, S. (2017) Influence of gas-liquid mass transfer on kinetic modeling: Carbonation of epoxidized vegetable oils. *Chemical Engineering Journal*, **313**:1168–1183. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S1385894716315741

Caló, V., Nacci, A., Monopoli, A., & Fanizzi, A. (2002) Cyclic carbonate formation from carbon dioxide and oxiranes in tetrabutylammonium halides as solvents and catalysts. *Organic Letters*, **4**:2561–2563. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/ol026189w

Carvalho Rocha, C., Onfroy, T., Pilmé, J., Denicourt-Nowicki, A., Roucoux, A., & Launay, F. (2016) Experimental and theoretical evidences of the influence of hydrogen bonding on the catalytic activity of a series of 2-hydroxy substituted quaternary ammonium salts in the styrene oxide/CO_2 coupling reaction. *Journal of Catalysis*, **333**:29–39. Retrieved from. https://doi.org/10.1016/j.jcat.2015.10.014

Cokoja, M., Wilhelm, M. E., Anthofer, M. H., Herrmann, W. A., & Kühn, F. E. (2015) Synthesis of cyclic carbonates from epoxides and carbon dioxide by using organocatalysts. *ChemSusChem*, **8**:2436–2454. Retrieved from. http://doi.wiley.com/10.1002/cssc.201500161

Cruz, V. L., Martinez, S., Martinez-Salazar, J., Polo-Cerón, D., Gómez-Ruiz, S., Fajardo, M., & Prashar, S. (2007) 3D-QSAR study of *ansa*-metallocene catalytic behavior in ethylene polymerization. *Polymer*, **48**:4663–4674.

Cuéllar-Franca, R. M., & Azapagic, A. (2015) Carbon capture, storage and utilisation technologies: A critical analysis and comparison of their life cycle environmental impacts. *Journal of CO2 Utilization*, **9**:82–102. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S2212982014000626

Danov, S. M., Kazantsev, O. A., Esipovich, A. L., Belousov, A. S., Rogozhin, A. E., & Kanakov, E. A. (2017) Recent advances in the field of selective epoxidation of vegetable oils and their derivatives: A review and perspective. *Catalysis Science & Technology*, **7**:3659–3675. Retrieved from. http://xlink.rsc.org/?DOI=C7CY00988G

Das, S., Ojha, P. K., & Roy, K. (2017) Development of a temperature dependent 2D-QSPR model for viscosity of diverse functional ionic liquids. *Journal of Molecular Liquids*, **240**:454–467. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0169743901001551

Desens, W., & Werner, T. (2016) Convergent activation concept for CO_2 fixation in carbonates. *Advanced Synthesis and Catalysis*, **358**:622–630.

Dharman, M. M., Yu, J.-I., Ahn, J.-Y., & Park, D.-W. (2009) Selective production of cyclic carbonate over polycarbonate using a double metal cyanide–quaternary ammonium salt catalyst system. *Green Chemistry*, **11**:1754. Retrieved from. http://xlink.rsc.org/?DOI=b916875n

Dinda, S., Patwardhan, A. V., Goud, V. V., & Pradhan, N. C. (2008) Epoxidation of cottonseed oil by aqueous hydrogen peroxide catalysed by liquid inorganic acids. *Bioresource Technology*, **99**:3737–3744. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0960852407005639

Fayet, G., Raybaud, P., Toulhoat, H., & de Bruin, T. (2009) Iron bis(arylimino)pyridine precursors activated to catalyze ethylene oligomerization as studied by DFT and QSAR approaches. *Journal of Molecular Structure: THEOCHEM*, **903**:100–107. Retrieved from. https://linkinghub.elsevier.com/retrieve/pii/S0166128009000955

Ghose, A. K., & Crippen, G. M. (1986) Atomic physicochemical parameters for three-dimensional structure-directed quantitative structure-activity relationships I. Partition coefficients as a measure of hydrophobicity. *Journal of Computational Chemistry*, **7**:565–577. Retrieved from. http://doi.wiley.com/10.1002/jcc.540070419

Ghose, A. K., & Crippen, G. M. (1987) Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure-activity relationships. 2. Modeling dispersive and hydrophobic interactions. *Journal of Chemical Information and Modeling*, **27**:21–35. Retrieved from. http://pubs.acs.org/cgi-bin/doilookup/?10.1021/ci00053a005

Golbraikh, A., & Tropsha, A. (2002) Beware of Q^2! *Journal of Molecular Graphics & Modelling*, **20**:269–276. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S1093326301001231

Gomes, C. D. N., Jacquet, O., Villiers, C., Thuéry, P., Ephritikhine, M., & Cantat, T. (2012) A diagonal approach to chemical recycling of carbon dioxide: Organocatalytic transformation for the reductive functionalization of CO_2. *Angewandte Chemie International Edition*, **51**:187–190. Retrieved from. http://doi.wiley.com/10.1002/anie.201105516

Gonzalez, M., Teran, C., Saiz-Urra, L., & Teijeira, M. (2008) Variable selection methods in QSAR: An overview. *Current Topics in Medicinal Chemistry*, **8**:1606–1627. Retrieved from. http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1568-0266&volume=8&issue=18&spage=1606

Goud, V. V., Patwardhan, A. V., Dinda, S., & Pradhan, N. C. (2007) Epoxidation of karanja (*Pongamia glabra*) oil catalysed by acidic ion exchange resin. *European Journal of Lipid Science and Technology*, **109**:575–584. Retrieved from. http://doi.wiley.com/10.1002/ejlt.200600298

Gramatica, P., & Sangion, A. (2016) A historical excursus on the statistical validation parameters for QSAR models: A clarification concerning metrics and terminology. *Journal of Chemical Information and Modeling*, **56**:1127–1131. Retrieved from. http://pubs.acs.org/doi/10.1021/acs.jcim.6b00088

Guo, Z., Lin, Q., Wang, X., Yu, C., Zhao, J., Shao, Y., & Peng, T. (2014) Rapid synthesis of nanoscale double metal cyanide catalysts by ball milling for the cycloaddition of CO_2 and propylene oxide. *Materials Letters*, **124**:184–187. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0167577X14004339

Han, L., Choi, S.-J., Park, M.-S., Lee, S.-M., Kim, Y.-J., Kim, M.-I., … Park, D.-W. (2012) Carboxylic acid functionalized imidazolium-based ionic liquids: Efficient catalysts for cycloaddition of CO_2 and epoxides. *Reaction Kinetics, Mechanisms and Catalysis*, **106**:25–35. Retrieved from. http://link.springer.com/10.1007/s11144-011-0399-8

Ion, A., Parvulescu, V., Jacobs, P., & de Vos, D. (2009) Sc and Zn-catalyzed synthesis of cyclic carbonates from CO_2 and epoxides. *Applied Catalysis A: General*, **363**:40–44. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0926860X0900307X

Karelson, M., Lobanov, V. S., & Katritzky, A. R. (1996) Quantum-chemical descriptors in QSAR/QSPR studies. *Chemical Reviews*, **96**:1027–1044. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/cr950202r

Karmakar, G., Ghosh, P., & Sharma, B. (2017) Chemically modifying vegetable oils to prepare green lubricants. *Lubricants*, **5**:44. Retrieved from. http://www.mdpi.com/2075-4442/5/4/44

Katritzky, A. R., Karelson, M., & Lobanov, V. S. (1997) QSPR as a means of predicting and understanding chemical and physical properties in terms of structure. *Pure and Applied Chemistry*, **69**: 245–248. Retrieved from. http://www.degruyter.com/view/j/pac.1997.69.issue-2/pac199769020245/pac199769020245.xml

Katritzky, A. R., & Lobanov, V. S. (1995) QSPR: The correlation and quantitative prediction of chemical and physical properties from structure. *Chemical Society Reviews*, **24**:279–287.

Kim, S. (2016) Getting the most out of PubChem for virtual screening. *Expert Opinion on Drug Discovery*, **11**:843–855. Retrieved from. https://www.tandfonline.com/doi/full/10.1080/17460441.2016.1216967

Kiralj, R., & Ferreira, M. M. C. (2010) Is your QSAR/QSPR descriptor real or trash? *Journal of Chemometrics*, **24**:681–693. Retrieved from. http://doi.wiley.com/10.1002/cem.1331

Kumar R., Bansal V., Patel M. B. and Sarpal A. S. (2012) 1H nuclear magnetic resonance (NMR) determination of the iodine value in biodiesel produced from algal and vegetable oils. *Energy & Fuels*, **26**:7005–7008. Retrieved from. https://doi.org/10.1021/ef300991n

Langanke, J., Greiner, L., & Leitner, W. (2013) Substrate dependent synergetic and antagonistic interaction of ammonium halide and polyoxometalate catalysts in the synthesis of cyclic carbonates from oleochemical epoxides and CO_2. *Green Chemistry*, **15**:1173. Retrieved from. https://doi.org/10.1039/C3GC36710J

Liu, B., Zhang, Y.-Y., Zhang, X.-H., Du, B., & Fan, Z.-Q. (2016) Fixation of carbon dioxide concurrently or in tandem with free radical polymerization for highly transparent polyacrylates with specific UV absorption. *Polymer Chemistry*, **7**:3731–3739. Retrieved from. http://xlink.rsc.org/?DOI=C6PY00525J

Liu, R., Sun, H., & So, S.-S. (2001) Development of quantitative structure−property relationship models for early ADME evaluation in drug discovery. 2. Blood-brain barrier penetration. *Journal of Chemical Information and Computer Sciences*, **41**:1623–1632. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/ci010290i

Liu, S., & Wang, X. (2017) Polymers from carbon dioxide: Polycarbonates, polyurethanes. *Current Opinion in Green and Sustainble Chemistry*, **3**:61–66. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S2452223616300256

Longwitz, L., Steinbauer, J., Spannenberg, A., & Werner, T. (2018) Calcium-based catalytic system for the synthesis of bio-derived cyclic carbonates under mild conditions. *ACS Catalysis*, **8**: 665–672. Retrieved from. http://pubs.acs.org/doi/10.1021/acscatal.7b03367

Maldonado, A. G., & Rothenberg, G. (2010) Predictive modeling in homogeneous catalysis: A tutorial. *Chemical Society Reviews*, **39**: 1891. Retrieved from–1902. http://xlink.rsc.org/?DOI=b921393g

Martínez, S., Cruz, V. L., Ramos, J., & Martínez-Salazar, J. (2012) Polymerization activity prediction of Zirconocene single-site catalysts using 3D quantitative structure–activity relationship modeling. *Organometallics*, **31**:1673–1679. Retrieved from. http://pubs.acs.org/doi/10.1021/om2007776

Mazo, P. C., & Rios, L. A. (2012) Improved synthesis of carbonated vegetable oils using microwaves. *Chemical Engineering Journal*, **210**:333–338. Retrieved from. https://doi.org/10.1016/j.cej.2012.08.099

Miao, S., Wang, P., Su, Z., & Zhang, S. (2014) Vegetable-oil-based polymers as future polymeric biomaterials. *Acta Biomaterialia*, **10**: 1692–1704. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S1742706113004327

Mitra, I., Roy, P. P., Kar, S., Ojha, P. K., & Roy, K. (2010) On further application of rm^2 as a metric for validation of QSAR models. *Journal of Chemometrics*, **24**:22–33. Retrieved from. http://doi.wiley.com/10.1002/cem.1268

Mustafa, A., Lougou, B. G., Shuai, Y., & Wang, Z. (2020) Current technology development for CO_2 utilization into solar fuels and chemicals: A review. *Journal of Energy Chemistry*, **49**:96–123.

Narang, S., Berek, D., Upadhyay, S. N., & Mehta, R. (2016) Effect of electron density on the catalysts for copolymerization of propylene oxide and CO_2. *Journal of Polymer Research*, **23**:96. Retrieved from. http://link.springer.com/10.1007/s10965-016-0994-5

Narang, S., Mehta, R., & Upadhyay, S. N. (2017) Solvent-free cycloaddition of CO_2 and propylene oxide to cyclic carbonates using different ligand metal complexes. *Inorganic and Nano-Metal Chemistry*, **47**:909–916. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0149763417301318

North, M., & Styring, P. (2015) Perspectives and visions on CO_2 capture and utilisation. *Faraday Discussions*, **183**:489–502. Retrieved from. http://xlink.rsc.org/?DOI=C5FD90077H

Peña Carrodeguas, L., Cristòfol, À., Fraile, J. M., Mayoral, J. A., Dorado, V., Herrerías, C. I., & Kleij, A. W. (2017) Fatty acid based biocarbonates: Al-mediated stereoselective preparation of mono-, di- and tricarbonates under mild and solvent-less conditions. *Green Chemistry*, **19**:3535–3541. Retrieved from. http://xlink.rsc.org/?DOI=C7GC01206C

Poliakoff, M., Leitner, W., & Streng, E. S. (2015) The twelve principles of CO_2 CHEMISTRY. *Faraday Discussions*, **183**:9–17. Retrieved from. http://pubs.rsc.org/en/content/articlehtml/2015/fd/c5fd90078f

Pratim Roy, P., Paul, S., Mitra, I., & Roy, K. (2009) On two novel parameters for validation of predictive QSAR models. *Molecules*, **14**:1660–1701. Retrieved from. http://www.mdpi.com/1420-3049/14/5/1660

Ratanasak, M., Rungrotmongkol, T., Saengsawang, O., Hannongbua, S., & Parasuk, V. (2015) Towards the design of new electron donors for Ziegler–Natta catalyzed propylene polymerization using QSPR modeling. *Polymer*, **56**:340–345. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0032386114010489

Rothenberg, G. (2008) Data mining in catalysis: Separating knowledge from garbage. *Catalysis Today*, **137**:2–10.

Roy, K., Ambure, P., & Aher, R. B. (2017) How important is to detect systematic error in predictions and understand statistical applicability domain of QSAR models? *Chemometrics and Intelligent Laboratory Systems*, **162**:44–54. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0169743917300217

Roy, K., Ambure, P., Kar, S., & Ojha, P. K. (2018) Is it possible to improve the quality of predictions from an "intelligent" use of multiple QSAR/QSPR/QSTR models? *Journal of Chemometrics*, **32**:e2992. Retrieved from. https://doi.org/10.1002/cem.2992

Roy, K., & Ghosh, G. (2004) QSTR with extended topochemical atom indices. 2. Fish toxicity of substituted benzenes. *Journal of Chemical Information and Computer Sciences*, **44**:559–567. Retrieved from. http://doi.wiley.com/10.1002/qsar.200330864

Roy, K., & Mitra, I. (2011) On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. *Combinatorial Chemistry & High Throughput Screening*, **14**:450–474. Retrieved from. http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1386-2073&volume=14&issue=6&spage=450

Roy, K., & Mitra, I. (2012) On the use of the metric rm^2 as an effective tool for validation of QSAR models in computational drug design and predictive toxicology. *Mini-Reviews in Medicinal Chemistry*, **12**:491–504. Retrieved from. http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1389-5575&volume=12&issue=6&spage=491

Roy, K., Mitra, I., Kar, S., Ojha, P. K., Das, R. N., & Kabir, H. (2012a) Comparative studies on some metrics for external validation of QSPR models. *Journal of Chemical Information and Modeling*, **52**:396–408.

Roy, K., Mitra, I., Ojha, P. K., Kar, S., Das, R. N., & Kabir, H. (2012b) Introduction of rm^2 (rank) metric incorporating rank-order predictions as an additional tool for validation of QSAR/QSPR models. *Chemometrics and Intelligent Laboratory Systems*, **118**:200–210. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0169743912001402

Samanta, S., Selvakumar, S., Bahr, J., Wickramaratne, D. S., Sibi, M., & Chisholm, B. J. (2016) Synthesis and characterization of polyurethane networks derived from soybean-oil-based cyclic carbonates and bioderivable diamines. *ACS Sustainable Chemistry & Engineering*, **4**:6551–6561. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/acssuschemeng.6b01409

Schäffner, B., Blug, M., Kruse, D., Polyakov, M., Köckritz, A., Martin, A., … Buchholz, S. (2014) Synthesis and application of carbonated fatty acid esters from carbon dioxide including a life cycle analysis. *ChemSusChem*, **7**:1133–1139. Retrieved from. http://doi.wiley.com/10.1002/cssc.201301115

Scior, T., Bender, A., Tresadern, G., Medina-Franco, J. L., Martínez-Mayorga, K., Langer, T., … Agrafiotis, D. K. (2012) Recognizing pitfalls in virtual screening: A critical review. *Journal of Chemical Information and Modeling*, **52**:867–881. Retrieved from. http://pubs.acs.org/doi/10.1021/ci200528d

Stec, M., Spietz, T., Więcław-Solny, L., Tatarczuk, A., & Krótki, A. (2015) Predicting normal densities of amines using quantitative structure-property relationship (QSPR). *SAR and QSAR in Environmental Research*, **26**:893–904. Retrieved from. http://www.tandfonline.com/doi/full/10.1080/1062936X.2015.1095239

Sternberg, A., Jens, C. M., & Bardow, A. (2017) Life cycle assessment of CO_2-based C1-chemicals. *Green Chemistry*, **19**:2244–2259. Retrieved from. http://xlink.rsc.org/?DOI=C6GC02852G

Sun, H., & Zhang, D. (2007) Density functional theory study on the Cycloaddition of carbon dioxide with propylene oxide catalyzed by Alkylmethylimidazolium chlorine ionic liquids. *The Journal of Physical Chemistry A*, **111**:8036–8043. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/jp073873p

Tamami, B., Sohn, S., & Wilkes, G. L. (2004) Incorporation of carbon dioxide into soybean oil and subsequent preparation and studies of nonisocyanate polyurethane networks. *Journal of Applied Polymer Science*, **92**:883–891. Retrieved from. https://doi.org/10.1002/app.20049

Tenhumberg, N., Büttner, H., Schäffner, B., Kruse, D., Blumenstein, M., & Werner, T. (2016) Cooperative catalyst system for the synthesis of oleochemical cyclic carbonates from CO_2 and renewables. *Green Chemistry*, **18**:3775–3788. Retrieved from. http://xlink.rsc.org/?DOI=C6GC00671J

Tharun, J., Dharman, M. M., Hwang, Y., Roshan, R., Park, M. S., & Park, D.-W. (2012) Tuning double metal cyanide catalysts with complexing agents for the selective production of cyclic carbonates over polycarbonates. *Applied Catalysis A: General*, **419–420**:178–184. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0926860X12000397

Todeschini, R., Ballabio, D., & Grisoni, F. (2016) Beware of unreliable Q^2! A comparative study of regression metrics for predictivity assessment of QSAR models. *Journal of Chemical Information and Modeling*, **56**:1905–1913. Retrieved from. http://pubs.acs.org/doi/10.1021/acs.jcim.6b00277

Todeschini, R., & Consonni, V. (2009) In R. Todeschini & V. Consonni (Eds.), *Molecular descriptors for chemoinformatics*. Weinheim, Germany. Retrieved from: Wiley-VCH Verlag GmbH & Co. KGaA. http://doi.wiley.com/10.1002/9783527628766

Tropsha, A. (2010) Best practices for QSAR model development, validation, and exploitation. *Molecular Informatics*, **29**:476–488. Retrieved from. http://doi.wiley.com/10.1002/minf.201000061

Tropsha, A., Gramatica, P., & Gombar, V. (2003) The importance of being Earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR and Combinatorial Science*, **22**:69–77. Retrieved from. http://doi.wiley.com/10.1002/qsar.200390007

Vieira, M. O., Monteiro, W. F., Neto, B. S., Ligabue, R., Chaban, V. V., & Einloft, S. (2018) Surface active ionic liquids as catalyst for CO_2 conversion to propylene carbonate. *Catalysis Letters*, **148**:108–118. Retrieved from. http://link.springer.com/10.1007/s10562-017-2212-4

Wang, J., Zhao, Y., Li, Q., Yin, N., Feng, Y., Kang, M., & Wang, X. (2012b) Pt doped $H_3PW_{12}O_{40}/ZrO_2$ as a heterogeneous and recyclable catalyst for the synthesis of carbonated soybean oil. *Journal* *of Applied Polymer Science*, **124**:4298–4306. Retrieved from. http://doi.wiley.com/10.1002/app.35418

Wang, L., Huang, T., Chen, C., Zhang, J., He, H., & Zhang, S. (2016) Mechanism of hexaalkylguanidinium salt/zinc bromide binary catalysts for the fixation of CO_2 with epoxide: A DFT investigation. *Journal of CO2 Utilization*, **14**:61–66. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S2212982016300129

Wang, L., Jin, X., Li, Y., Li, P., Zhang, J., He, H., & Zhang, S. (2015) Insight into the activity of efficient acid–base bifunctional catalysts for the coupling reaction of CO_2. *Molecular Physics*, **113**: 3524–3530. Retrieved from. http://www.tandfonline.com/doi/full/10.1080/00268976.2015.1037804

Wei, R., Zhang, X., Du, B., Fan, Z., & Qi, G. (2013a) Highly active and selective binary catalyst system for the coupling reaction of CO_2 and hydrous epoxides. *Journal of Molecular Catalysis A: Chemical*, **379**:38–45. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S1381116913002768

Wei, R., Zhang, X., Du, B., Fan, Z., & Qi, G. (2013b) Synthesis of bis(cyclic carbonate) and propylene carbonate via a one-pot coupling reaction of CO_2, bisepoxide and propylene oxide. *RSC Advances*, **3**:17307. Retrieved from. http://xlink.rsc.org/?DOI=c3ra42570c

Wold, S., Sjöström, M., & Eriksson, L. (2001) PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **58**:109–130. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0169743901001551

Xia, W., Budge, S. M., & Lumsden, M. D. (2016) ^{1}H-NMR characterization of epoxides derived from polyunsaturated fatty acids. *Journal of the American Oil Chemists' Society*, **93**:467–478. Retrieved from. http://doi.wiley.com/10.1007/s11746-016-2800-2

Xu, L., & Zhang, W.-J. (2001) Comparison of different methods for variable selection. *Analytica Chimica Acta*, **446**:475–481. Retrieved from. http://www.tandfonline.com/doi/abs/10.1081/AL-120025260

Yang, H., Guo, J., Wen, Y., Ren, T., Wang, L., & Zhang, J. (2017) Solvent effect on the fixation of CO_2 catalyzed by quaternary ammonium-based ionic liquids bearing different numbers of hydroxyl groups: A combined molecular dynamics simulation and ONIOM study. *Molecular Catalysis*, **441**:134–139. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S2468823117304443

Yao, S., Shoji, T., Iwamoto, Y., & Kamei, E. (1999) Consideration of an activity of the metallocene catalyst by using molecular mechanics, molecular dynamics and QSAR. *Computational and Theoretical Polymer Science*, **9**:41–46. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S1089315698000518

Yap, C. W. (2011) PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, **32**:1466–1474. Retrieved from. http://doi.wiley.com/10.1002/jcc.21707

Yu, B., Zou, B., & Hu, C.-W. (2018) Recent applications of polyoxometalates in CO_2 capture and transformation. *Journal of CO2 Utilization*, **26**:314–322. Retrieved from. https://linkinghub.elsevier.com/retrieve/pii/S2212982018301793

Zhao, Y. H., Abraham, M. H., & Zissimos, A. M. (2003) Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *The Journal of Organic Chemistry*, **68**:7368–7373. Retrieved from. http://pubs.acs.org/doi/abs/10.1021/jo034808o

Zheng, J. L., Burel, F., Salmi, T., Taouk, B., & Leveneur, S. (2015) Carbonation of vegetable oils: Influence of mass transfer on reaction kinetics. *Industrial and Engineering Chemistry Research*, **54**: 10935–10944. Retrieved from. http://pubs.acs.org/doi/10.1021/acs.iecr.5b02006

Zheng, J. L., Tolvanen, P., Taouk, B., Eränen, K., Leveneur, S., & Salmi, T. (2018) Synthesis of carbonated vegetable oils: Investigation of microwave effect in a pressurized continuous-flow recycle batch reactor. *Chemical Engineering Research and Design*, **132**: 9–18. Retrieved from. http://linkinghub.elsevier.com/retrieve/pii/S0263876217307104