Chemoinformatics-Driven Design of New Physical Solvents for Selective CO 2 Absorption

Alexey Orlov, Daryna Yu. Demenko, Charles Bignaud, Alain Valtz, Gilles Marcou, Dragos Horvath, Christophe Coquelet, Alexandre Varnek, Frédérick de Meyer

Chemoinformatics-driven design of new physical solvents for selective CO_2 absorption

Alexey A. Orlov[§], Daryna Yu. Demenko[§], Charles Bignaud[†], Alain Valtz[‡], Gilles Marcou[§], Dragos Horvath[§], Christophe Coquelet[‡], Alexandre Varnek[§], Frédérick de Meyer[†‡*].*

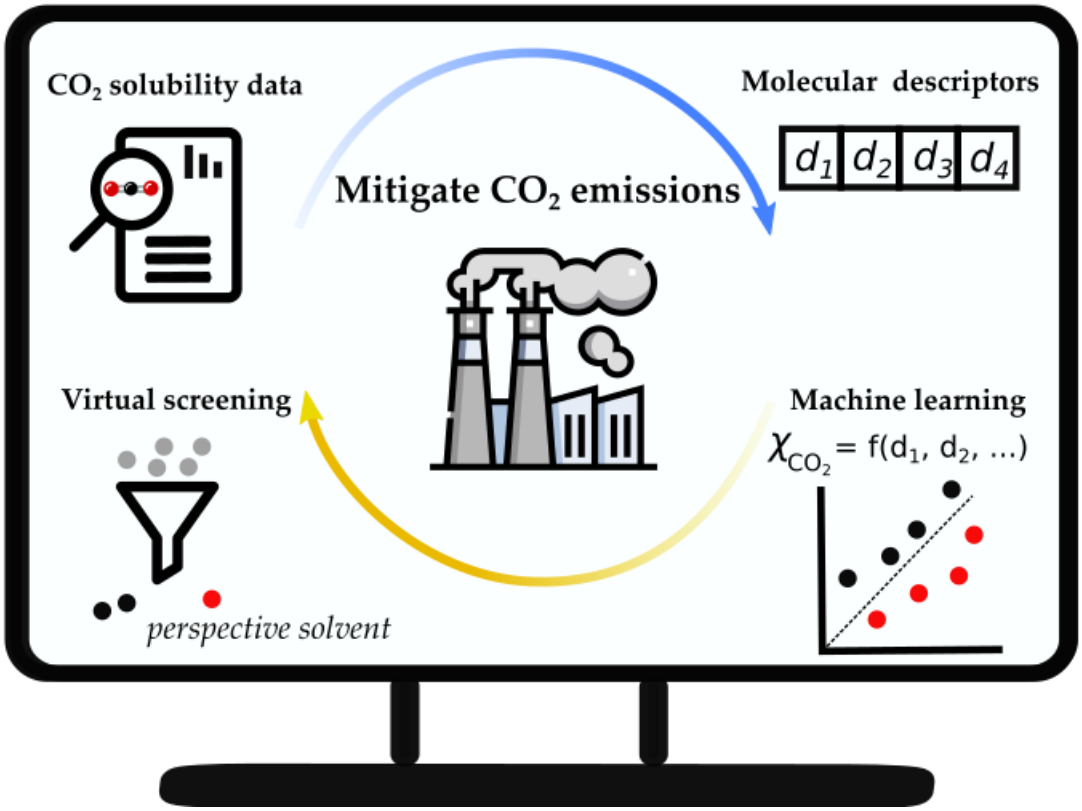
§ Laboratory of Chemoinformatics, Faculty of Chemistry, University of Strasbourg, Strasbourg, 67081 France

† TotalEnergies S.E., Exploration Production, Development and Support to Operations, Liquefied Natural Gas – Acid Gas Entity, CCUS R&D Program, Paris, 92078 France

‡ MINES ParisTech, PSL University, Centre de thermodynamique des procédés (CTP), 35 rue St Honoré, 77300 Fontainebleau, France

Keywords: gas solubility, industrial gases, carbon dioxide, methane, nitrogen, hydrogen, carbon monoxide, chemoinformatics, machine learning

14 **TOC GRAPHIC**



15

ABSTRACT

The removal of CO_2 from gases is an important industrial process in the transition to a low-carbon economy. The use of selective physical (co-)solvents is especially perspective in cases when the amount of CO_2 is large as it enables one to lower the energy requirements for solvent regeneration. However, only a few physical solvents have found industrial application and the design of new ones can pave the way to more efficient gas treatment techniques. Experimental screening of gas solubility is a labor-intensive process, and solubility modeling is a viable strategy to reduce the number of solvents subject to experimental measurements. In this paper, a chemoinformatics-based modeling workflow was applied to build a predictive model for the solubility of CO_2 and four other industrially important gases (CO, CH_4, H_2, N_2). A dataset containing solubilities of gases in 280 solvents was collected from literature sources and supplemented with the new data for six solvents measured in the present study. A modeling workflow based on the usage of several state-of-the-art machine learning algorithms was applied to establish quantitative structure-solubility relationships. The best models were used to perform virtual screening of the industrially produced chemicals. It enabled the identification of compounds with high predicted CO_2 solubility and selectivity towards the other gases. The prediction for one of the compounds − 4-Methylmorpholine was confirmed experimentally.

SYNOPSIS STATEMENT

Developing better solvents for selective CO_2 capture is crucial for reaching net-zero emissions targets.

INTRODUCTION

Global warming due to increasing levels of greenhouse gases CO_2 and CH_4 in the atmosphere has become a major public issue. Several companies and countries have announced ambitious plans to reach net zero CO_2 emissions by 2050. According to the International Energy Agency, Carbon Capture, Utilization or Storage (CCUS) will likely play an important role in achieving this goal.[1] Numerous materials for CO_2 capturing from gases were suggested, including chemical and physical solvents, zeolites, metal oxides, metal-organic frameworks, and membranes.[2,3] The applicability of a certain technology in each case depends on many factors, including the concentrations of CO_2 and of other components in the gas, the pressure of the gas feed, the temperature, etc. In cases, wherein the partial pressure of CO_2 in a gas mixture is sufficiently large, physical solvents represent a perspective alternative to the conventionally used aqueous amines mixtures because of the lower energy requirement. Indeed, a large part of a physical solvent can be regenerated by pressure swing and air stripping, while the regeneration of chemical solvents requires heating and steam stripping[3,4] Pre-combustion CO_2 capture is a key example of a case with a CO_2 partial pressure sufficiently high to use physical solvents. In a pre-combustion process the feed (e.g. coal, natural gas, biomass, etc.) is converted into syngas (H_2 and CO) via gasification, steam reforming, auto thermal reforming or partial oxidation and subsequently the CO is further converted into CO_2 and H_2 via the water gas shift reaction. Typical CO_2 concentrations are in the range of 15 to 60 mol% for a total pressure of 2 to 7 MPa, thus, the CO_2 can be captured with a physical solvent[5]. The other components are mainly H_2, but also CO, N_2, CH_4, H_2O (saturation). The composition strongly depends on the feedstock and on the process. At the moment, only a limited number of physical solvents such as methanol (Rectisol® process), propylene carbonate (Fluor® process), N-acetyl and N-formyl morpholines

(Morphysorb® process), 1-methylpyrrolidin-2-one (Purisol® process), polyethylene glycol ethers (Selexol® process), found application in the industrial CO_2 capture processes. Physical solvents are also often added to chemical solvents (so-called hybrid solvents), for example, to increase the selectivity of absorption towards a specific gas component, to lower the regeneration energy, etc. Examples are sulfolane (Sulfinol® process), thiodiglycol (Hysweet ® process), etc. The search for new physical (co-)solvents is thus an important task.[6]

New suitable solvents should satisfy many criteria, among which are a decent capacity to absorb CO_2, a competitive price, a low volatility (to avoid solvent losses), a low viscosity, etc. Another very important criterium is the selectivity towards CO_2 which should be high enough to obtain a CO_2 stream of acceptable purity for re-utilization or storage, and to avoid losses of valuable chemicals like CH_4. Depending on the source of the CO_2, the selectivity criteria are different. For the removal of CO_2 from natural gas the co-absorption of mainly CH_4, but also of N_2 and H_2O and eventually H_2S should be limited. For a steam methane reformer (SMR), which in a near future is likely to play a key role in the massive production of blue hydrogen from natural gas, the absorption of CO_2 with a high selectivity towards H_2 is important[7,8], but also towards CO, H_2O, N_2 and CH_4[7,8]. The CO_2 present in flue gas from boilers should be removed with a high selectivity towards N_2, but the co-absorption of water and SO_x, NO_x should be low too.

Experimental screening of gas solubility is a time and labor-intensive process, and solubility modeling is a viable strategy to reduce the costs of the required experiments. There were numerous approaches suggested for modeling gas solubility in pure physical solvents. In the work by Pirig et al.[9] a five-parameter linear equation based on the experimentally measured properties was used to model the solubility in 58 solvents. In the work by Li et al.[10], artificial

neural networks were used to model CO_2 mole fraction solubility in 11 solvents (alcohols, ethers, ketones) at different temperatures and pressures. Several structural features (number of C-H, O-H, C-O, C=O bonds, number of rotatable bonds, etc.) and physical properties of the compounds (density, dipole moment, etc.) were used as descriptors. It enabled achieving high precision of predictions for certain types of physical solvents under varying experimental conditions. Nonetheless, the major disadvantage of the modeling approaches based on experimentally measured parameters − the limited number of compounds for which the parameters are available, complicates their usage for large-scale virtual screening of solvent candidates.

Alternative strategies for modeling CO_2 solubility, that require less preliminary knowledge of the experimental properties, were also suggested. In the works of Li et al.[11] and Shi et al.[12] molecular simulations were used to predict CO_2 solubility in nine and twenty-seven physical solvents respectively. Although high predictive accuracy was achieved in both works, molecular simulations are also not very convenient for the large-scale virtual screening, as they are time-consuming and require significant computational resources. Alternatively, the conductor-like screening model for real solvents (COSMO-RS),[13] a method combining quantum chemical calculations with statistical thermodynamics, was suggested to rapidly screen large sets of various materials. In the work by Kim et al.[14] CO_2 and CH_4 Henry's law coefficients were predicted by COSMO-RS for 63 common liquid solvents and 10 ionic liquids at 300 K. Unfortunately, there was no comparison of the predicted solubilities with the available experimental data.

A pool of gas solubility data accumulated in scientific literature supports applying machine learning for quantitative structure-solubility relationships (QSPR) modeling. In this approach chemical structures of compounds are encoded as vectors of molecular descriptors and machine

learning algorithms are then applied for modeling the property of interest. As compared to the modeling based on the usage of experimentally determined parameters or resource-intensive molecular simulations, this method allows efficient screening of large numbers of compounds. To our knowledge, there was only one work related to QSPR modeling of CO_2 solubility in physical solvents published.[15] In the paper of Gorji et al.,[15] Henry coefficients for 22 solvents composed only of carbon, oxygen, and hydrogen elements at different temperatures were used to build a multiple linear regression model with Dragon[16] descriptors. Although good predictive performance was achieved, the applicability domain (AD) of this model is limited to the specific classes of compounds used for model building. Except for recent publications on H_2S solubility modeling[17,18], to our knowledge there were no papers describing the application of the chemoinformatics-driven methods for physical solubility modeling of the major components encountered in natural gas treatment or in the pre-combustion CO_2 capture process : carbon dioxide (CO_2), methane (CH_4), carbon monoxide (CO), hydrogen (H_2), and nitrogen (N_2). Hence, the investigation of the perceptiveness of using chemoinformatics for the rational design of new solvents for the absorption of CO_2 and other industrial gases is an important task.

MATERIALS AND METHODS

Data collection and preprocessing

A dataset containing mole fraction solubility values (χ) for 280 liquid solvents at 298.15 K and 1 atm was collected (Table S1, Supporting Information) from IUPAC reports[19–23], scientific literature[24–75], and patents[76,77]. The mole fraction solubility for a binary (gas-liquid) system is defined[19] as:

$$\chi = \frac{n(g)}{n(g)+n(l)} \qquad (1)$$

n(g) – an amount of substance in a gas phase, n(l) – an amount of substance in a liquid phase.

Median values were then taken for the solvents associated to several reliable measurements of χ at the particular temperature. The final dataset used for modeling is present in Table S1 (Supporting Information). Extrapolation and interpolation of the data to 298.15 K were performed assuming a linear variation of Henry's coefficients with temperature or by the equations suggested in the IUPAC report or corresponding papers. The compounds that are structural outliers with respect to the training set majority, comprising water, carbon disulfide, octamethylcyclotetrasiloxane and hydrazines were not included to the dataset. These compounds contain rare or unique fragments significantly affecting their gas-absorbing properties, and, hence, confident predictions cannot be obtained for them by statistical modeling.

The collected mole fraction solubilities were converted to the Kuenen coefficients S using the following formula[78]:

$$S = \frac{R \times T \times P}{Mw} \times \frac{\chi}{1-\chi} \qquad (2)$$

S – Kuenen coefficient (m^3kg^{-1}), R – ideal gas constant (8.314 $m^3PaK^{-1}mol^{-1}$), T and P – standard temperature and pressure (273.15 K and 101.325 kPa), Mw – molecular weight of compound ($kgmol^{-1}$), χ – mole fraction solubility value.

The Kuenen coefficient is the volume of saturated gas reduced at 273.15 K and 1 atm pressure, which is dissolved by unit mass of pure solvent at the temperature of measurement and partial pressure of 1 atm. This parameter is widely used in industrial applications, as it enables one to

directly estimate the efficiency of the particular solvent related to its cost and dimensions of the required industrial unit (design-capital expenses cost CAPEX). Here, Kuenen coefficients were used for the data analysis and models interpretation.

The selectivity index SI was calculated using the following formula:

$$SI = \frac{\chi_{CO2}}{\chi_{gas}} \quad (3)$$

SI – selectivity index, experimental or predicted χ_{CO2} and χ_{gas} – mole fraction solubilities of CO_2 and other gases respectively.

All χ values were also transformed to a logarithmic scale, i.e. the negative value of the decimal logarithm was taken (Figure S1).

Modeling

Standardization

All compound structures were standardized using in-house standardization procedures based on KNIME,[79] which included aromatization, stereochemistry depletion, etc.

Descriptors

193 different ISIDA fragment descriptor sets were generated using the Fragmentor17 software.[80,81] ISIDA fragments represent either sequences (the shortest topological paths with an explicit representation of all atoms and bonds), atom-centered fragments (all connected atoms to a certain topological distance), or triplets (all the possible combinations of 3 atoms in a graph with the topological distance between each pair indicated). The number of fragments in each set

varied from 30-40 (for short sequences of atoms/bonds) to 400-1200 (for long sequences up to 6 atoms) for different gases.

Quantum chemical descriptors resulted from DFT calculations in the gas phase, with model wB97X-D 6-31G* performed with the Spartan 18.0 program[82]. Default QSAR descriptors available in Spartan including energy, dipole moment, E_{HOMO}, and E_{LUMO} were calculated.

Machine learning algorithms

Random forest (RF): RF algorithm[83] implemented in sci-kit learn library (v. 0.22.1)[84,85] was used. The following hyperparameters were tuning during optimization (grid search): number of trees (100, 300, 1000), number of features (all features, one-third of all features, $\log_2$ of the number of features), the maximum depth of the tree (5, 10, full tree), bootstrapping (with and without the usage of bootstrap samples for building the tree).

XGBoost (XGB): XGBoost algorithm[86] as implemented in XGBoost python module (v.1.2.0)[87] was used. The following hyperparameters were tuning during optimization (grid search): number of trees (50, 100, 300, 500), number of features (all features, 70% of all features), number of samples (all samples, 70% of all samples), the maximum depth of the tree (3, 5, 10), learning rate (0.3, 0.1, 0.5, 0.05), the minimum sum of instance weight needed in a node (1, 5, 10). All other parameters were left as default.

Support vector regression (SVR): SVR algorithm[88] implemented in sci-kit learn library (v. 0.22.1), was used. The descriptors were scaled to the [0,1] range before applying the algorithm. The following hyperparameters were tuning during optimization (grid search): kernel (linear, rbf,

poly, sigmoid), kernel coefficient (1, 0.1, 0.01, 0.001, 0.0001), regularization parameter (0.1, 1, 10, 100, 1000).

Model validation workflow

The modeling workflow was implemented using sci-kit learn library (v. 0.22.1) in python 3.7 scripting language. Identical modeling workflows were used for solubility modeling (expressed as -lg χ) of all gases. At the first stage of the modeling, a machine learning algorithm: RF, SVR and XGB were tested in 5-fold cross-validation, which was repeated 5 times (Figure S2). For each descriptor set, the model's measures of performance were calculated and several models with a coefficient of determination $Q^2_{CV} \geq 0.7$ were selected for consensus modeling.

The following equations were used to calculate the measures of the model's performance in cross-validation:

$$Q^2_{CV} = \frac{\sum_{j=1}^{5}(1-\frac{\sum_{i=1}^{n}(y_{i,exp}- y_{i,pred})^2}{\sum_{i=1}^{n}(y_{i,exp}- \bar{y})^2})}{5} \quad (4)$$

$$RMSE_{CV} = \frac{\sum_{j=1}^{5}\sqrt{\sum_{i=1}^{n}\frac{(y_{i,exp}- y_{i,pred})^2}{n}}}{5} \quad (5)$$

$$MAE_{CV} = \frac{\sum_{j=1}^{5}\sum_{i=1}^{n}\frac{|y_{i,exp}- y_{i,pred}|}{n}}{5} \quad (6)$$

Above, n is the number of compounds in the entire learning set, $y_{i,exp}$, $y_{i,pred}$ experimental and values predicted in 5-fold cross-validation for compound i from the learning set, j is the index of the repetition of the 5-fold cross-validation procedure. For each measure of the model's performance, the standard deviation over 5 repetitions was calculated.

Each of the selected models was then associated with an Applicability Domain (AD), defined as a bounding box.[89] Hence, the pool of selected models extracted from the given data set was used as a consensus predictor, returning for each input solvent candidate a mean value of solubility estimates and its standard deviation, taken over the predictions returned by each model in the pool, if the compounds appeared outside AD of all the models, or, alternatively, over the predictions returned by only those models having the candidate within their AD.

Outlying data points were defined as the data points for which absolute errors ($|\chi exp - \chi pred|$) from cross-validation were larger than $2 \times RMSE_{CV}$ threshold.

Y-randomization test

The absence of chance correlation was checked through the Y-randomization procedure. Y-randomization test was performed in the following way: $-lg\chi$ values (y values) were shuffled, surrogate models from the cross-validation were built using shuffled values and the values from the corresponding cross-validation test set were calculated. This procedure was repeated 100 times for each fold and the maximum values of the coefficient of determination were compared with the coefficient of determination obtained for the original $-lg\chi$ values.

Virtual screening

An in-house dataset comprising 4,082 industrially produced compounds and their structural analogs was screened in the following way. Only structures containing the same atoms (C, H, N, O, S, P, halogens) as in the learning set were kept. All structures were standardized and ISIDA descriptors were calculated for them as described above. Individual ISIDA models refitted to the entire dataset with the hyperparameters selected in the cross-validation were used

to compose the final consensus model as described above. Then, predictions were made using the ISIDA consensus model. Only compounds that were inside the applicability domain defined as bounding box for at least three ISIDA fragment types were considered.

Software implementation

The developed model was implemented into the ISIDA-Predictor software.[81]

Experimental measurement of CO_2 solubility

A "static-synthetic" technique based on a closed-circuit method[90–92] was used for the determination of CO_2 solubility in the solvents. In this method, which is explained in detail in the supplementary information (Text S1, Figures S3, S4), the system pressure is measured at constant temperature for different overall compositions. To determine the global compositions, the quantities of pure substances charged into the stirred equilibrium cell, which is evacuated and placed in a thermostatic liquid bath, need to be known precisely. The purified and degassed solvents are charged into the cell as compressed liquids using thermostatted piston injectors. Then, the gas is added stepwise as a liquefied gas using the same injection pumps or as a gaseous component using a thermo-regulated gas bomb. Knowing the pressure, temperature, and volume of the gas bomb, the amount of gas inside the bomb can be calculated using correlated PvT data of the gas. Thus, the injected amount of gas can be obtained from the pressure difference in the bomb before and after each injection.

Since only temperature, pressure, total loadings, gas-liquid interface level and total volumes are measured, the compositions of the coexisting phases need to be determined by the evaluation of the raw data. From the known amount of solvent, the liquid phase volume is determined using

precise information about the density of the liquid solution inside the equilibrium chamber. From the total volume of the cell, the remaining gas phase volume can be calculated precisely (see supplementary information). At given equilibrium conditions (temperature, gas phase volume, and gas pressure) the amounts of gas in the gas phase and thus, also in the liquid phase are obtained. In this approach, several effects influence the resulting liquid phase compositions. These effects are the small amounts of solvents in the gas phase, the compressibility of the solvent under the gas pressure, the partial molar volume of the dissolved gas and the solvent activity coefficient. All effects are considered in an isothermal and isochoric algorithm by solving the mass and volume balances.

The partial pressure is obtained during the iterative procedure:

$$P_{gas} = P_{sys} - P_{solvent} \qquad (6)$$

where Pgas – partial pressure of the acid gas in the system, Psys – total pressure in the system, Psolvent – partial pressure of a solvent vapour. This equation is valid at low pressure and in the absence of chemical reactions in the gas phase. The method to calculate the uncertainty of the measured experimental data is explained in the supplementary information. The uncertainty of the measured CO_2 solubility at 1 atm. is equal or lower than 1%.

RESULTS AND DISCUSSION

Data collection, preprocessing and analysis

IUPAC reports on gas solubilities in non-aqueous solvents contain to our knowledge the most complete and carefully analyzed publicly available data on gas solubility. The data from these reports were used to compose the “cores” of our datasets. As the largest number of data points

for various solvents was available at 298.15 K and 1 atm, the data at this temperature and pressure was chosen for modeling. Since the mole fraction values can vary significantly depending on the experimental methods being used, we chose only the data points which were considered as the most reliable by IUPAC's or Total's experts. Data from recent publications either at 298.15 K or obtained by extrapolation or interpolation of the data measured at close temperatures were also added to the dataset.

Besides the data collected from IUPAC reports and literature, data points for six compounds, hexametapol (HMPA), 1,3-Dimethylimidazolidin-2-one (DMI), thiodiglycol (TDG), and three tertiary amines 2-[2-hydroxyethyl(methyl)amino]ethanol (MDEA), 2-[ethyl(2-hydroxyethyl)amino]ethanol (EDEA), and 2-[2-(diethylamino)ethoxy]ethanol (DEAE-EO) were measured experimentally and added to the dataset (Figure 1a; Table S1). The choice of the solvents was motivated by their wide application in industrial processes and the absence of consistent data at 298.15 K for them in the literature. TDG is employed in a commercial mixed chemical/physical solvent formulation for sour gas treating (HySWEET technology) developed by TotalEnergies S.E.[93] HMPA and DMI are being used as solvents for gases, polymers, and in organic synthesis. Aqueous amines are used as chemical solvents, and only little is known about the physical solubility of gases in pure amines. One of the few examples is MDEA, which, in an aqueous solution, is commonly used for industrial gas treatment, and for which the CO_2 mole fraction solubility can be estimated from Skylogianni.[94] The value is extremely high (~0.04 at 1 atm and 313K) as compared to other physical solvents (see below). Hence, considering the growing interest in water-lean solvents[95], including the ones based on pure amines[96], we have chosen three industrial amines (MDEA, EDEA, DEAE-EO) for the experimental assessment of CO_2 physical solubility.

All the solvents showed close to linear variation of mole fraction solubility vs partial pressure of the gas in the pressure range 0-2 atm (Figure 1a). Estimated mole fraction solubilities for TDG and DMI (0.0041; 0.0150) are in good agreement with the data that can be obtained by extrapolation from recent publications.[27,97] On the contrary, CO_2 mole fraction solubility in HMPA (0.024) is lower, than the one suggested by IUPAC's expert[19] (0.031), but is close to the one obtained by Schay et al.[19] (0.028). The mole fractions values in MDEA (0.1) and EDEA (0.08) are remarkably high and in agreement with the aforementioned data for MDEA from the Skylogianni et al. obtained at higher temperatures.[94] At the same time, the CO_2 solubility in DEAE-EO is much lower (0.02) and is the same as in two other tertiary amines present in the dataset: triethylamine (0.02) and perfluorotributylamine (0.02).

The high CO_2 solubility in pure MDEA and EDEA triggers the question whether the absorption is purely physical. It is commonly reported in the literature that, contrary to primary and secondary amines, tertiary amines cannot chemically absorb CO_2 in the absence of water[98,99] (we have verified that no water was present in the solvent in the experiments performed in the present paper). This view has been challenged by Maddox[100] and more recently by Heldebrandt et al.[101,102] who studied the reaction of CO_2 with anhydrous tertiary amines. Both conclude that reversible Lewis acid-base adducts are formed at high pressure (note that in this work we compare solubilities at low pressure, 1 atm.). Heldebrandt[101] suggests that the difference in absorption capacity between pure amines can be explained by solvent polarity effects. Either way, anhydrous tertiary amines do absorb less CO_2 than aqueous tertiary amines, but some still absorb significant amounts of CO_2.

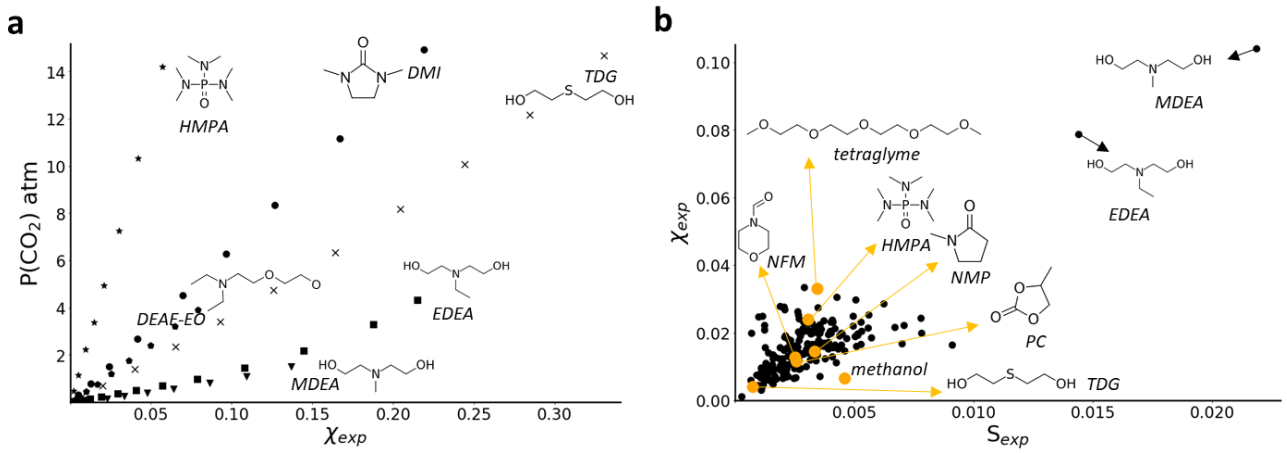


Figure 1. (a) Variation of mole fraction with partial pressure for CO_2 in TDG (×), DMI (★), HMPA (●), MDEA (▼), EDEA (■), and DEAE-EO (□) at 298.15 K experimentally measured for this paper. (b) Plot of experimental molecular fraction values (χ_{exp}) vs Kuenen coefficients (S_{exp}) at 298.15 K and 1 atm for CO_2.

There were 211 mole fraction solubility values collected for CO_2. The largest mole fraction CO_2 solubility was for tertiary amines MDEA and EDEA (Figure 1b). Among other classes of compounds with large CO_2 solubility were phosphoric acid esters, long chain ethers, and esters. To estimate the efficiency of a solvent related to its cost and dimensions of the required industrial unit, mole fractions were converted to Kuenen coefficients. MDEA and EDEA also have the largest Kuenen coefficients. By contrast to the trend observed for mole fractions, the largest Kuenen coefficient values in other compound classes were for small polar compounds: nitriles, ketones (acetone, butan-2-one), tetrahydrofuran (THF). Notably, the solvents, which are used in industrial gas treatment processes are not among the best ones in terms of CO_2 solubility (Figure 1b).

Other gases are less studied as compared to CO_2. There were less than 105 mole fraction solubility values collected for each of other gases (N_2, H_2, CO, CH_4). It is worth noting, that polar CO and non-polar CH_4, N_2, H_2 showed similar solubility trends. The largest values were

for non-polar compounds, including perfluorated alkanes, and long-chain n-alkanes (Figure S5). The minimal χ values were for polar solvents, such as methanol, *N*,*N*-dimethylformamide (DMF).

The collected experimental data were used to analyze the trends in the selectivity of CO_2 absorption towards other gases (Figure S6). The only industrial solvent for which the data were available for all the solvents is methanol, which is not selective at 298.15 K and 1 atm. The industrially used solvents were among the best ones in terms of the CO_2/CH_4 selectivity. The highest selectivity index (SIexp) was for dimethyl sulfoxide (DMSO, SIexp=24) and N-formyl morpholine (NFM, SIexp=21). Among other most selective solvents were industrially used propylene carbonate (PC) and N-Methyl-2-pyrrolidone (NMP). The large selectivity stems from the extremely low solubility of CH_4 in these solvents. The same observation was made for all other gases: the most selective are the polar solvents, such as DMF, 1,4-dioxane, DMSO, etc. (Figure 1b, Figures S3-S4).

Quantitative structure solubility relationships

Application of the machine learning allowed one to establish quantitative structure-solubility relationships. Reasonable predictive accuracy was achieved in the repeated cross-validation procedure for all the gases (Table 1). None of the models has shown chance correlation in the y-scrambling procedure.

Table 1. Performance estimation for modeling of mole fraction solubility expressed as -lgχ.

Gas	ISIDA consensus models	Q^2_{CV}	$RMSE_{CV}$	MAE_{CV}
CO_2	20 RF, 17 XGBoost	0.71±0.01	0.12±0.01	0.08±0.01

CH_4	9 SVR, 15 XGBoost	0.77±0.02	0.15±0.01	0.10±0.01
CO	5 SVR, 8 XGBoost	0.78±0.03	0.12±0.01	0.09±0.01
H_2	2 RF, 3 SVR, 23 XGBoost	0.77±0.04	0.15±0.01	0.10±0.01
N_2	9 SVR, 12 XGBoost	0.75±0.06	0.17±0.01	0.11±0.01

347

348 Since the collected datasets are small, the presence of compounds containing rare fragments, or
349 compounds with noise in the experimental data lead to unstable modeling results. Several
350 compounds, which were systematically mispredicted (the absolute error >0.7 log units) in the
351 cross-validation procedure were removed: dodecanal and dodecene for the H_2 model, and
352 dimethyl ether and dodecanal for the CO model. The values for all these compounds are
353 significantly different from their close structural analogs. They were obtained by interpolation
354 and additional experimental confirmation is required to assess whether the values are reliable.
355 After the removal of outliers, the models with reasonable figures of merit were obtained for each
356 gas (Table 1). The lowest mean absolute error (MAE_{CV}) was for the CO_2 model, which is based
357 on the largest pool of data. Note, that MAE_{CV} is close to the variance in the experimental data.
358 For example, the standard deviation for propylene carbonate based on IUPAC's data[19] and the
359 recently published data[36] can be estimated as 0.05 log units.

360 To check if some other descriptor types can lead to significantly better results, we calculated
361 quantum chemical descriptors using Spartan software. The results of modeling were on average
362 comparable to those obtained by the usage of ISIDA fragments (Table S2). Considering the
363 advantages of ISIDA fragments, i.e. speed of calculation and intuitive interpretation of structure-
364 property relationships, we further focused on this descriptor type.

For each model, the compounds for which absolute errors were larger than $2\times RMSE_{CV}$ threshold were analyzed (Figure 2, Table S3). These compounds either contain rare fragments or can be considered as "solubility cliffs": small changes in structure (e.g. replacement of hydrogen atom by methyl group) lead to large changes in solubility (see Figure S7 and discussion below). For example, hexafluorobenzene – the only polyhalogenated aromatic compound and MDEA – one of few representatives of alkanolamines in the dataset were among the compounds with the largest errors for the CO_2 model. The datasets for the gases are rather chemically diverse. Many compounds containing rare fragments appear outside AD of the models in the cross-validation. There were 15% of compounds appearing to be outside AD in the cross-validation for CO_2, while about 20-25% of compounds were outside AD for other gases. The presence of compounds with rare fragments leads to high variance of predictions in the cross-validation. From the learning curve (Figure S8), one can see that adding data improves the performance on the validation sets, and, therefore, decreases the gap between prediction accuracy on validation and training sets. Hence, further accumulation of the experimental data on gas solubility organic solvents is required for building more robust models with enlarged applicability domains and the extended range of temperature and pressure values.

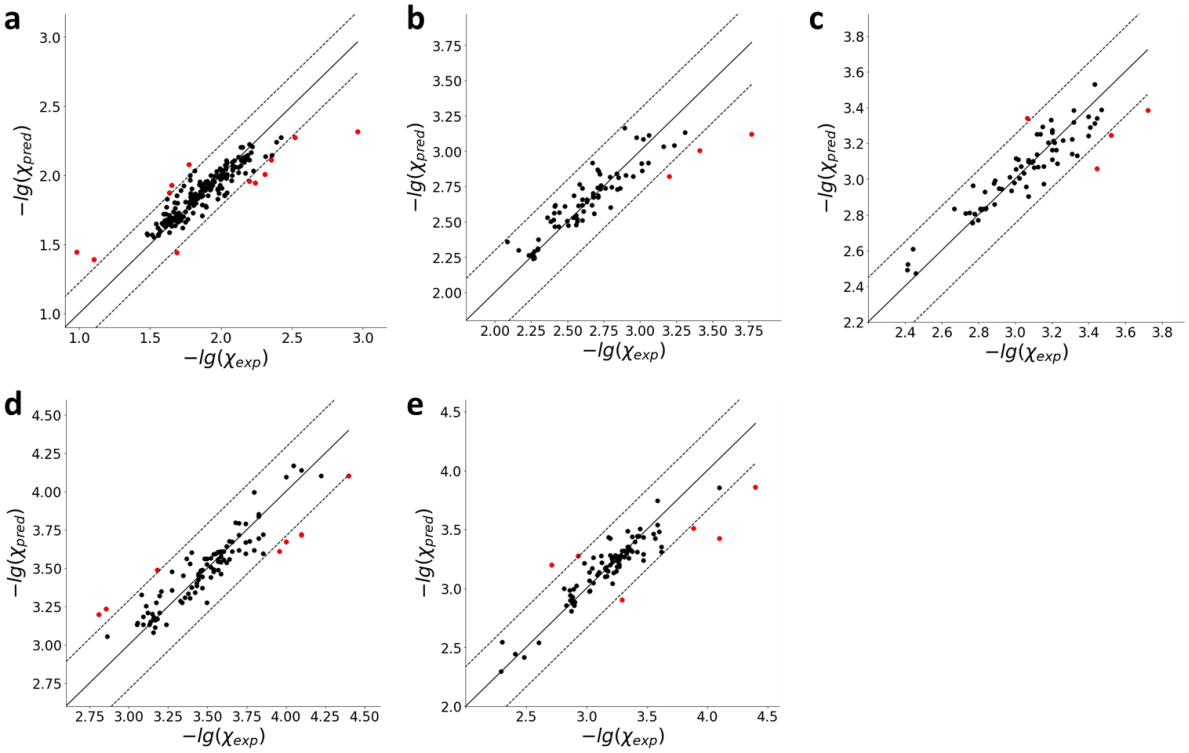


Figure 2. Plot of predicted ($-lg\chi_{pred}$) vs experimental ($-lg\chi_{exp}$) values for ISIDA consensus model in cross-validation procedure for CO_2 (a), CH_4 (b), CO (c), H_2 (d), N_2 (e). The predicted values are calculated as an average of 5 folds. Compounds for which absolute errors were larger than $2\times RMSE_{CV}$ are shown in red. Dash lines indicate $\pm 2\times RMSE_{CV}$ threshold.

In contrast to the above examples, solubility of CO_2 in alcohols, glycols and ethers was systematically studied. Yet, one of the largest absolute errors were for glycols (glycerol, ethane-1,2-diol), which is related to a sharp change in solubility with the replacement of -OH group to -OCH_3 (Figure S7). For example, the mole fraction solubility in glycerol is more than three-time smaller, than the solubility in its closest structural analog − propylene glycol (PG). At the same time solubility in another structural analog − diethylene glycol (DEG), containing the same number of carbon and oxygen atoms as glycerol, is six time higher. This phenomenon can be

394 explained by considering forces driving the process of gas dissolution. The mechanistic
395 interpretation of this process assumes the formation of a cavity capable of accommodating a gas
396 molecule by breaking solvent-solvent bonds and introduction and fixation of a gas molecule in
397 this cavity due to gas-solvent interactions. Hence, solubility of gases in liquids depends upon two
398 types of interactions: gas-solvent and solvent-solvent.[103,104] Strong gas-solvent and weak solvent-
399 solvent interactions lead to greater solubility. In line with that, CO_2 solubility in glycols and their
400 ethers follows the cohesive energy density values trend: the solubility is increasing from glycerol
401 to DEG with the decreasing cohesive energy density (glycerol: 1142 MPa; DEG: 615 MPa).[105]
402 However, the cohesive energy density is not the only factor affecting the solubility of CO_2.
403 Although carbon dioxide is nonpolar, its appreciable polarizability and ability to accept hydrogen
404 bonds from suitable donor solvents[106] makes structure solubility landscape more complex. For
405 example, CO_2 solubility is lower in hexane than in dimethyl ether of ethylene glycol (DMEG),
406 while the cohesive energy density of DMEG is higher than that of hexane (DMEG: 317 MPa;
407 hexane: 222 MPa)[105,107]. Thus, the interplay between cohesive energy density and solvent-CO_2
408 interactions should be taken into account in the process of the design of new solvents.

409 *Virtual screening*

410 To find new solvents with high CO_2 solubility and high selectivity towards other gases, we
411 performed the virtual screening of the in-house library of industrially produced chemicals and
412 their close structural analogs comprising more than 4,000 chemicals (Figure S9). It is worth
413 noting that the experimentally measured physico-chemical properties such as melting and boiling
414 points, density, flash points, etc. were available only for a small fraction of the dataset and thus,
415 we did not check if the compounds possess plausible values of properties at 298.15 K and 1 atm.
416 Most of the screened compounds (87%) appeared to be inside AD of the CO_2 model. There were

numerous compounds found with high predicted CO_2 mole fractions values and Kuenen coefficients, several of which were superior to the existing industrially used solvents (Figure 3a). Among the best CO_2 solvents according to mole fraction solubilities were tertiary amines and long-chain esters (e.g., dioctyl adipate, $\chi_{pred} = 0.27$), while the largest Kuenen coefficients were for tertiary amines and the close structural analogs of the compounds with the largest Kuenen coefficients from the learning set: ethers (e.g. ethyl methyl ether, $S_{pred} = 0.0068$), ketones (e.g. methoxyacetone, $S_{pred} = 0.0053$) and nitriles (e.g. butyronitrile, $S_{pred} = 0.0052$). One of the tertiary amines with the largest Kuenen coefficient ($S_{pred} = 0.0054$) − 4-Methylmorpholine (NMM) was selected for the experimental measurement of solubility. NMM showed linear variation of mole fraction solubility vs partial pressure of the gas in the pressure range indicating pure physical solubility (Figure 3b). The experimental mole fraction solubility (0.26) matched the predicted one (0.24) well. The CO_2 solubility in NMM is appreciably higher than in industrially used NFM, which instead of a tertiary amine group contains an amide group. We have measured the dynamic viscosity of NMM: 0,92 cP at 20 °C. The dynamic viscosity of water at 20 °C is 1 cP. The viscosity of NMM is thus comparable to water. NMM is much less viscous than e.g. pure MDEA (100 cP at 20 °C) or pure EDEA (90 cP at 20 °C), which is a significant advantage. On the other hand, the boiling point of NMM is 116 °C (MDEA 243 °C). NMM is thus more volatile than MDEA. The NMM solvent should thus be used at a lower temperature, to minimize the solvent losses. This is e.g. also done in the Rectisol process which uses methanol (boiling point 65 °C).

The environment, health and safety (EHS) of solvents for CO_2 capture is a potential issue. In principle, amine emissions should not be an obstacle because the causes are well known and counter-measures can be put in place (operating temperature and pressure, water wash, Brownian

demister, reclaiming units, etc.). One clear advantage of using physical solvents is that there is much less thermal and oxidative solvent degradation because the regeneration is not thermal and the high operating pressure avoids oxygen ingress. According to the safety datasheet[108] NMM is flammable, corrosive and harmful, but the substance contains no components considered to be either persistent, bioaccumulative and toxic.

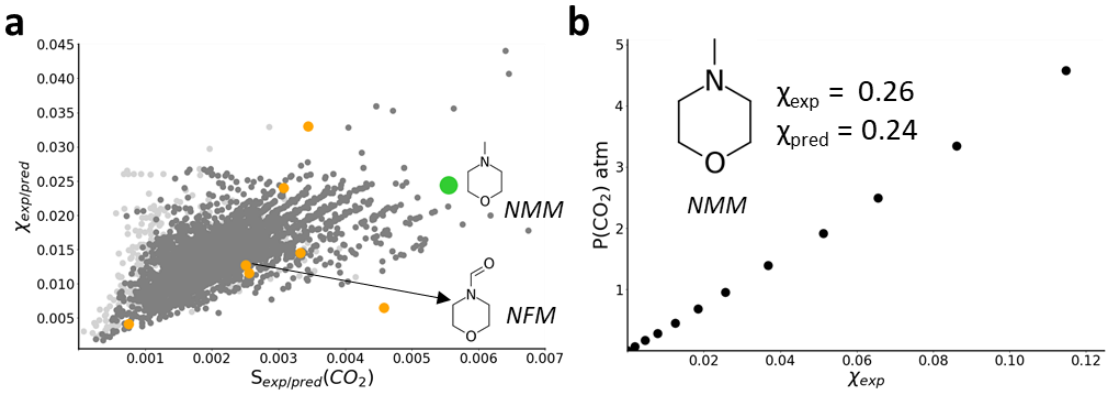


Figure 3. (**a**) A plot of CO_2 molecular fraction values vs Kuenen coefficients. Experimental values – black and orange (solvents used in the industry). **(b)** Variation of mole fraction with partial pressure for CO_2 in NMM at 298.15 K experimentally measured in this paper; χ_{exp} – experimental mole fraction value at 1 atm and 298.15 K, χ_{pred} – predicted value. Predicted values for compounds inside AD – grey, outside AD – light grey, for NMM – green. The position of the NFM on the plot is shown for comparison.

The selectivity of the NMM calculated from predicted mole fraction values is comparable to the one of industrially used solvents (Figure S10). Other tertiary amines were also among the most selective solvents. For example, the 3-(Dimethylamino)-1,2-propanediol was among the most selective solvents for all gases. Among other classes of solvents with the highest selectivity

456 indexes were cyclic amides (e.g. 5-(hydroxymethyl)-1-methylpyrrolidin-2-one) and ketones (e.g.
457 methoxyacetone).

458 To conclude, the rational approach to the design of new physical solvents based on the usage
459 of machine learning for modeling of structure-solubility relationships was suggested in this
460 paper. The collected data on solubility of gases were used to build QSPR models, which were
461 then applied to identify compounds potentially superior to the existing ones via virtual screening
462 of industrially produced chemicals. We have identified pure tertiary amines with a remarkable
463 CO_2 absorption capacity. Previously, the team of Heldebrandt et al.[101] has investigated the use of
464 pure, anhydrous amines for high pressure CO_2 absorption. They compared the performance of
465 anhydrous EDEA to the Fluor solvent (propylene carbonate), to Selexol, and to aqueous MDEA
466 for a representative absorber. Despite the attractiveness due to lower energy consumptions, the
467 use of anhydrous or water-lean amines faces numerous challenges, for example, their lower
468 absorption capacity and their higher viscosity.[6] In this work we have focused on the gas
469 solubility. A further extension of the chemoinformatics workflow for the prediction of other
470 industrial important solvent properties might be very useful in the identification of the most
471 suitable physical (co-)solvent (optimal absorption properties, selectivity, viscosity,[109] EHS
472 impact,[110] etc.) for a given application.

473

474 ASSOCIATED CONTENT

475 **Supporting Information**.

476

AUTHOR INFORMATION

Corresponding Author

*Professor Alexandre Varnek. Laboratory of Chemoinformatics, Faculty of Chemistry, University of Strasbourg, 4, Blaise Pascal Str., 67081, Strasbourg, France. email: varnek@unistra.

*Doctor Frédérick de Meyer. TotalEnergies S.E., Exploration Production, Development and Support to Operations, Liquefied Natural Gas – Acid Gas Entity, CCUS R&D Program, Paris, 92078 France. email: frederick.de-meyer@totalenergies.com.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by the Carbon Capture Utilization and Storage (CCUS) transverse R&D program from TotalEnergies S.E.

ACKNOWLEDGMENT

The authors are grateful to Dr. Fanny Bonachera for her help with the implementation of models to the Predictor software. The authors sincerely acknowledge the Laboratory for Thermophysical Properties in Oldenburg, Germany, for performing a part of the experimental work.

ABBREVIATIONS

Caroxin D – 1,1,2,2,3,3,4,4-octafluoro-1,4-bis(1,1,1,2,3,3,3-heptafluoropropan-2-yloxy)butane

Caroxin F – 1,1,1,2,2,3,3,4,4,5,5,6,6-tridecafluoro-6-(1,1,1,2,3,3,3-heptafluoropropan-2-yloxy)hexane

DEG – 2-(2-hydroxyethoxy)ethanol (diethylene glycol)

DEGM – 2-(2-Methoxyethoxy)ethan-1-ol (diethylene glycol monomethyl ether)

diglyme – 1-methoxy-2-(2-methoxyethoxy)ethane

DMF – N,N-dimethylformamide

DMI – 1,3-Dimethylimidazolidin-2-one

DMSO – methylsulfinylmethane (dimethyl sulfoxide)

EG – ethane-1,2-diol (ethylene glycol)

glycerol – propane-1,2,3-triol

HMPA – N-[bis(dimethylamino)phosphoryl]-N-methylmethanamine (hexametapol)

M2CA – methyl 2-cyanoacetate

MDEA – 2-[2-hydroxyethyl(methyl)amino]ethanol 2-[ethyl(2-hydroxyethyl)amino]ethanol

EDEA – 2-[ethyl(2-hydroxyethyl)amino]ethanol

DEAE-EO – 2-[2-(diethylamino)ethoxy]ethanol

methoxyacetone – 1-methoxypropan-2-one

NMM – 4-methylmorpholine

NFM – morpholine-4-carbaldehyde (N-formylmorpholine)

NMP – 1-methylpyrrolidin-2-one

PC – 4-methyl-1,3-dioxolan-2-one (propylene carbonate)

TDG – 2-(2-hydroxyethylsulfanyl)ethanol (thiodiglycol)

pentaglyme – 1-methoxy-2-[2-[2-[2-(2-methoxyethoxy)ethoxy]ethoxy]ethoxy]ethane

perflubron – 1-bromo-1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-heptadecafluorooctane

perfluoroheptane – 1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,7-hexadecafluoroheptane

perfluoro(methylcyclohexane) – 1,1,2,2,3,3,4,4,5,5,6-undecafluoro-6-(trifluoromethyl)cyclohexane

perfluoroctane – 1,1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,8-octadecafluorooctane

perfluorotributylamine – 1,1,2,2,3,3,4,4,4-nonafluoro-N,N-bis(1,1,2,2,3,3,4,4,4-nonafluorobutyl)butan-1-amine

THF – oxolane (tetrahydrofuran)

TPrP – tripropyl phosphate

χ – mole fraction solubility

S – Kuenen coefficient

SI – Kuenen coefficients selectivity index

squalane – 2,6,10,15,19,23-hexamethyltetracosane

551

552 REFERENCES

553 (1) Net Zero by 2050, A Roadmap for the Global Energy Sector, IEA Report, May 2021.
554 https://www.iea.org/reports/Net-Zero-by-2050 (Accessed 17.06.2021).

555 (2) Sifat, N. S.; Haseli, Y. A Critical Review of CO2 Capture Technologies and Prospects for
556 Clean Power Generation. *Energies* 2019, *12* (21), 4143.
557 https://doi.org/10.3390/en12214143.

558 (3) Abdulsalam, J.; Mulopo, J.; Amosa, M. K.; Bada, S.; Falcon, R.; Oboirien, B. O. Towards a
559 Cleaner Natural Gas Production: Recent Developments on Purification Technologies. *Sep.*
560 *Sci. Technol.* 2019, *54* (15), 2461–2497. https://doi.org/10.1080/01496395.2018.1547761.

561 (4) N.Borhani, T.; Wang, M. Role of Solvents in CO2 Capture Processes: The Review of
562 Selection and Design Methods. *Renew. Sustain. Energy Rev.* 2019, *114*, 109299.
563 https://doi.org/10.1016/j.rser.2019.109299.

564 (5) Wang, X.; Song, C. Carbon Capture From Flue Gas and the Atmosphere: A Perspective.
565 *Front. Energy Res.* 2020, *8*, 560849. https://doi.org/10.3389/fenrg.2020.560849.

566 (6) Wanderley, R. R.; Pinto, D. D. D.; Knuutila, H. K. From Hybrid Solvents to Water-Lean
567 Solvents – A Critical and Historical Review. *Sep. Purif. Technol.* 2021, *260*, 118193.
568 https://doi.org/10.1016/j.seppur.2020.118193.

569 (7) Collodi, G.; Azzaro, G.; Ferrari, N.; Santos, S. Techno-Economic Evaluation of Deploying
570 CCS in SMR Based Merchant H2 Production with NG as Feedstock and Fuel. *Energy*
571 *Procedia* 2017, *114*, 2690–2712. https://doi.org/10.1016/j.egypro.2017.03.1533.

572 (8) Yan, Y.; Thanganadar, D.; Clough, P. T.; Mukherjee, S.; Patchigolla, K.; Manovic, V.;
573 Anthony, E. J. Process Simulations of Blue Hydrogen Production by Upgraded Sorption

574 Enhanced Steam Methane Reforming (SE-SMR) Processes. *Energy Convers. Manag.* 2020,
575 *222*, 113144. https://doi.org/10.1016/j.enconman.2020.113144.
576 (9) Pirig N. Ya.; Polyuzhin I. V.; Makitra R. G. Carbon Dioxide Solubility. *Russ. J. Appl.*
577 *Chem.* 1993, *4* (66), 691–695.
578 (10) Li, H.; Yan, D.; Zhang, Z.; Lichtfouse, E. Prediction of CO2 Absorption by Physical
579 Solvents Using a Chemoinformatics-Based Machine Learning Model. *Environ. Chem. Lett.*
580 2019, *17* (3), 1397–1404. https://doi.org/10.1007/s10311-019-00874-0.
581 (11) Li, H.; Tang, Z.; He, Z.; Gui, X.; Cui, L.; Mao, X. Structure-Activity Relationship for CO2
582 Absorbent. *Energy* 2020, *197*, 117166. https://doi.org/10.1016/j.energy.2020.117166.
583 (12) Shi, W.; Thompson, R. L.; Macala, M. K.; Resnik, K.; Steckel, J. A.; Siefert, N. S.;
584 Hopkinson, D. P. Molecular Simulations of CO_2 and H_2 Solubility, CO_2 Diffusivity, and
585 Solvent Viscosity at 298 K for 27 Commercially Available Physical Solvents. *J. Chem.*
586 *Eng. Data* 2019, *64* (9), 3682–3692. https://doi.org/10.1021/acs.jced.8b01228.
587 (13) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the
588 Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* 1995, *99* (7), 2224–2235.
589 https://doi.org/10.1021/j100007a062.
590 (14) Kim, J.; Maiti, A.; Lin, L.-C.; Stolaroff, J. K.; Smit, B.; Aines, R. D. New Materials for
591 Methane Capture from Dilute and Medium-Concentration Sources. *Nat. Commun.* 2013, *4*
592 (1), 1694. https://doi.org/10.1038/ncomms2697.
593 (15) Gorji, A. E.; Gorji, Z. E.; Riahi, S. Quantitative Structure-Property Relationship (QSPR) for
594 Prediction of CO2 Henry's Law Constant in Some Physical Solvents with Consideration of
595 Temperature Effects. *Korean J. Chem. Eng.* 2017, *34* (5), 1405–1415.
596 https://doi.org/10.1007/s11814-017-0018-0.

(16) *Kode Srl, Dragon (Software for Molecular Descriptor Calculation) Version 7.0.8, 2017, https://chm.kode-solutions.net*.

(17) Orlov, A. A.; Marcou, G.; Horvath, D.; Cabodevilla, A. E.; Varnek, A.; Meyer, F. de. Computer-Aided Design of New Physical Solvents for Hydrogen Sulfide Absorption. *Ind. Eng. Chem. Res.* 2021, *60* (23), 8588–8596. https://doi.org/10.1021/acs.iecr.0c05923.

(18) H. Rostami; Riahi, S. Quantitative Structure–Property Relationship Study on Solubility of Hydrogen Sulfide in Organic Solvent; Kish, Iran, 2014.

(19) Carbon Dioxide in Non-Aqueous Solvents at Pressures Less than 200 KPa. In IUPAC Solubility Data Series (Volume 50); Fogg, P. G. T.; Ed.; Pergamon: Amsterdam, 61988; Pp. 1-483, 1992.

(20) Methane. In IUPAC Solubility Data Series (Volume 27/28); Clever, H. L.; Young, C. L.; Eds.; Pergamon: Amsterdam, 61988; Pp 1−783, 1987.

(21) Carbon Monoxide. In IUPAC Solubility Data Series (Volume 43); Cargill, R.W.; Ed; Eds.; Pergamon: Amsterdam, 61988; Pp 1−783, 1990.

(22) Hydrogen and Deuterium. In IUPAC Solubility Data Series (Volume 5/6); Young, C. L.; Ed; Pergamon: Amsterdam, 61988; Pp 1−646, 1981.

(23) Nitrogen and Air. In IUPAC Solubility Data Series (Volume 10); Battino, R; Ed; Pergamon: Amsterdam, 61988; Pp 1−570, 1982.

(24) Décultot, M.; Ledoux, A.; Fournier-Salaün, M.-C.; Estel, L. Solubility of CO2 in Methanol, Ethanol, 1,2-Propanediol and Glycerol from 283.15 K to 373.15 K and up to 6.0 MPa. *J. Chem. Thermodyn.* 2019, *138*, 67–77. https://doi.org/10.1016/j.jct.2019.05.003.

(25) Yamamoto, H.; Kamei, H.; Tokunaga, J. Solubilities of Argon, Oxygen and Nitrogen in 1,2-Propanediol + Water Mixed Solvent at 298.15 K and 101.33 kPa. J. Chem. Eng. Japan. 1994, 27 (4), 455-459.

(26) Li, Y.; Liu, Q.; Huang, W.; Yang, J. Solubilities of CO2 Capture Absorbents Methyl Benzoate, Ethyl Hexanoate and Methyl Heptanoate. *J. Chem. Thermodyn.* 2018, *127*, 25–32. https://doi.org/10.1016/j.jct.2018.07.010.

(27) Li, X.; Jiang, Y.; Han, G.; Deng, D. Investigation of the Solubilities of Carbon Dioxide in Some Low Volatile Solvents and Their Thermodynamic Properties. *J. Chem. Eng. Data* 2016, *61* (3), 1254–1261. https://doi.org/10.1021/acs.jced.5b00893.

(28) Henni, A.; Tontiwachwuthikul, P.; Chakma, A. Solubilities of Carbon Dioxide in Polyethylene Glycol Ethers. *Can. J. Chem. Eng.* 2008, *83* (2), 358–361. https://doi.org/10.1002/cjce.5450830224.

(29) Zhao, Z.; Xing, X.; Tang, Z.; Zhao, Y.; Fei, W.; Liang, X.; He, Z.; Zhang, S.; Guo, D. Solubility of CO2 and H2S in Carbonates Solvent: Experiment and Quantum Chemistry Calculation. *Int. J. Greenh. Gas Control* 2017, *59*, 123–135. https://doi.org/10.1016/j.ijggc.2017.02.011.

(30) Deng, D.; Han, G.; Jiang, Y.; Ai, N. Solubilities of Carbon Dioxide in Five Biobased Solvents. J. Chem. Eng. Data 2015, 60 (1), 104–111.

(31) Yogish, K. Solubility of CO2 in Some Physical Solvents. *J. Chem. Eng. Jpn.* 1991, *24* (1), 135–137. https://doi.org/10.1252/jcej.24.135.

(32) Miller, M. B.; Chen, D.-L.; Luebke, D. R.; Johnson, J. K.; Enick, R. M. Critical Assessment of CO2 Solubility in Volatile Solvents at 298.15 K. *J. Chem. Eng. Data* 2011, *56* (4), 1565–1572. https://doi.org/10.1021/je101161d.

(33) Hansen, C. M. *Hansen Solubility Parameters: A User's Handbook*, 2nd ed.; CRC Press: Boca Raton, 2007.

(34) Gennaro, A.; Isse, A. A.; Vianello, E. Solubility and Electrochemical Determination of CO2 in Some Dipolar Aprotic Solvents. *J. Electroanal. Chem. Interfacial Electrochem.* 1990, *289* (1–2), 203–215. https://doi.org/10.1016/0022-0728(90)87217-8.

(35) Anouti, M.; Dougassa, Y. R.; Tessier, C.; El Ouatani, L.; Jacquemin, J. Low Pressure Carbon Dioxide Solubility in Pure Electrolyte Solvents for Lithium-Ion Batteries as a Function of Temperature. Measurement and Prediction. *J. Chem. Thermodyn.* 2012, *50*, 71–79. https://doi.org/10.1016/j.jct.2012.01.027.

(36) Li, Y.; You, Y.; Huang, W.; Yang, J. Solubility Measurement and Thermodynamic Properties Calculation for Several CO_2 + Ether Absorbent Systems. *J. Chem. Eng. Data* 2019, *64* (3), 1020–1028. https://doi.org/10.1021/acs.jced.8b00936.

(37) Li, Y.; Zheng, D.; Dong, L.; Xiong, B. Solubilities of Carbon Dioxide in 2-Methoxyethyl Acetate, 1-Methoxy-2-Propyl Acetate and 3-Methoxybutyl Acetate. *J. Chem. Thermodyn.* 2014, *74*, 126–132. https://doi.org/10.1016/j.jct.2014.01.019.

(38) Li, Y.; Liu, Q.; Huang, W.; Yang, J. Below the Room Temperature Measurements of Solubilities in Ester Absorbents for CO2 Capture. *J. Chem. Thermodyn.* 2018, *127*, 71–79. https://doi.org/10.1016/j.jct.2018.07.021.

(39) Flowers, B. S.; Mittenthal, M. S.; Jenkins, A. H.; Wallace, D. A.; Whitley, J. W.; Dennis, G. P.; Wang, M.; Turner, C. H.; Emel'yanenko, V. N.; Verevkin, S. P.; Bara, J. E. 1,2,3-Trimethoxypropane: A Glycerol-Derived Physical Solvent for CO_2 Absorption. *ACS Sustain. Chem. Eng.* 2017, *5* (1), 911–921. https://doi.org/10.1021/acssuschemeng.6b02231.

(40) Li, Y.; Huang, W.; Zheng, D.; Mi, Y.; Dong, L. Solubilities of CO2 Capture Absorbents 2-Ethoxyethyl Ether, 2-Butoxyethyl Acetate and 2-(2-Ethoxyethoxy)Ethyl Acetate. *Fluid Phase Equilibria* 2014, *370*, 1–7. https://doi.org/10.1016/j.fluid.2014.02.029.

(41) Schappals, M.; Breug-Nissen, T.; Langenbach, K.; Burger, J.; Hasse, H. Solubility of Carbon Dioxide in Poly(Oxymethylene) Dimethyl Ethers. *J Chem Eng Data* 2017, 5.

(42) Gui, X.; Tang, Z.; Fei, W. Solubility of CO2 in Alcohols, Glycols, Ethers, and Ketones at High Pressures from (288.15 to 318.15) K. *J. Chem. Eng. Data* 2011, *56* (5), 2420–2429. https://doi.org/10.1021/je101344v.

(43) Gui, X.; Wang, W.; Wang, C.; Zhang, L.; Yun, Z.; Tang, Z. Vapor–Liquid Phase Equilibrium Data of CO_2 in Some Physical Solvents from 285.19 K to 313.26 K. *J. Chem. Eng. Data* 2014, *59* (3), 844–849. https://doi.org/10.1021/je400985u.

(44) Jou, F.-Y.; Otto, F. D.; Mather, A. E. Solubility of H2S and CO2 in Diethylene Glycol at Elevated Pressures. *Fluid Phase Equilibria* 2000, *175* (1), 53–61. https://doi.org/10.1016/S0378-3812(00)00440-4.

(45) F. Blanchard; B. Carre; F. Bonhomme; P. Biensan; D. Lemordat. Solubility of Carbon Dioxide in Alkylcarbonates and Lactones. *Can. J. Chem.* No. 81, 385–391.

(46) Wu, F.; Zhao, Q.; Tao, L.; Danaci, D.; Xiao, P.; Hasan, F. A.; Webley, P. A. Solubility of Carbon Monoxide and Hydrogen in Methanol and Methyl Formate: 298–373 K and 0.3–3.3 MPa. *J. Chem. Eng. Data* 2019, *64* (12), 5609–5621. https://doi.org/10.1021/acs.jced.9b00676.

(47) Qureshi, M. S.; Le Nedelec, T.; Guerrero-Amaya, H.; Uusi-Kyyny, P.; Richon, D.; Alopaeus, V. Solubility of Carbon Monoxide in Bio-Oil Compounds. *J. Chem. Thermodyn.* 2017, *105*, 296–311. https://doi.org/10.1016/j.jct.2016.10.030.

(48) Brunner, E. Solubility of Hydrogen in 10 Organic Solvents at 298.15, 323.15, and 373.15 K. *J. Chem. Eng. Data* 1985, *30* (3), 269–273. https://doi.org/10.1021/je00041a010.

(49) Brunner, E. Solubility of Hydrogen in Diols and Their Ethers. *J. Chem. Thermodyn.* 1980, *12* (10), 993–1002. https://doi.org/10.1016/0021-9614(80)90140-8.

(50) Purwanto; Deshpande, R. M.; Chaudhari, R. V.; Delmas, H. Solubility of Hydrogen, Carbon Monoxide, and 1-Octene in Various Solvents and Solvent Mixtures. *J. Chem. Eng. Data* 1996, *41* (6), 1414–1417. https://doi.org/10.1021/je960024e.

(51) Krüger, M. B.; Selle, C.; Heller, D.; Baumann, W. Determination of Gas Concentrations in Liquids by Nuclear Magnetic Resonance: Hydrogen in Organic Solvents. *J. Chem. Eng. Data* 2012, *57* (6), 1737–1744. https://doi.org/10.1021/je2013582.

(52) Henni, A.; Tontiwachwuthikul, P.; Chakma, A. Solubility Study of Methane and Ethane in Promising Physical Solvents for Natural Gas Sweetening Operations. *J. Chem. Eng. Data* 2006, *51* (1), 64–67. https://doi.org/10.1021/je050172h.

(53) Hesse, P. J.; Battino, R.; Scharlin, P.; Wilhelm, E. Solubility of Gases in Liquids. 21. Solubility of He, Ne, Ar, Kr, N2, O2, CH4, CF4, and SF6in 2,2,4-Trimethylpentane AtT= 298.15 K. *J. Chem. Thermodyn.* 1999, *31* (9), 1175–1181. https://doi.org/10.1006/jcht.1999.0529.

(54) Battino, R.; Rettich, T. R.; Tominaga, T. The Solubility of Nitrogen and Air in Liquids. *J. Phys. Chem. Ref. Data* 1984, *13* (2), 563–600. https://doi.org/10.1063/1.555713.

(55) Bo, S.; Battino, R.; Wilhelm, E. Solubility of Gases in Liquids. 19. Solubility of He, Ne, Ar, Kr, Xe, N2, O2, CH4, CF4, and SF6 in Normal 1-Alkanols n-ClH2l+1OH (1 .Ltoreq. l .Ltoreq. 11) at 298.15 K. *J. Chem. Eng. Data* 1993, *38* (4), 611–616. https://doi.org/10.1021/je00012a035.

(56) Hesse, P. J.; Battino, R.; Scharlin, P.; Wilhelm, E. Solubility of Gases in Liquids. 20. Solubility of He, Ne, Ar, Kr, N2, O2, CH4, CF4, and SF6 in n-Alkanes n-ClH2l+2 ($6 \leq l \leq 16$) at 298.15 K. *J. Chem. Eng. Data* 1996, *41* (2), 195–201. https://doi.org/10.1021/je9502455.

(57) Pardo, J.; López, M. C.; Santafé, J.; Royo, F. M.; Urieta, J. S. Solubility of Gases in Butanols. I. Solubilities of Nonpolar Gases in 1-Butanol from 263.15 to 303.15 K at 101.33 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1995, *109* (1), 29–37. https://doi.org/10.1016/0378-3812(95)02712-N.

(58) Pardo, J.; López, M. C.; Mayoral, J. A.; Royo, F. M.; Urieta, J. S. Solubility of Gases in Butanols. III. Solubilities of Non-Polar Gases in 2-Butanol from 263.15 to 303.15 K at 101.33 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1997, *134* (1–2), 133–140. https://doi.org/10.1016/S0378-3812(97)00064-2.

(59) Pardo, J.; López, M. C.; Santafé, J.; Royo, F. M.; Urieta, J. S. Solubility of Gases in Butanols II. Solubilities of Nonpolar Gases in 2-Methyl-1-Propanol from 263.15 to 303.15 K at 101.33 KPa Partial Pressure of Gas. *Fluid Phase Equilibria* 1996, *119* (1), 165–173. https://doi.org/10.1016/0378-3812(95)02984-2.

(60) Pardo, J.; Mainar, A. M.; Lopez, M. C.; Royo, F.; Urieta, J. S. Solubility of Gases in Butanols IV. Solubilities of Nonpolar Gases in 2-Methyl-2-Propanol at 303.15 K and 101.33 KPa Partial Pressure of Gas. Fluid Ph. Equilibria. 1999, 155 (1), 127-137.

(61) Weng, W.-L.; Chen, J.-T.; Chang, J.-S.; Chang, S.-L. Vapor–Liquid Equilibria for Nitrogen with 2-Hexanol, 2-Heptanol, or 2-Octanol Binary Systems. *Fluid Phase Equilibria* 2006, *248* (2), 168–173. https://doi.org/10.1016/j.fluid.2006.08.005.

(62) Gallardo, M. A.; Melendo, J. M.; Urieta, J. S.; Losa, C. G. Solubility of Non-Polar Gases in Cyclohexanone between 273.15 and 303.15 K at 101.32 KPa Partial Pressure of Gas. *Can. J. Chem.* 1987, *65* (9), 2198–2202. https://doi.org/10.1139/v87-368.

(63) Gallardo, M. A.; López, M. C.; Urieta, J. S.; Losa, C. G. Solubility of He, Ne, Ar, Kr, Xe, H2, D2, N2, O2, CH4, C2H4, C2H6, CF4, SF6 and CO2 in Cyclopentanone from 273.15 K to 303.15 K and Gas Partial Pressure of 101.33 KPa. *Fluid Phase Equilibria* 1989, *50* (1), 223–233. https://doi.org/10.1016/0378-3812(89)80292-4.

(64) Gallardo, M.A.; López, M.C.; Urieta, J.S.; Gutierrez Losa, C. Solubility of 15 non-polar gases (He, Ne, Ar, Kr, Xe, H2, D2, N2, O2, CH4, C2H4, C2H6, CF4, SF6 and CO2) in cycloheptanone. Fluid Ph. Equilibria. 1990, 58 (1–2), 1990, 159-172.

(65) Arai, C; Yoshitama, T.; Nishihara, K.; Sano, Y. Gas Solubilities in Esters of Oleic Acid. *Kagaku Kougaku Ronbunsyu* 1989, *15* (6), 1193–1195.

(66) Lizano, L. P.; López, M. C.; Royo, F. M.; Urieta, J. S. Solubility of Non Polar Gases in Formaldehyde Diethyl Acetal Between-10 and 30°C, and 1 Atm Partial Pressure of Gas. *J. Solut. Chem.* 1990, *19* (7), 721–728. https://doi.org/10.1007/BF00647390.

(67) Urieta, J. S.; Gibanel, F.; Martínez-López, J. F.; Pardo, J. I.; Mainar, A. M. Solubilities of Gases in Cycloethers. The Solubility of 13 Nonpolar Gases in 2,5-Dimethyltetrahydrofuran at 273.15 to 303.15 K and 101.32 KPa. *J. Chem. Thermodyn.* 2019, *132*, 306–315. https://doi.org/10.1016/j.jct.2018.12.037.

(68) Gibanel, F.; López, M. C.; Royo, F. M.; Rodríguez, V.; Urieta, J. S. Solubility of Nonpolar Gases in Tetrahydropyran at 0 to 30°C and 101.33 KPa Partial Pressure of Gas. *J. Solut. Chem.* 1994, *23* (11), 1247–1256. https://doi.org/10.1007/BF00974033.

(69) Gibanel, F.; López, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutiérrez Losa, C. Solubility of Nonpolar Gases in Hexamethylenoxide. *Fluid Phase Equilibria* 1988, *42*, 261–268. https://doi.org/10.1016/0378-3812(88)80063-3.

(70) Mainar, A.M.; Pardo, J.; Royo, F.M.; Lopez, M.C.; Urieta, J.S. Solubility of nonpolar gases in 2,2,2-trifluoroethanol at 25C and 101.33 kPa partial pressure of gas. J. Solution. Chem. 1996, 25 (6) , 589-595.

(71) Lopez, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutierrez Losa, C. Solubility of Nonpolar Gases in Halogenated Compounds. 1. Solubility of Hydrogen, Deuterium, Nitrogen, Oxygen, Methane, Ethylene, Ethane, Carbon Tetrafluoride, Sulfur Hexafluoride and Carbon Dioxide in Chlorocyclohexane at 263.15-303.15 K and 101.32 KPa of Partial Pressure of Gas. *J. Chem. Eng. Data* 1987, *32* (4), 472–474. https://doi.org/10.1021/je00050a027.

(72) Lopez, M. C.; Gallardo, M. A.; Urieta, J. S.; Gutierrez Losa, C. Solubility of Nonpolar Gases in Halogenated Compounds. 2. Solubility of Hydrogen, Deuterium, Nitrogen, Oxygen, Methane, Ethylene, Ethane, Carbon, Tetrafluoride, Sulfur Hexafluoride and Carbon Dioxide in Bromocyclohexane at 263.15- to 303.15 K and 101.32 KPa Partial Pressure of Gas. *J. Chem. Eng. Data* 1989, *34* (2), 198–200. https://doi.org/10.1021/je00056a015.

(73) Nitta, T.; Nakamura, Y.; Ariyasu, H.; Katayama, T. Solubilities of Nitrogen in Binary Solutions of Acetone with Cyclohexane, Benzene, Chloroform and 2-Propanol. *J. Chem. Eng. Jpn.* 1980, *13* (2), 97–103. https://doi.org/10.1252/jcej.13.97.

(74) Akimoto, T.; Nitta, T.; Katayama, T. Nitrogen Solubility and Vapor Pressure of Binary Mixed Solvents Containing Benzene, Carbon Tetrachloride, Cyclohexane and 1-Hexane. *J. Chem. Eng. Jpn.* 1984, *17* (6), 637–641. https://doi.org/10.1252/jcej.17.637.

(75) Mainar, A. M.; Pardo, J.; García, J. I.; Royo, F. M.; Urieta, J. S. Solubility of Gases in Fluoroorganic Alcohols. *J. Chem. Soc. Faraday Trans.* 1998, *94* (24), 3595–3599. https://doi.org/10.1039/A807488G.

(76) Atlani, M.; Loutaty, R.; Wakselman, C.; Yacono, C. Method of Purifying a Gas Mixture Containing Undesirable Gas Compounds. 4504287, March 12, 1985.

(77) Barber, R.F.G; Ritter, T.J.; Sweeney, C.W. Removing Sulfur Compounds from Gases. 2245889A, January 15, 1992.

(78) Gamsjäger, H.; Lorimer, J. W.; Salomon, M.; Shaw, D. G.; Tomkins, R. P. T. The IUPAC-NIST Solubility Data Series: A Guide to Preparation and Use of Compilations and Evaluations (IUPAC Technical Report). *Pure Appl. Chem.* 2010, *82* (5), 1137–1159. https://doi.org/10.1351/PAC-REP-09-10-33.

(79) KNIME. https://www.knime.com/open-for-innovation-0 (accessed 2021-03 -03).

(80) Varnek, A.; Fourches, D.; Hoonakker, F.; Solov'ev, V. P. Substructural Fragments: An Universal Language to Encode Reactions, Molecular and Supramolecular Structures. *J. Comput. Aided Mol. Des.* 2005, *19* (9–10), 693–703. https://doi.org/10.1007/s10822-005-9008-0.

(81) Varnek, A.; Fourches, D.; Horvath, D.; Klimchuk, O.; Gaudin, C.; Vayer, P.; Solov'ev, V.; Hoonakker, F.; Tetko, I.; Marcou, G. ISIDA - Platform for Virtual Screening Based on Fragment and Pharmacophoric Descriptors. *Curr. Comput. Aided-Drug Des.* 2008, *4* (3), 191–198. https://doi.org/10.2174/157340908785747465.

(82) Spartan 18.0; Wavefunction, Inc.:www.wavefun.com.

(83) Breiman, L. Random Forests. *Mach. Learn.* 2001, *45* (1), 5–32. https://doi.org/10.1023/A:1010933404324.

(84) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, É. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* 2011, *12* (85), 2825–2830.

(85) https://scikit-learn.org/stable/ (accessed 23.10.20).

(86) Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; ACM: San Francisco California USA, 2016; pp 785–794. https://doi.org/10.1145/2939672.2939785.

(87) XGBoost, https://xgboost.readthedocs.io/en/latest/python/python_intro.html (accessed 31.05.21).

(88) Cortes, C.; Vapnik, V. Support-Vector Networks. *Mach. Learn.* 1995, *20* (3), 273–297. https://doi.org/10.1007/BF00994018.

(89) Sahigara, F.; Mansouri, K.; Ballabio, D.; Mauri, A.; Consonni, V.; Todeschini, R. Comparison of Different Approaches to Define the Applicability Domain of QSAR Models. *Molecules* 2012, *17* (5), 4791–4810. https://doi.org/10.3390/molecules17054791.

(90) Descamps, C.; Coquelet, C.; Bouallou, C.; Richon, D. Solubility of Hydrogen in Methanol at Temperatures from 248.41 to 308.20K. *Thermochim. Acta* 2005, *430* (1), 1–7. https://doi.org/10.1016/j.tca.2004.12.001.

(91) Soubeyran, A.; Rouabhi, A.; Coquelet, C. Thermodynamic Analysis of Carbon Dioxide Storage in Salt Caverns to Improve the Power-to-Gas Process. *Appl. Energy* 2019, *242*, 1090–1107. https://doi.org/10.1016/j.apenergy.2019.03.102.

(92) Dicko, M.; Coquelet, C.; Theveneau, P.; Mougin, P. Phase Equilibria of H2S-Hydrocarbons (Propane, n-Butane, and n-Pentane) Binary Systems at Low Temperatures J. Chem. Eng. Data 2012, 57, 5, 1534–1543.

(93) Cadours, R.; Shah, V.; Weiss, C.; Roquet, D.; Lallemand, F. Industrial Operation of HySWEET®, a New Hybrid Solvent for Improved Mercaptan Removal. In *Proceedings of the 2nd Annual Gas Processing Symposium*; Elsevier, 2010; pp 221–228. https://doi.org/10.1016/S1876-0147(10)02024-0.

(94) Skylogianni, E.; Wanderley, R. R.; Austad, S. S.; Knuutila, H. K. Density and Viscosity of the Nonaqueous and Aqueous Mixtures of Methyldiethanolamine and Monoethylene Glycol at Temperatures from 283.15 to 353.15 K. *J. Chem. Eng. Data* 2019, *64* (12), 5415–5431. https://doi.org/10.1021/acs.jced.9b00607.

(95) Heldebrant, D. J.; Koech, P. K.; Glezakou, V.-A.; Rousseau, R.; Malhotra, D.; Cantu, D. C. Water-Lean Solvents for Post-Combustion CO_2 Capture: Fundamentals, Uncertainties, Opportunities, and Outlook. *Chem. Rev.* 2017, *117* (14), 9594–9624. https://doi.org/10.1021/acs.chemrev.6b00768.

(96) Zheng, R. F.; Barpaga, D.; Mathias, P. M.; Malhotra, D.; Koech, P. K.; Jiang, Y.; Bhakta, M.; Lail, M.; V. Rayer, A.; Whyatt, G. A.; Freeman, C. J.; Zwoster, A. J.; Weitz, K. K.; Heldebrant, D. J. A Single-Component Water-Lean Post-Combustion CO_2 Capture Solvent with Exceptionally Low Operational Heat and Total Costs of Capture – Comprehensive Experimental and Theoretical Evaluation. *Energy Environ. Sci.* 2020, *13* (11), 4106–4113. https://doi.org/10.1039/D0EE02585B.

(97) Vahidi, M.; Shokouhi, M. Experimental Solubility of Carbon Dioxide and Hydrogen Sulfide in 2,2′-Thiodiglycol. *J. Chem. Thermodyn.* 2019, *133*, 202–207. https://doi.org/10.1016/j.jct.2019.02.024.

(98) Versteeg, G. F.; van Swaaij, W. P. M. On the Kinetics between CO2 and Alkanolamines Both in Aqueous and Non-Aqueous Solutions—II. Tertiary Amines. *Chem. Eng. Sci.* 1988, *43* (3), 587–591. https://doi.org/10.1016/0009-2509(88)87018-0.

(99) Versteeg, G. F.; van Swaaij, W. P. M. On the Kinetics between CO2 and Alkanolamines Both in Aqueous and Non-Aqueous Solutions—I. Primary and Secondary Amines. *Chem. Eng. Sci.* 1988, *43* (3), 573–585. https://doi.org/10.1016/0009-2509(88)87017-9.

(100) Maddox, R. N.; Mains, G. J.; Rahman, M. A. Reactions of Carbon Dioxide and Hydrogen Sulfide with Some Alkanolamines. *Ind. Eng. Chem. Res.* 1987, *26* (1), 27–31. https://doi.org/10.1021/ie00061a006.

(101) Rainbolt, J. E.; Koech, P. K.; Yonker, C. R.; Zheng, F.; Main, D.; Weaver, M. L.; Linehan, J. C.; Heldebrant, D. J. Anhydrous Tertiary Alkanolamines as Hybrid Chemical and Physical CO2 Capture Reagents with Pressure-Swing Regeneration. *Energy Environ. Sci.* 2011, *4* (2), 480–484. https://doi.org/10.1039/C0EE00506A.

(102) Mathias, P. M.; Jasperson, L. V.; VonNiederhausern, D.; Bearden, M. D.; Koech, P. K.; Freeman, C. J.; Heldebrant, D. J. Assessing Anhydrous Tertiary Alkanolamines for High-Pressure Gas Purifications. *Ind. Eng. Chem. Res.* 2013, *52* (49), 17562–17572. https://doi.org/10.1021/ie4020974.

(103) Battino, R.; Clever, H. L. The Solubility of Gases in Liquids. *Chem. Rev.* 1966, *66* (4), 395–463. https://doi.org/10.1021/cr60242a003.

865 (104) Pierotti, R. A. A Scaled Particle Theory of Aqueous and Nonaqueous Solutions. *Chem.*
866 *Rev.* 1976, *76* (6), 717–726. https://doi.org/10.1021/cr60304a002.
867 (105) Abboud, J.-L. M.; Notari, R. Critical Compilation of Scales of Solvent Parameters. Part I.
868 Pure, Non-Hydrogen Bond Donor Solvents. *Pure Appl. Chem.* 1999, *71* (4), 645–718.
869 https://doi.org/10.1351/pac199971040645.
870 (106) Lewis, M.; Wu, Z.; Glaser, R. Polarizabilities of Carbon Dioxide and Carbodiimide.
871 Assessment of Theoretical Model Dependencies on Dipole Polarizabilities and Dipole
872 Polarizability Anisotropies. *J. Phys. Chem. A* 2000, *104* (48), 11355–11361.
873 https://doi.org/10.1021/jp002927r.
874 (107) Zeng, W.; Du, Y.; Xue, Y.; Frisch, H. L. Solubility Parameters. In *Physical Properties of*
875 *Polymers Handbook*; Mark, J. E., Ed.; Springer New York: New York, NY, 2007; pp 289–
876 303. https://doi.org/10.1007/978-0-387-69002-5_16.
877 (108) Sigma-Aldrich Catalog, https://www.sigmaaldrich.com/FR/En/Sds/Aldrich/M56557
878 (Accessed 11.08.21).
879 (109) Kauffman, G. W.; Jurs, P. C. Prediction of Surface Tension, Viscosity, and Thermal
880 Conductivity for Common Organic Solvents Using Quantitative Structure−Property
881 Relationships. *J. Chem. Inf. Comput. Sci.* 2001, *41* (2), 408–418.
882 https://doi.org/10.1021/ci000139t.
883 (110) Singh, A. K.; Bilal, M.; Iqbal, H. M. N.; Raj, A. Trends in Predictive Biodegradation for
884 Sustainable Mitigation of Environmental Pollutants: Recent Progress and Future Outlook.
885 *Sci. Total Environ.* 2021, *770*, 144561. https://doi.org/10.1016/j.scitotenv.2020.144561.
886