



Computer-Aided Molecular Design and Selection of CO₂ Capture Solvents Based on Thermodynamics, Reactivity and Sustainability

Journal:	<i>Molecular Systems Design & Engineering</i>
Manuscript ID	ME-ART-06-2016-000049.R1
Article Type:	Paper
Date Submitted by the Author:	24-Jul-2016
Complete List of Authors:	<p>Papadopoulos, Athanasios; Centre for Research and Technology-Hellas, Chemical Process and Energy Resources Institute Badr, Sara; Institute for Chemical and Bio-Engineering, ETH Zurich Chremos, Alexandros; Imperial College London; National Institute of Standards and Technology Forte, Esther; Imperial College, Dept. of Chemical Eng. and Chemical Technology Zarogiannis, Theodoros; Centre for Research and Technology-Hellas, Chemical Process and Energy Resources Institute Seferlis, Panos; Aristotle University of Thessaloniki, Department of Mechanical Engineering; Centre for Research and Technology-Hellas, Chemical Process and Energy Resources Institute Papadokonstantakis, Stavros; Institute for Chemical and Bio-Engineering, ETH Zurich; Chalmers University of Technology, Energy and Environment Galindo, Amparo; Imperial College London, Chemical Engineering Jackson, George; Imperial College, Department of Chemical Engineering Adjiman, Claire; Imperial College London, Department of Chemical Engineering and Chemical Technology</p>

“Design, System, Application” statement

We propose a Computer Aided Molecular Design (CAMD) method which employs optimization to support the synthesis and selection of high performance molecules for use in process systems and to guide experimental efforts. The method can be used to address challenging applications where a) the desired molecules exhibit phase and chemical equilibrium, b) numerous combinations of molecules need to be evaluated, and c) multiple criteria must be considered to capture the effects of molecular chemistry on the process system performance. The method is applied to the design of solvents in chemical absorption processes for the separation of carbon dioxide (CO₂) from gas streams. The molecular design problem is first approached via a fast screening stage where molecules are evaluated based on the simultaneous consideration of multiple performance indices pertaining to thermodynamics, reactivity and sustainability. A few high-performance solvents are further evaluated using an advanced group contribution equation of state to predict reliably the highly non-ideal equilibrium behavior of solvent-water-CO₂ mixtures. Several promising novel solvents for CO₂ capture are proposed and can now be assessed experimentally. The proposed method can readily be applied to other chemical absorption processes to accelerate the identification of novel solvents.

Computer-Aided Molecular Design and Selection of CO₂
Capture Solvents Based on Thermodynamics, Reactivity and
Sustainability

Athanasios I. Papadopoulos^a, Sara Badr^b, Alexandros Chremos^{c,f}, Esther Forte^c,
Theodoros Zarogiannis^{a,d}, Panos Seferlis^{d,al}, Stavros Papadokonstantakis^{b,e}, Amparo
Galindo^c, George Jackson^c, Claire S. Adjiman^c

^aChemical Process and Energy Resources Institute, Centre for Research and
Technology-Hellas, 57001 Themi, Greece

^bSwiss Federal Institute of Technology, Institute for Chemical- and Bioengineering,
Zurich, Switzerland

^cImperial College London, Department of Chemical Engineering, Centre for Process
Systems Engineering, London SW7 2AZ, UK

^dDepartment of Mechanical Engineering, Aristotle University of Thessaloniki, 54124
Thessaloniki, Greece

^eChalmers University of Technology, Department of Energy and Environment,
Kemivägen 4, 41296 Gothenburg, Sweden

^fCurrent address: Materials Science and Engineering Division, National Institute of
Standards and Technology, Gaithersburg, MD, 20899, U.S.A.

^l Corresponding author: seferlis@auth.gr

Abstract

The identification of improved carbon dioxide (CO₂) capture solvents remains a challenge due to the vast number of potentially-suitable molecules. We propose an optimization-based computer-aided molecular design (CAMD) method to identify and select, from hundreds of thousands of possibilities, a few solvents of optimum performance for CO₂ chemisorption processes, as measured by a comprehensive set of criteria. The first stage of the approach consists in a fast screening stage where solvent structures are evaluated based on the simultaneous consideration of important pure component properties reflecting thermodynamic, kinetic, and sustainability behaviour. The impact of model uncertainty is considered through a systematic method that employs multiple models for the prediction of performance indices. In a second stage, high-performance solvents are further selected and evaluated using a more detailed thermodynamic model, namely the group-contribution statistical associating fluid theory for square well potentials (SAFT- γ SW), to predict accurately the highly non-ideal chemical and phase equilibrium of the solvent-water-CO₂ mixtures. The proposed CAMD method is applied to the design of novel molecular structures and to the screening of a dataset of commercially available amines. New molecular structures and commercially-available compounds that have received little attention as CO₂ capture solvents are successfully identified and assessed using the proposed approach. We recommend that these solvents given priority in experimental studies to identify new compounds.

Keywords: Solvents, Computer-aided molecular design, Sustainability, SAFT, CO₂ capture

1. Introduction

In view of the pressing need to curtail carbon dioxide (CO₂) emissions to the atmosphere in a world that remains highly dependent on carbon-based fuels for its energy provision, a large number of technological solutions for CO₂ capture are currently being investigated¹. Chemical absorption processes^{2, 3} represent a well-established industrial technology for CO₂ capture compared to many of the alternatives that are being developed: it can be retrofitted onto existing plants and the conditions for both absorption and solvent regeneration (desorption) are relatively easy to meet. However, some major downsides must be overcome to enable the widespread adoption of this technology. These include the high energy input required for solvent regeneration and negative environmental impacts associated with the solvents, to name but a few⁴.

Many of these issues can be linked to the solvents most commonly used in such processes, namely amine-based organic compounds. In view of the shortcomings for these solvents, significant efforts are being devoted to identifying solvents that can significantly improve the economic and sustainability performance of chemisorption processes, compared to existing options. The identification of improved solvents, however, is very challenging due to a) the highly non-ideal solvent-CO₂-water chemical interactions, b) the countless combinations of potential capture solvents and blend candidates and c) the complex interactions between solvent and process that require the simultaneous consideration of multiple criteria⁵ in the selection of solvents with optimum capture features. To address this last aspect, one must avoid focusing on a narrow property target (e.g., Oexmann and Kather⁶ argue convincingly against searching for low heat of absorption solvents) and instead consider a comprehensive set of properties related to thermodynamics, kinetics and sustainability.

A large number of experimental studies have been undertaken to identify solvents and solvent blends that perform better than the current benchmark solvent, aqueous solutions of

monoethanolamine (MEA). Based on the collective experience of a large community of researchers, several rules-of-thumb have been formulated to help generate molecular structures that are likely to perform well as CO₂ chemisorption solvents. For example, Singh et al.⁷⁻⁹ and Singh and Versteeg¹⁰ performed a large set of experiments to derive useful insights on the effects of amine structure on important parameters such as absorption rate and solvent capacity. Other researchers have developed qualitative design rules which can be used for the heuristic screening of solvents and a selection of these heuristics is listed in Table 1, with more details provided in Section A of the Supporting Information. While qualitative relationships that link structure and properties offer a very useful guide to search the vast space of possibilities, they also reduce the diversity of the options considered, thereby decreasing the potential for novel designs. Additionally, they do not reduce the significant cost of obtaining a measure of the performance of the postulated solvents, which requires the synthesis and characterization of each new solvent. For instance, the laboratory-based investigation of DEAB has required extensive effort as reflected in publications related to compound identification^{11,12} and investigation¹³⁻¹⁶. Thus, in addition to acquiring high-quality data for different solvents, there is a need to accelerate the search using computational methods¹⁷.

Table 1 – Examples of qualitative structure-property relations for CO₂ capture.

Label	Evidence	Compound type(s)	Structure-Property Relation
(1)	Puxty et al. ¹⁸ Yamada et al. ¹⁹ Tontiwachwuthikul et al. ¹²	Hydroxylamines	2-4 carbons between hydroxyl and amine functional groups increase absorption capacity.
(2)	Singh and Versteeg ¹⁰ Aronu et al. ²⁰	Diamines	Longer chains between amine groups, up to 6 carbons, increase basicity and improves desorption.
(3)	Singh and Versteeg ¹⁰ Singh et al. ^{8,9}	Hydroxylamines Diamines	At least 2 carbon atoms between amine and hydroxyl or first amine and amine branch improve desorption.
(4)	Singh and Versteeg ¹⁰	Linear hydroxylamines	Up to 4 carbon atoms between hydroxyl and amines improves

			desorption.
(5)	Zhang et al. ²¹	Branched alkylamines	An alkyl branch at the α carbon improves kinetics; at the β carbon, it leads to precipitation.
(6)	Sartori et al. ²²	Sterically hindered amines	<ul style="list-style-type: none"> • A primary amine is hindered when the amino group is attached to a tertiary carbon; • A secondary amine is hindered when the amino group is attached to at least one secondary or tertiary carbon.

Systematic computer-aided molecular design (CAMD) methods offer a promising alternative route to the identification of solvents that should be investigated experimentally, by allowing the fast, cost-effective, and automated evaluation of a larger and more diverse set of solvent properties compared to what is possible through purely experimental studies²³. CAMD methods have been particularly successful in addressing some of the challenges in the design and selection of solvents for a range of separation systems and other applications; they have been used to explore a very wide range of solvent structures, leading to the identification of either novel molecular structures or conventional, but previously overlooked, optimum molecules. The reader is referred to a few general works^{24, 25-27} and references therein to gain a historical perspective.

Because many of the predictive models used are relatively simple and yet cover a very broad set of molecular structures, uncertainty has emerged as an important issue in CAMD. Approaches to mitigate the impact of uncertainty have previously been proposed within systematic CAMD methods for the design of solvents as separation²⁸⁻³⁰ and reaction³¹ media, of polymers³², and of heat exchange fluid mixtures³³. An uncertainty quantification method for property predictions through GC was also recently proposed^{34, 35}.

While research on CAMD methods has led to the development of many sophisticated and successful CAMD techniques, the deployment of a systematic approach in the context of chemisorption is not without difficulty. The models typically used in chemisorption process design, while of high accuracy, are specific to a given solvent and require extensive experimental data on each solvent to be studied; this is the case of models of the thermodynamics and kinetics of aqueous alkanolamine and CO₂ mixtures (e.g., eNRTL models³⁶⁻³⁸), and of the kinetics of degradation reactions³⁹⁻⁴¹. Few predictive models that relate solvent molecular structure or solvent blend composition to properties, and hence performance, are available to support CAMD activities. As a result, recent research efforts to screen CO₂ capture solvents *quantitatively* have mostly been based on the use of simple approximation models or statistical correlations^{42, 43}. While such approaches led to some promising results, the narrow range of performance criteria considered and the relatively low accuracy of the predictive models have not allowed the full benefits of *in silico* solvent design to be realized.

Encouragingly, three CAMD-based approaches have been reported for the design of CO₂ capture solvents for chemisorption processes. Eden and co-workers^{44, 45} have proposed solvents to replace MDEA in an absorption/desorption process, using a range of physical properties to define the optimum molecular design space⁴⁴ and developing a new CAMD algorithm.⁴⁵ Salazar et al.⁴⁶ used different physical properties as performance criteria to identify few promising solvents. These solvents were then evaluated based on their stripping energy requirement, calculated using ASPEN Plus[®]⁴⁷ process models and with the reaction and thermodynamic models derived in part from similarity with MEA and DEA (diethanolamine), and in part by using undisclosed NRTL⁴⁸ parameters derived based on UNIFAC⁴⁹ groups to model the liquid phase non-ideality.

Among these CAMD-based works, only Salazar et al.⁴⁶ considered the multicomponent chemical and phase equilibria that are central to solvent performance as an explicit part of the design.

However, to the best of our knowledge, the physical model used has not been validated against experimental data and its predictive capabilities are unknown. In this context, the increasing availability of reliable predictive models of relevance to chemisorption provides opportunities to enhance CAMD approaches to CO₂ capture solvent design by broadening the range of performance metrics and increasing their reliability. Chemical reactivity, for instance, has been predicted using continuum-solvation models, among other computational chemistry methods, for the calculation of the amine base strength as a measure of molecular reactivity^{50, 51}. Although such models are computationally expensive, they could in the future be included in CAMD approaches that embed quantum mechanical models^{52, 53}.

Another interesting route is the use of a predictive equation of state, such as the statistical associating fluid theory equation for potentials of variable range, SAFT-VR^{54, 55}, which has been applied within a CAMD methodology to the design of solvents for the separation of CO₂ from a natural gas stream through physical absorption using a single family of compounds⁵⁶ or within a group-contribution version⁵⁷ for a broader range of compounds⁵⁸. The use of the SAFT platform to design CO₂ capture solvents for physical absorption was also proposed by Bardow et al.⁵⁹, Oyarzun et al.⁶⁰, Stavrou et al.⁶¹ and Lampe et al.⁶² using the PC-SAFT EoS, the perturbed-chain version of SAFT⁶³ and the PCP-SAFT EoS, its extension to polar fluids⁶⁴⁻⁶⁶. In the context of chemisorption, the SAFT thermodynamic platform has been used to predict the phase and chemical equilibria of CO₂ capture fluids through a parameterization of a group contribution version of the statistical associating fluid theory for potentials of variable range equation of state, SAFT- γ SW^{67, 68}. This approach has been applied successfully to model primary alkanolamines and their mixtures with water and CO₂⁶⁹⁻⁷¹. With such a model, it becomes possible to treat aqueous solutions of CO₂ and amines for which no data are available, and to investigate the effect of decision variables such as composition and temperature, as well as the performance of mixtures containing several amine components.

Drawing on these advances, we aim in our work to identify a set of promising solvents that offer a simultaneously high performance in many criteria associated with effective chemical absorption of CO₂ and solvent regeneration. For the first time in CO₂ capture research, we take an approach that combines thermodynamic behaviour, reactivity and sustainability as performance criteria in solvent design or selection, and that includes predictive models of the phase and chemical equilibria of the solvents identified. Furthermore, we recognise the inherent uncertainty in many of the methods we use, and propose a strategy to address this issue and reduce the risk of premature elimination of solvent candidates. The proposed approach can easily be extended to include other aspects not considered here provided that appropriate structure-property models are available (e.g., the selectivity of the solvent towards CO₂ can be predicted using thermodynamic models similar to those used to predict CO₂ loading). Other considerations not accounted for here, such as propensity to oxidative degradation, can be investigated for the solvents in the final set obtained. The proposed computer-based solvent design framework is applied in two ways: (i) a database of commercially available molecules is screened systematically for high-performance capture options (solvent *selection*), (ii) novel molecules that are found to be optimum capture solvents are identified (solvent *design*).

The article is organised as follows. In Section 2, an overview of the proposed systematic CAMD methodology is given. The approach is generic and can be applied to different molecular design applications depending on the properties and property models selected. In Section 3, the specific choices required for the application of the method to CO₂ capture are discussed. The solvents identified are presented in Section 4; the approach is assessed based on the predicted performance of the solvents generated relative to MEA, their novelty, and any indication from the literature that they are viable options.

2. Overview of generic design and selection framework

An overview of the proposed generic framework for solvent design and selection is presented in this section, together with a description of each step. Our aim is to enable the identification of efficient CO₂ capture solvents by investigating two cases:

- The *selection* of solvents from a database of pre-existing chemical compounds (Selection problem).
- The *design* of novel solvents using an optimization-based CAMD method (Design problem).

Both cases are addressed within a single framework whose major characteristics include the use of a wide set of properties as criteria for solvent selection and design, the use of simple models to develop an initial list of candidate solvents, the explicit consideration of uncertainty in decision-making and the pruning of the list of candidate solvents using a more advanced predictive model (here, we use the SAFT- γ SW EoS). Uncertainty is accounted for by generating a distribution of solvent ranks that results from the use of different prediction models for the same properties. The distribution of ranks is widely used in non-parametric statistics⁷² to detect patterns in data without requiring information on error distributions or other assumptions on the data under consideration. These ideas are implemented within a systematic, two-stage approach, as illustrated in Figure 1. Each step is introduced in the remainder of this section.

[Insert Figure1.tif here]

Figure 1 - Unified Computer-Aided Molecular Design framework for Selection (left) or Design (right).

2.1 Stage 1 of the methodology

The aim of Stage 1 is to quickly identify and eliminate solvents of inferior performance, producing a “long list” of candidates. In the design problem, the algorithm used⁷³ can rapidly evaluate around 200,000 molecular structures to generate the few tens of solvents comprising the

“long list”. Solvents are designed and/or evaluated using GC and non-GC-based models for the prediction of properties that are specifically selected to reflect the major driving forces for CO₂ capture. The selected properties provide a set of solvent performance measures that link the chemical structure of the solvent to thermodynamic, reactivity and sustainability characteristics of importance in the capture/regeneration process (see section 3 for details).

2.1.1 Step 1.1 – Generation of an initial solvent list

Step 1.1 applies to the Design problem, but not to the Selection problem, in which the database constitutes the initial solvent list. In this step, candidate solvents are generated from the design space defined by the set of functional groups and the design rules that determine which molecules should be considered. The long list is generated by using a CAMD approach based on multiple property criteria to identify a set of highly performing options⁷³. It combines GC models with a multi-objective optimization approach to synthesize optimum novel or existing solvents. This approach enables the simultaneous evaluation of several properties as performance criteria and hence it is particularly suitable for the generation of a rich pool of optimum solvent candidates reflecting important property trade-offs. In optimization terms, this pool represents the non-dominated (Pareto) front of optimum solutions⁷³. Every solvent in this front is such that it is not possible to improve the value of one of its properties without deterioration in at least one of its other properties. Additional details regarding the method may be found in Papadopoulos and Linke⁷³ and Papadopoulos et al.³³. The resulting non-dominated list of solvents includes high performance candidates but is sufficiently small so that every solvent can be further evaluated in subsequent steps.

2.1.2 Step 1.2 – Calculation of aggregate property indices

In Step 1.2, several aggregate property indices are generated for each molecule in the long list or database. GC models and non-GC models are used in order to calculate all properties of interest. Whenever possible, each property is evaluated with different property models. Each combination

of property models is used to compute an aggregate index and the collection of aggregate indices forms the basis for a balanced solvent selection that avoids biases towards specific properties.

To explain how the aggregate indices are computed, let us consider a set of properties $P_r = \{1, \dots, N_p\}$ that are linked to performance. Let each property $j \in P_r$ be calculated from a set of available models $l_j = \{1, \dots, N_j^{md}\}$ consisting of a total of N_j^{md} models per property. The number of all possible model combinations is given by

$$N_c = \prod_{j \in P_r} N_j^{md} \quad (1)$$

Furthermore consider a solvent i in the set $G = \{1, \dots, N_s\}$ of N_s candidate molecules. One can then define an aggregate property index $J_{i,k}$ for solvent i based on the k^{th} model combination, $k \in \{1, \dots, N_c\}$, corresponding to some choice l_j of model ($l_j \in \{1, \dots, N_j^{md}\}$) for a property j :

$$J_{i,k} = \sum_{j \in P_r} a_j \cdot x_{i,l_j}^* \quad (2)$$

where a_j represents a unit coefficient that is positive if property j needs to be minimized and negative if it needs to be maximized. x_{i,l_j}^* represents a scaled value of property j for solvent i obtained using model l_j . Details on the scaling approach are reported in Section B.1 of the Supporting Information.

The set of indices $J_{i,k}$, $i \in \{1, \dots, N_s\}$, $k \in \{1, \dots, N_c\}$, may be represented as a matrix \mathbf{J} of dimensions $(N_s \times N_c)$ with each model combination k (column of \mathbf{J}) indicating the performance of each solvent i (row of \mathbf{J}). As an example of matrix \mathbf{J} , assume a set G of $N_s = 4$ solvents is used and a set P_r of $N_p = 2$ properties with $N_1^{md} = 2$ and $N_2^{md} = 3$ models (hence $N_c = 6$) is available for the calculation of property 1 and 2, respectively. The elements of matrix \mathbf{J} are calculated as shown in Table 2.

Table 2 - Example of matrix **J** for 4 solvents and 6 property model combinations

Solvent	$k = 1, j = \{1, 2\}$	$k = 2, j = \{1, 2\}$...	$k = 6, j = \{1, 2\}$
	$l_1 = 1, l_2 = 1$	$l_1 = 1, l_2 = 2$		$l_1 = 2, l_2 = 3$
1	$J_{1,1} = a_1x_{1,1}^* + a_2x_{1,2}^*$	$J_{1,2} = a_1x_{1,1}^* + a_2x_{1,2}^*$...	$J_{1,6} = a_1x_{1,1}^* + a_2x_{1,2}^*$
2	$J_{2,1} = a_1x_{2,1}^* + a_2x_{2,2}^*$	$J_{2,2} = a_1x_{2,1}^* + a_2x_{2,2}^*$...	$J_{2,6} = a_1x_{2,1}^* + a_2x_{2,2}^*$
3	$J_{3,1} = a_1x_{3,1}^* + a_2x_{3,2}^*$	$J_{3,2} = a_1x_{3,1}^* + a_2x_{3,2}^*$...	$J_{3,6} = a_1x_{3,1}^* + a_2x_{3,2}^*$
4	$J_{4,1} = a_1x_{4,1}^* + a_2x_{4,2}^*$	$J_{4,2} = a_1x_{4,1}^* + a_2x_{4,2}^*$...	$J_{4,6} = a_1x_{4,1}^* + a_2x_{4,2}^*$

2.1.3 Step 1.3 – Identification of candidate molecules

In Step 1.3, a rank-based approach is used to identify the top candidate solvents from the initial list (i.e., from the database in the selection problem or the Pareto set in the design problem), rather than focusing on the numerical value of the $J_{i,k}$ indices. If only one model were used for each property, making **J** a one-dimensional vector, the solvents could be ranked from best to worst in order of increasing value of $J_{i,k}$. When several models are used for one property, the differences between the multiple columns in **J** provide a measure of the underlying model uncertainty and make it possible to identify solvents with potentially high performance in spite of model uncertainty.

To this end, the elements in each column of **J** are sorted in ascending order independently for every column (model combination) k , yielding a matrix **J'** of independently rank-ordered columns. The transformations needed to go from **J** to **J'** can be used to construct a matrix of ranked lists, **L**, in which an entry $L_{r,k}$ is an integer that denotes the identity of the r^{th} ranked solvent according to model combination k . Consider an example with the same dimensionality as that in Table 2: the elements of matrix **L** will be of the form presented in Figure 2a. The solvents may follow a different order for different k values (i.e., different combinations of models). For example, solvent 3 ranks first for $k = 2$, second for $k = 1, 3, 5, 6$, and third for $k = 4$. Also notice that all four solvents appear in the top 2 positions (ranks 1 and 2). For this small solvent set, the

number of appearances of each solvent in the top 2 positions can be used to select solvents to be taken forward for further analysis, as shown in Figure 2b: solvents 1 and 3 show the strongest predicted performance, closely followed by solvent 2, while solvent 4 appears to be an under-performer.

[Insert Figure2.tif here]

Figure 2 - a) Example of rank-ordered solvents in matrix **L** where each entry represents the identity of a solvent. b) Number of occurrences of each molecule in the top 2 solvents, based on (a).

The frequent appearance of specific solvents at the top implies an agreement in the predictions made by multiple different models. On the other hand, even solvents that appear with lower frequency at the top are also considered thanks to the use of the distribution of ranks to select candidate molecules. This allows a reduction of the original solvent list without bias due to uncertainty in the predictions. This is important because there is often little or no prior knowledge on the predictive accuracy of the models employed. As a result, agreement in the predictions obtained from different models does not necessarily imply a higher accuracy than in the case where fewer models result in similar or different predictions. With the proposed approach, assumptions related to accuracy are avoided. More details are provided in Section B.2 of the Supporting Information.

The multi-criterion approach adopted for Stage 1, which combines a range of properties and multiple property models, provides a robust evaluation of the solvents in the initial candidate list, in spite of the use of relatively simple models. As a result, a few effective solvents emerge as the best performers: this reduced solvent list is progressed to Stage 2.

2.2 Stage 2 of the methodology

In Stage 2 the most promising candidates from Stage 1 are further evaluated using a combination of ranking, structure-property relations and, where feasible, more detailed predictive models to

assess mixture behaviour. This allows the solvent list to be narrowed down further and priorities to be set for further research and experimental investigation.

2.2.1 Step 2.1 – Prioritisation using average indices and heuristics

The solvent candidates are prioritised by using average values of the property indices. A single average index, $J_{i,ave}$, is calculated for each solvent i , using average values of the properties for which multiple models are available, as follows:

$$J_{i,ave} = \sum_{j \in P_r} a_j \cdot x_{i,j}^{ave}, \quad (3)$$

where $x_{i,j}^{ave}$ is the average value of property j for solvent i (calculation details are available in section B.2 of the Supporting Information). The resulting ranked list can be used to prioritise further investigation, with the number of solvents to be pursued depending on the resources available. The list of candidate solvents can also be further refined by comparing the molecular structures identified to existing solvents and by gathering evidence from the literature on the effectiveness or lack thereof of similar molecules. This information is especially useful in the context of solvent design, where the structures designed in Stage 1 may represent multiple isomers due to the lack of full connectivity information in most group contribution models. Molecular structures of interest can be fully specified based on prior knowledge and heuristics, such as those shown in Table 1.

2.2.2 Step 2.2 – Evaluation of mixture properties with advanced property models

As a final stage before experimental investigation, advanced property models can be applied to the top-ranked solvents arising from Step 2.1 to assess more accurately their likely behaviour within the application of interest. This may require the calculation of mixture properties (e.g., absorption of CO₂) over a range of pressures and temperatures, of reactivity or other relevant properties. This stage may lead to a revision of the rank-ordered list of solvents to help prioritise experimental effort.

3. Application of proposed framework to CO₂ chemisorption

In this section, we discuss the specific application of the generic framework presented in Section 2 to the selection and design of CO₂ capture solvents. At the heart of the framework is the identification of relevant property criteria spanning thermodynamics, reactivity and sustainability, as well as appropriate models to relate these to the solvent molecules; the models and criteria relevant to Stage 1 are introduced in Section 3.1, while those for Stage 2 are discussed in Section 3.2. In Section 3.3, we define the “molecular search space”, i.e., the rules that define the new and existing molecules that are considered as potential capture solvents.

3.1 Selection of relevant properties and property models for Stage 1

Numerous solvent properties can be considered as screening criteria to select optimum CO₂ capture solvents; Hoff et al.⁷⁴ recommended a large set of relevant properties, independently of the effort required for their calculation. Several authors^{7-10, 12, 42,43} have provided significant evidence that, prior to using rigorous models or lab-scale experimentation, simpler models, indices or guidelines facilitate a quick yet reliable screening of CO₂ capture solvents. Several properties are selected for Stage 1 based a) on their potential to reflect important thermodynamic, reactivity and sustainability characteristics relevant to this application, b) on the availability of models that enable their fast calculation, and c) on the availability of sufficient data so that these models may be applied for a wide range of molecular structures. These properties are summarized in Table 3, together with Stage 2 properties, and their connections with broader criteria for design⁵ are highlighted. The property models used are presented in Tables B1 and B2 of the Supporting Information. The rationale for the property choices made and the relevant criteria are described in the remainder of this section.

Table 3- Properties considered as criteria for solvent design and selection in Stage 1 and their impact on absorption/desorption process characteristics and the key performance indicators (KPIs) highlighted by Kim and Svendsen⁵: energy consumption, environmental impact and cost (CAPEX - capital expenditure and OPEX - operating expenditure).

Criterion	Impact on absorption/desorption process
Thermodynamic	
-CO ₂ solubility <ul style="list-style-type: none"> Solubility parameter (δ) (Stage 1) Chemical and phase equilibrium (Stage 2) 	Ability and capacity to dissolve CO ₂ / All design and operating parameters are affected (CAPEX and OPEX, energy consumption).
-Vapour pressure <ul style="list-style-type: none"> Amine pressure (P_{vp}) (Stage 1) Phase equilibrium (Stage 2) 	Solvent losses (OPEX and environmental impact).
-Liquid heat capacity (C_p) (Stage 1)	Sensible heat/ Reboiler duty/ Heat exchanger size (CAPEX and OPEX, energy consumption).
-Density <ul style="list-style-type: none"> Amine density (ρ) (Stage 1) Phase equilibrium (Stage 2) 	Equipment size-capacity (directly CAPEX).
-Surface tension (σ) (Stage 1)	Mass transfer coefficients-packing material characteristics (directly CAPEX).
-Viscosity (η) (Stage 1)	Mass transfer coefficients-packing material characteristics (directly CAPEX).
-Boiling point temp. (T_{bp}) (Stage 1)	Solvent evaporation losses/ Reboiler duty (OPEX, energy consumption, environmental impact).
-Melting point temp. (T_m) (Stage 1)	Solvent solidification (directly OPEX).
Reactivity	
-Amine basicity (pK_a) (Stage 1) -Chemical equilibrium (Stage 2)	Solvent ability to react and solvent kinetics/ All design and operating parameters are affected (CAPEX and OPEX, energy consumption).
Sustainability	
-Environmental health and safety hazards (<i>EHS</i>) (Stage 1)	Inherent hazards of solvent in the capture processes/ Cradle-to-gate environmental impact of solvent (per kg of solvent used)/ Environmental impact of the reclaimer waste (e.g. characterization as hazardous waste according to toxicity and corrosion indices) if landfilled.
-Cumulative energy demand (<i>CED</i>) (Stage 1)	
-Global warming potential (<i>GWP</i>) (Stage 1)	
-Eco-indicator 99 (<i>EI99</i>) (Stage 1)	

3.1.1 Thermodynamic property criteria at Stage 1

The relative energy difference (*RED*) is derived from the solubility parameter (δ) and is used as a measure of miscibility. Two compounds with similar values of δ are likely to be miscible. The

solubility parameter can be related to the cohesive energy density^{75, 76}, i.e., to the energy required to overcome the solvent-solvent interactions⁴². Solvents with $RED < 1$ are generally considered to be favourable for the dissolution of a particular solute^{77, 78}. Clearly, this concept is very useful for the identification of CO₂ capture solvents. Calculation details are reported in Section B.4 of the Supporting Information.

The vapour pressure (P_{vp}) is an indicator of solvent losses and should be as low as possible. Several amine-based solutions used as solvents present a high P_{vp} , hence, water sprinklers are often used in the absorption column to reduce losses.

The liquid heat capacity (C_p) plays an important role in determining the heat required to regenerate the amine-based solution during desorption. This heat can be divided into three components⁷⁹; the sensible heat, the heat of vaporization (mostly that of water, which is the main component evaporated in the reboiler), and the heat of absorption necessary to desorb one mole of CO₂. The sensible heat represents the heat required to raise the temperature of the solution from the absorption to the desorption temperature. This is directly proportional to the heat capacity, C_p , of this solution. Hence, a solution containing an amine with a low C_p may be assumed to contribute towards reducing the sensible heat requirements. C_p is predicted via a GC method⁸⁰ developed specifically for amine-based solvents used in CO₂ capture.

The density (ρ), surface tension (σ) and viscosity (n) of the liquid are three properties tightly associated with the design and operating features of the absorption column. In particular the solvent liquid density, ρ , should be high because this leads to reduced solvent flowrate, equipment size and pumping power requirements. Furthermore, ρ , σ , and n have an effect on the mass transfer coefficient, which increases with increasing values of the density and decreasing values of the surface tension and viscosity, as shown, for instance, in Dvorak et al.⁸¹.

The melting point temperature (T_m) should be lower than the lowest absorption/desorption temperature to avoid solvent solidification under processing conditions.

The boiling point temperature (T_{bp}) should be higher than the highest absorption/desorption temperature to avoid excessive vaporization of the pure solvent under processing conditions.

T_m and T_{bp} are considered in the calculations as an average from Marrero and Gani⁸² and Hukkerikar et al.³⁴.

3.1.2 Reactivity property criterion

The basicity (pK_a) is an important solvent property as it provides information on the reactivity of amines³ and the prevailing reactions⁸³. Reaction rates increase with increasing amine basicity⁸⁴, hence amines with higher basicity are desirable. Quantum chemistry continuum-solvation models^{50, 51} provide a robust prediction of pK_a , but the screening of large solvent sets requires less computationally-demanding approaches. There are two empirical methods⁸⁵ which utilize simple rules for the calculation of pK_a , namely the Perrin-Dempsey-Serjeant (PDS) method and the Taft equation method. PDS has been shown to result in pK_a values for several amine solvents which are as accurate as predictions using quantum chemistry calculations⁵⁰. The Taft method also provides sufficiently accurate predictions for many well-known amine solvents. Both methods are based on an analysis of the effects of substituents in molecules and employ rules for determining the pK_a when a particular substituent is attached to a given position within a molecular structure. In our work pK_a is obtained from the publicly available version of the Marvin 6.0.5⁸⁶ software which combines calculations of molecular partial charge distributions⁸⁷ with PDS and Taft method principles⁸⁸.

3.1.3 Sustainability property criteria

The sustainability characteristics of the investigated solvents are evaluated and compared using two categories of indices: a cradle-to-process gate life cycle assessment (LCA) showing the cumulative environmental impacts of the solvent production process chain; and a safety, health and environmental hazard assessment (EHS) that enables the estimation of harm potential in accidental scenarios within a CO₂ capture plant in the presence of the tested solvents. Details on

the employed calculation methods and values for all metrics are reported in Section C of the Supporting Information. **The LCA metrics** used for a functional unit of 1 kg of solvent produced are the Cumulative Energy Demand, the Global Warming Potential and the Eco-Indicator-99.

The Cumulative Energy Demand (CED), measured in MJ-equivalent, is a resource-oriented indicator which reflects the total primary energy demand during the production cycle and therefore also reflects to some extent the economic impacts⁸⁹.

The Global Warming Potential (GWP)⁹⁰, measured in kg-CO₂-equivalent, is a damage-oriented indicator reflecting the effect of greenhouse gas emissions, considered here for a period of 100 years.

The Eco-Indicator-99 (EI-99)⁹¹ provides an aggregated score (i.e., EI-99 points) for the end-point impact categories of human health, ecosystem-quality and resource consumption.

The required cradle-to-process gate data for the environmental impact of the production of solvents are obtained either from the Ecoinvent database⁹² or estimated using the state-of-the-art FineChem tool⁹³ from molecular descriptors. The FineChem tool provides estimations with an average error of 30-40%, but is very useful in the case of data gaps, especially for the design of new molecules.

The EHS hazard categories represent the safety hazards regarding accidental release of energy or material, the health aspects associated with the long term effects on workers due to workplace exposure and the environmental hazards referring to the damage to flora and fauna. Koller et al.⁹⁴ used a prioritized list of physical and chemical substance properties to estimate a comprehensive set of characteristic dangerous properties for each hazard category. In the **safety category** the properties include mobility, fire and explosion and acute toxicity effects. In the **health category** properties include irritation and chronic toxicity effects. In the **environmental category** properties include air and water mediated effects as well as persistency and the potential for accumulation in the food chain.

Note that the environmental index of the solvent also refers to a part of the environmental impact of the reclaimer waste, if landfilled (e.g., sludge including the part of the solvent that is not recycled by the reclaimer). The toxicity and corrosion indices used in the EHS characterize indirectly the reclaimer waste as hazardous or not (and thus requiring special treatment). The LCA indices (i.e., CED, GWP and EI99) refer to the environmental impact of the make-up solvent, i.e., per kg of solvent. The LCA relevant reclaimer impact is partly described by the solvent physical and hazard properties, but also from the amount of purge stream, which depends on the amount and type of degradation products not included in this study. Furthermore, energy related gate-to-gate sustainability aspects are represented by many of the indicators used for solvent design and selection (e.g. boiling point and vapour pressure of the solvent affect not only the reboiler duty but also the reclaimer energy demands).

A comprehensive sustainability assessment should also include the degradation potential of the solvent under the capture process conditions, as well as the type of degradation products. These issues are directly associated with the amount of solvent make-up (i.e., the complementary factor of the “per-kg of solvent” LCA indices) and indirectly with the environmental burden of the reclaimer (i.e., because of the purged amount treated by the reclaimer in order to avoid accumulation of the degradation products in the main capture system). Moreover, certain degradation products (e.g., nitrosamines) are associated with additional health hazards to those described by the solvent inherent health indices. However, there is severe lack of mechanistic studies relating degradation potential to the molecular structure of the solvent. Some rules have been proposed for instance, requiring the presence of secondary amines for the formation of stable nitrosamines, or supporting the use of tertiary amines (typically in blends because of their slower absorption kinetics) because they have much lower nitrosation rates. However, they are too broad and qualitative to play a useful role in a CAMD approach.

3.1.4 Summary of property criteria at Stage 1

Based on the criteria identified in Sections 3.1.1 to 3.1.3, an optimal solvent at the end of Stage 1 can be interpreted as a solution of the following multi-objective optimization problem:

$$\begin{aligned}
 &\max \quad \rho, pK_a \\
 &\min \quad C_p, P_{vp}, \sigma, n, EHS, CED, GWP, EI99, RED \\
 &\text{s.t.} \quad T_m < T_{Abs} \\
 &\quad \quad T_{bp} > T_{Des}
 \end{aligned} \tag{4}$$

where T_{Abs} and T_{Des} are the average temperatures of the absorption and desorption columns (313K and 393K). The inequalities for the two temperatures represent minimum requirements to avoid solvent solidification or significant vaporization at processing conditions. Based on equation (1) and the property models listed in Tables B1 and B2 of the Supporting Information, the total number of $J_{i,k}$ values that can be obtained for each solvent is 252.

3.1.5 Formulation of Step 1.1

In the case of the design of new solvents (Figure 1), the first task (Step 1.1) is to derive a set of Pareto-optimal solvents from the molecular search space. The solvent performance criteria used during this CAMD step are calculated exclusively from GC models, hence this step accounts only for properties which may be calculated using such models. Uncertainty is not considered in this step and the following models are used from Tables B1 and B2 of the Supporting Information: 1.2, 2.1, 3.1, 4.2, 5.1, 6.1, 7.1 and 8.1. Group-contribution models are not available for pK_a, EHS, CED, GWP and $EI99$ so these are not taken into account in this step of the methodology. These properties are calculated and considered for further evaluations of the Pareto optimal solvents in step 1.2. Given the property models chosen for Stage 1, the multi-objective optimization problem is thus reduced to:

$$\begin{aligned}
& \max \quad \rho \\
& \min \quad C_p, P_{vp}, \sigma, n \\
& \text{s.t.} \quad T_m < T_{Abs} \\
& \quad \quad T_{bp} > T_{Des} \\
& \quad \quad RED \leq RED^{up}
\end{aligned} \tag{5}$$

Note that RED is used as a constraint during CAMD, with RED^{up} being an upper limit which may be set to 1 as noted previously or to a higher value, for example using MEA as a reference point. The RED of MEA is 3.94 using model 1.2 of Table B1 of the Supporting Information. The multi-objective formulation in problem (5) is solved to identify the set of non-dominated optimum solvents using the approach of Papadopoulos and Linke⁷³.

3.2 Selection of relevant models for Stage 2

At Stage 2 (Step 2.2), more detailed models are required to estimate mixture behaviour and in particular the absorption capacity of the candidate solvents. Most models of the thermodynamic behaviour of reactive mixtures of aqueous solvents and CO₂ require extensive experimental data on the specific solvent of interest. However, more predictive models have recently been derived from the SAFT^{95, 96} equation of state. A systematic approach to the modelling of the thermodynamics of CO₂ chemisorption solvents based on limited data has been developed for monoethanolamine (MEA)⁹⁷ and *n*-alkylamines⁹⁸, and further extended by Rodriguez et al.⁹⁹ using the SAFT-VR EoS applied to various aqueous mixtures of alkanolamines and carbon dioxide. To enhance the predictive capabilities of this approach, models to predict the phase behaviour and thermophysical properties of aqueous alkanolamine mixtures with CO₂ using the GC formalism embedded within the SAFT- γ SW framework^{67, 68} have been proposed⁶⁹⁻⁷¹. This provides a much firmer basis for the design of novel solvents, enabling the design of molecular structure and mixture composition to be varied. Thus, in Step 2.1, the fluid-phase behaviour of mixtures of the candidate solvents with carbon dioxide and water is evaluated using the SAFT- γ SW EoS, provided that the relevant building blocks are available in the corresponding group

contribution table (cf. Section D of the Supporting Information). Both phase and chemical equilibria are taken into account simultaneously via this model, enabling a more detailed assessment of the selected solvents. Partial pressure profiles are used to assess the relative performance of the solvents in terms of absorption potential. Saturated vapour pressures and densities can also be evaluated with the approach.

3.3 The molecular search space

The final specification required to enable the application of our approach to CO₂ capture solvents is the definition of the search space. The search space for solvent design is defined by a set of 13 functional groups, namely CH₃-, -CH₂-, >CH-, >C<, -OH, -CH₂-NH₂, -CH₂-NH-, -CH₂-N<, >CH-NH₂, >CH-NH-, CH₃-NH-, CH₃-N<, ≥C-NH₂. These groups are chosen due to the availability of the input data required for the calculation of pure-component molecular properties through GC methods and their frequent occurrence in existing CO₂ capture solvents.

Using these 13 groups as building blocks, a database of existing molecules is also developed by conducting a broad search for acyclic, aliphatic and hydroxylamines available in an in-house data repository at ETH Zurich, in publicly available databases^{100, 101} and in the commercial catalogue of Sigma-Aldrich¹⁰². The search results in a database of 126 amines and hydroxylamines, each with a unique CAS registry number.

The database is partitioned into two classes: firstly, the Reference class (R), containing 25 solvents previously considered extensively in CO₂ capture in the open literature (Table E1 in Supporting Information), serves as a benchmark to assess the viability of the new capture solvents proposed in our work and to assess the potential performance improvements; secondly, the Commercial class (C) containing 101 acyclic, aliphatic amines for which we had limited or no knowledge regarding their suitability as CO₂ capture solvents, prior to this work (Tables E2, E3 and E4 in Supporting Information). The C-class solvents are assumed to be commercially available because they were retrieved from public and commercial databases. The database is not

exhaustive but provides a varied set of amines with different structural characteristics (e.g., primary, secondary, tertiary, branching at different positions).

4. Results and discussion

The proposed systematic approach is applied in Selection mode, using classes R and C, and in Design mode, using the set of 13 functional groups. In this section, we present the top solvents identified after Stages 1 and 2, and analyse the results with the aim to address two questions:

1. How successful is the proposed approach at identifying realistic solvents for CO₂ capture?
2. Which molecules are most promising for further investigation as CO₂ capture solvents?

4.1 Selection of top candidate solvents from the Reference and Commercial classes

4.1.1 Stage 1 of solvent selection

In the Selection problem, the matrix \mathbf{J}' of property indices includes all solvents in the R and C classes and is therefore of size 126 x 252. The application of Steps 1.2 and 1.3 results in a matrix of ranked lists, \mathbf{L} ; the molecules that occur most frequently in the top 15 solvents are selected from \mathbf{L} . This yields a set of 29 candidates, including 9 from class R and 20 from class C. The 29 solvents are presented in Tables 4-6, arranged based on structural characteristics, to reflect the qualitative structure-property relations of Table 1.

Table 4- Amines with [OCCN] or [CCCN] patterns. The square brackets give a hydrogen-free representation of the molecular formula, with atoms in parentheses indicating branching that starts from the previous atom. Each molecule is further labelled by an acronym (e.g., AMP) and a label that indicates its class and number within the class (e.g., R2, for the second solvent in class R). Where an acronym is shown in italics, this indicates that the compound has previously been considered as a potential CO₂ capture solvent in at least one publication or patent. The rows indicate different types of amines (primary, secondary or tertiary). In each row, molecular size increases from left to right. Further details are

provided in Tables E1-E4 and section F of the Supporting Information together with a discussion of the compounds in terms of their structural characteristics and likely CO₂ capture potential based on literature.

NH ₂	-	[Insert R1_Table4.skc here] 2AP (R1) [OCC(C)N]	[Insert C1_Table4.skc here] 2A1B (C1) [OCC(CC)N]	[Insert C2_Table4.skc here] 2A1PN (C2) [OCC(CCC)N]	[Insert C3_Table4.skc here] 2A1H (C3) [OCC(CCCC)N]
		[Insert R2_Table4.skc here] AMP (R2) [OCC(C)(C)N]			
NH	[Insert R3_Table4.skc here] MMEA (R3) [OCCNC]	[Insert R4_Table4.skc here] EMEA (R4) [OCCNCC]	[Insert C4_Table4.skc here] P̄AE (C4) [OCCNCCC]		[Insert R5_Table4.skc here] BEA (R5) [OCCNCCCC]
	[Insert C5_Table4.skc here] 1M2P (C5) [OC(C)CNC]	[Insert C6_Table4.skc here] IPAE (C6) [OCCNC(C)C]	[Insert C7_Table4.skc here] DIBA (C7) [CCC(C)NC(C)CC]	[Insert C8_Table4.skc here] DsBA (C8) [CC(C)CNCC(C)C]	
N	[Insert R6_Table4.skc here] DMMEA (R6) [OCCN(C)(C)]	[Insert R7_Table4.skc here] DEEA (R7) [OCCN(CC)(CC)]	-	[Insert C9_Table4.skc here] DPE (C9) [CCCN(CC)(CCC)]	-

Table 5- Amines with [OCCCN] and [CCCCN] patterns. The nomenclature is as for Table 4. Further details are provided in Tables E1-E4 of the Supporting Information.

NH ₂	[Insert R8_Table5.skc here] MPA (R8) [OCCCN]	[Insert C10_Table5.skc here] 4A2B (C10) [OC(C)CCN]	-	-
NH	-	-	[Insert C11_Table5.skc here] DBA (C11) [CCCCNCCCC]	-

	[Insert C12_Table5.skc here] 3DAP (C12) [OCCCN(CC)(CC)]	[Insert R9_Table5.skc here] DEAB (R9) [OC(C)CCN(CC)(CC)]	[Insert C13_Table5.skc here] ND1B (C13) [CCCCN(CC)(CC)]	[Insert C14_Table5.skc here] 1EDB (C14) [CCCC(CC)N(C)(C)]
N	-----			
	-	-	[Insert C15_Table5.skc here] TMEDA (C15) [N(C)(C)CCN(C)(C)]	[Insert C16_Table5.skc here] 2P12P (C16) [N(C)(C)C(C)(C)CN(C)(C)]

Table 6- Amines with longer carbon chains. The nomenclature is as for Table 4. Further details are provided in Tables E1-E4 of the Supporting Information material.

NH ₂	[Insert C17_Table6.skc here] 4AP (C17) [OCCCC(C)N]	[Insert C18_Table6.skc here] 5AP (C18) [OCCCCCN]	[Insert C19_Table6.skc here] HEXA (C19) [CCCCCCN]
N	[Insert C20_Table6.skc here] 4D1B (C20) [OCCCCN(C)(C)]	-	-

MEA, which was included in Class R, is not found amongst the top solvents due to its low environmental health and safety performance (it exhibits a high *EHS* value). When sustainability indices are not considered MEA appears in the top 15 solvents. Solvents in class R have been investigated to a wider extent in published literature compared to C class solvents with respect to their CO₂ absorption capacity and/or reaction kinetics as well as other properties (Table 7). Additional details are provided in Sections E and F of the Supporting Information.

Table 7- Indicative literature sources and properties investigated for selected class R amines

Solvent	Source	Investigated properties
2AP (R1)	Da Silva ¹⁰³ Fernandes et al. ¹⁰⁴	Carbamate stability Protonation constant, standard molar enthalpy and entropy
AMP (R2)	Sartori et al. ²² Harbou et al. ¹⁰⁵ Zheng et al. ¹⁰⁶	Chemical behaviour due to steric hindrance Process behaviour (e.g., reboiler duty, plant operation) Reaction kinetics
MMEA (R3)	Ma'mun et al. ¹⁰⁷ Suda et al. ¹⁰⁸ Ali et al. ¹⁰⁹	Absorption rate Absorption rate, mass transfer coefficient Reaction kinetics
EMEA (R4)	Kumar ¹¹⁰ Yamada et al. ¹⁹	Equilibrium solubility Absorption capacity
BEA (R5)	Ma'mun et al. ¹⁰⁷ Yamada et al. ¹⁹ Ali et al. ¹⁰⁹	Absorption rate Absorption capacity Reaction kinetics
DMMEA (R6)	Tong ¹¹¹ Versteeg et al. ³	Equilibrium solubility Reaction kinetics
DEEA (R7)	Sutar et al. ¹¹² Xu et al. ¹¹³	Equilibrium solubility and reaction kinetics Equilibrium solubility and heat of absorption
MPA (R8)	Henni et al. ¹¹⁴ Dong et al. ¹¹⁵	Reaction kinetics Equilibrium solubility
DEAB (R9)	Sema et al. ^{11, 13} Maneeintr et al. ¹⁴ Naami et al. ¹⁵ Shi et al. ¹⁶	Equilibrium solubility and reaction kinetics Equilibrium solubility and cyclic capacity Mass transfer coefficient in packing material Vapour liquid equilibrium model

An examination of the solvents in class C shows that:

- Nine of the twenty class C solvents are potentially novel capture options: very limited or no evidence that these molecules have been considered for application to CO₂ capture could be found in the open literature. The performance of these solvents for each property is shown in Figure 3, relative to MEA.
- Solvents 2A1B, 2A1PN and 2A1H have been considered in mixtures with methyldiethanolamine (MDEA) where they were tested experimentally in terms of stability and corrosiveness with respect to carbon steel¹¹⁶. The tests were motivated by the discovery that an aqueous mixture comprising a tertiary alkanolamine and a primary

alkanolamine with a secondary carbon atom attached to the amino group (such as 2A1B, 2A1PN and 2A1H) is not only effective in removing acid gases, but it also exhibits unexpectedly low degradation, corrosiveness and metal solubility properties. However, very few details are disclosed on these compounds, hence they may be worth of further investigation in mixtures with MDEA and other amines of similar structure.

- Solvents DIBA, DsBA, DBA and HEXA are phase-change solvents^{21, 117, 118} associated with significant reductions in the desorption energy requirements because they can be partly separated in a non-thermal process after absorption and before entering the desorption column. This class of solvents contains also several other options¹¹⁸, which are very promising for use either in their pure aqueous form or in mixtures. A review of other types of phase-change solvents is available in Wang and Li¹¹⁹.
- Solvents PAE, IPA¹²⁰, TMEDA¹²¹ and 5AP⁷⁻⁹ have been investigated considerably less than the other solvents and mainly with respect to their CO₂ absorption capacity hence they may be worth of further investigation.

Note that the solubility parameter values of solvents 4A2B, 3DAP, ND1B, 1EDB and 2P12P are within 14-17 (J/cm³)^{1/2} which is the range of solubility parameter values exhibited by the biphasic solvents HEXA, DsBA and DBA^{7-9, 21}. These solvents are therefore expected to exhibit low solubility in water and may also be biphasic candidates. Finally, the EHS index shows clearer, more favourable trends than the LCA metrics (Table C1a and C1b in the Supporting Information) for these solvents. In this regard, 1EDB exhibits much lower acute and chronic toxicity values than the rest of the solvents in this set, and while DPE, 4A2B and 4D1B have higher acute toxicity indices, they are still less than half the MEA index value. DPE, 4A2B, 2P12P and 4D1B, however, all have flash points lower than 35 °C.

[Insert Figure3.tif here]

Figure 3- Comparative performance with respect to MEA (dashed lines) of commercial molecules which have not been traced (or for which data may be rare) as CO₂ capture solvents. The blue bars to the left of the dashed lines indicate better performance than MEA. The red bars to the right of the dashed lines indicate worse performance than MEA.

4.1.2 Stage 2 of solvent selection

In Step 2.1, average performance indices are computed for all 29 solvents and the top 10 solvents are selected for further exploration. Two rankings are derived: one based on all properties, shown in Figure 4a, and one excluding the sustainability indices, shown in Figure 4b. The results are reported using MEA as the reference solvent. These figures highlight some interesting trade-offs between different properties:

- V_m , C_p , $EI-99$: the performance of almost all solvents deteriorates compared to MEA. V_m reflects density hence all solvents exhibit lower density than MEA.
- P_{vp} and CED : the performance of all solvents is slightly better or slightly worse than MEA in Figure 4a, while it deteriorates when sustainability indices are not considered (Figure 4b).
- σ , RED , n , pKa : the performance of almost all solvents improves compared to MEA. When sustainability is not considered (Figure 4b), σ and n of the selected solvents appear to be better than for the solvents in Figure 4a, while pKa appears to be slightly worse.
- EHS , and GWP : the performance is worse when sustainability indices are not considered, but better performance than MEA is observed in most cases.

[Insert Figure4a.tif here]

[Insert Figure4b.tif here]

Figure 4- Comparative results for top 10 solvents selected based on (a) set of equations (4), (b) set of equations (4) but without considering sustainability indices EHS , CED , GWP , $EI-99$. The results are

standardized with respect to MEA (shown by a dashed vertical line). The blue bars to the left of the dashed lines indicate better performance than MEA. The red bars to the right of the dashed lines indicate worse performance than MEA. Solvents in the reference class (R) are reported first, followed by solvents in the commercial class (C), regardless of their ranking in the top 10.

As seen in Figure 4, 2AP, MMEA, EMEA and MPA are repeated in both cases, indicating that they are not affected by the variation in the employed performance indices. DBA is the only solvent from the C class that is shared between the two sets in Figure 4. The consideration of sustainability indices has a significant effect on the selection of solvents from the C class. The evaluation of the selected solvents based on mixture behaviour (Step 2.2) will be considered in Section 4.3, where solvents that have been found via Selection or Design are compared.

4.1.3 Performance of selected solvents

The solvents listed in Figure 4 are first assessed using literature sources for prior evidence of their potential as CO₂ capture options. This is useful as part of the prioritisation process, but also in the context of assessing the efficacy of the proposed approach to solvent selection and design. Where we have been unable to identify prior use of a solvent as a CO₂ capture option, the qualitative structure-property relations listed in Table 1 are used. Further analysis for all 29 solvents identified at the end of Stage 1 is provided in section F of the Supporting Information. A summary of the main performance characteristics of the solvents and relevant literature sources is presented in Table 8. From the top 10 solvents, those that have previously been investigated in the literature present very favourable characteristics compared to MEA. This indicates that the selection procedure points toward useful CO₂ capture options, despite the use of simple property prediction models.

Table 8- Main performance characteristics and literature sources for top 10 solvents of Figure 4a.

Solvent		Main characteristics	Source
2AP	(R1)	Closest similar structure is MIPA which presents very similar solubility as MDEA at the same conditions but has higher pK_a than MDEA.	Rebolledo-Morales et al. ¹²²
MMEA	(R3)	Higher absorption rate and capacity than MEA.	Ma'mun ¹⁰⁷
EMEA	(R4)	Higher CO ₂ loading than MEA, DEA and MMEA (at 30% w/w) over a wide range of pressures.	Kumar ¹¹⁰
MPA	(R8)	Higher reaction rate and absorption capacity than MEA. In mixtures with AMP it exhibits considerably reduced corrosiveness, oxidative degradation and formation of nitrosamines than mixtures of MEA and AMP.	Henni et al. ¹¹⁴ Dong et al. ¹¹⁵ Hoff et al. ¹²³
DEAB	(R9)	Designed specifically for high CO ₂ capacity, reaction constant similar to AMP and DEA but higher than MDEA and much lower regeneration energy than MEA.	Sema et al. ¹¹ Maneeintr et al. ¹⁴
2A1B	(C1)	Exhibiting high stability and low corrosiveness in mixtures with MDEA.	Rooney ¹¹⁶
2A1PN	(C2)	Similar to 2A1B.	Rooney ¹¹⁶
1M2P	(C5)	Not traced as a CO ₂ capture solvent. Structure is similar to MMEA.	-
IPAE	(C6)	Slightly higher absorption capacity but lower reaction rate than the similar PAE which exhibits a high absorption capacity, similar to EMEA and BEA.	Yamada et al. ^{19, 120}
DBA	(C11)	A biphasic solvent exhibiting phase change and regeneration at 90°C, hence of much lower energy requirements than usual.	Zhang et al. ²¹
4AP	(C17)	Not traced as a CO ₂ capture solvent. Structure is similar to 2AP.	-

Focusing on sustainability performance, the EHS index generally shows clearer trends than the LCA metrics (Tables C2a and C2b in the Supporting Information) for the top 10 solvents. DMMEA, DEEA, DBA and HEXA are flammable with flash points lower than 50 °C. DEEA is a recognized skin sensitizer. HEXA also exhibits aquatic toxicity that is an order of magnitude higher than MEA. 2A1PN, 2AIB, IPAE, 4AP, DEAB are predicted to have lower acute toxicity

than MEA due to their relatively high oral lethal dosage (LD50oral) values. On the other hand, MEA is characterized by higher toxicity indices due to its low Immediately Dangerous to Life and Health (IDLH) threshold value generated by the National Institute for Occupational Safety and Health (NIOSH). This indicator, however, is not available for the rest of the molecules in the dataset.

4.2 Design of top candidate solvents from functional groups

4.2.1 Stage 1 of solvent design

Using the molecular design space given by the chosen functional groups, the solution of problem (5) leads to the identification of 157 molecular structures on the Pareto front. The molecular structures are defined by the number of groups of each type in the molecule, without full connectivity information, so that each molecular structure may correspond to several isomers. In this case, 2492 isomers can be constructed from the 157 structures identified.

The resulting database of designed structures (class D) is analyzed using the same strategy as classes C and R. The matrices \mathbf{J}' and \mathbf{L} are thus of size 157 x 252. The calculation of the thermodynamic properties is carried out using first-order functional groups except for the pK_a that can only be calculated using a specific structure as input in the Marvin 6.0.5⁸⁶ software. For this purpose, we use the isomers exhibiting the highest pK_a value as the representative for each solvent structure. As in the case of solvent selection, top structures are identified using the criteria set out in formulation (4): five structures are obtained with the use of the sustainability indices, and five without.

The set of candidates identified in this way consists of the 10 molecular structures (listed in Table 9) appearing most frequently with top ranks, and representing 114 isomers. The structures are reported in the form of functional group combinations with their frequency of appearance in the molecule because this is how they are designed using the CAMD method. The results in Table 9 provide a first indication of the group combinations which may favour CO₂ capture.

Table 9- Top molecular structures obtained from the class of designed solvents (D).

ID	Groups in molecule	Frequency of groups
D1	[-CH ₃ , >CH ₂ , ->C-NH ₂ , -CH ₂ -NH-, -CH ₂ -N<]	[4 1 1 1 1]
D2	[-CH ₃ , >CH ₂ , >C<, -OH, ->C-NH ₂]	[4 1 1 1 1]
D3	[-CH ₃ , >CH ₂ , ->C-NH ₂ , -CH ₂ -NH-, CH ₃ -N<]	[3 1 1 1 1]
D4	[-CH ₃ , >CH ₂ , ->C-NH ₂ , -CH ₂ -NH-]	[3 2 1 2]
D5	[-CH ₃ , >CH ₂ , ->C-NH ₂ , CH ₃ -NH-, -CH ₂ -NH-]	[2 1 1 1 1]
D6	[-CH ₃ , >CH ₂ , >CH-, -CH ₂ -NH-, -CH ₂ -N<,]	[5 1 1 1 2]
D7	[-CH ₃ , >CH ₂ , >CH-, -OH, ->C-NH ₂]	[3 2 1 1 1]
D8	[-CH ₃ , -OH, ->C-NH ₂ , -CH ₂ -NH-]	[2 1 1 1]
D9	[-CH ₃ , >CH ₂ , -OH, -CH ₂ -NH-]	[1 2 1 2]
D10	[-CH ₃ , >CH ₂ , -OH, ->C-NH ₂ , -CH ₂ -NH-]	[2 1 1 1 1]

Structures D1-D5 are obtained using equations (4) without considering the sustainability indices.

Structures D2, D7-D10 are obtained using the entire set of equations (4). Some of the structures obtained are multifunctional amines containing up to 3 amine groups while some contain only a single amine group. The fact that only D2 is common to both sets of structures shows that the consideration of sustainability plays an important role in assessing the performance of the designed solvents. Structure D6 is not in the top 5 structures of the case examined in Figure 5a, due to undesirably high C_p . It is reported here because one of its isomers is a novel molecule which closely resembles Bis-(3-dimethylaminopropanol) (TMBPA)²⁰. This is a solvent containing 3 amine groups and that has been found experimentally to exhibit very favourable CO₂ capture features in mixtures with piperazine, including 70% higher cyclic capacity and 41% higher CO₂ removal than a 5M MEA solution²⁰.

The performance of structures D1-D10 relative to MEA is presented in Figure 5. Except for *EHS*, all the sustainability indices point to lower performance than MEA. This is a reasonable outcome because the sustainability indices were not considered during CAMD in Step 1.1, but as a subsequent measure to identify the solvents with lower impacts among those obtained from

CAMD. In this case (Figure 5b) *EHS* and *EI99* improve but *CED* and *GWP* deteriorate slightly, possibly due to the effect of the remaining property indices.

[Insert Figure5a.tif here]

[Insert Figure5b.tif here]

Figure 5- Comparative performance with respect to MEA of novel solvents designed using the proposed CAMD approach: a) Solvents D1-D5 were obtained using the set of equations (4) without considering the sustainability indices (although their values are reported here), b) solvents D2, D7-D10 were obtained using the entire set of equations (4). TMBPA (E7) is a commercial solvent previously tested as a CO₂ capture option²⁰. D6 was not in the top choices but it is reported here to compare its performance to that of E7 since the structure of a D6 isomer is very similar to E7. Figure 5a compares solvents without considering sustainability hence these properties are not reported for E7 and D6.

Furthermore, the apparent tradeoff between the EHS and LCA indices could be attributed to the fact that molecules which are structurally larger and more complex than MEA would have longer production chains with higher environmental impacts, while they could be more stable with higher boiling points and thus lower mobility index values. The mobility index is also used as a correction factor for the toxicity and flammability indices to account for the relative ease of vapour formation promoting the risk of inhalation or catching fire. All the molecules in Table 9 show improved toxicity indices and slightly better flammability indices due to the improved mobility indices. However, they suffer from worse environmental indices such as aquatic toxicity and persistency in the environment. This could be a result of the more complex structure that is less naturally occurring and thus less prone to degradation in the environment. Tables C3a and C3b in the Supporting Information present, respectively, the LCA metrics and the various hazard indices and the parameters used to calculate them for the molecular structures listed in Table 9.

4.2.2 Stage 2 of solvent design

The main task in Step 2.1 is to identify specific molecules from the set of isomers represented in Table 9. To this end, we exploit the qualitative Structure-Property Relations (1) and (2) reported

in Table 1 to identify isomers that are likely to exhibit favourable CO₂ capture performance. The structures are transformed into the generic, hydrogen free representations of [OCCN], [OCCCN] (Structure-Property Relation (1)) and [NC...CN] with up to 6 carbon atoms (Structure-Property Relation (2)) which can then be processed computationally. The proposed motifs are sought systematically in the structures of the isomers using the publically available Small Molecules Subgraph Detector software¹²⁴ which employs a maximum common subgraph method¹²⁵ to identify similarities between molecular structures. Note that Structure-Property Relation (2) was considered by Aronu et al.²⁰ who started from a structure similar to TMEDA (C15) and heuristically identified TMBPA (E7) for further testing as a CO₂ capture solvent. We adopt a similar rational approach for the selection of isomer candidates but we combine two structure-property relations in our search. The results are shown in Table 10.

Table 10- Isomers corresponding to structures of Table 9 based on Structure-Property Relations (1) and (2). IUPAC names have been assigned using Marvin 6.0.5⁸⁶. Data are available for AEEA in Table E1 and for the other solvents in Tables E5 and E7 of the Supporting Information material.

[Insert D1_Table10.skf here] D1: [3-amino-4-(dimethylamino)-3-methylbutyl] (methyl)amine	[Insert D3_Table10.skf here] D3: [3-amino-3-(dimethylamino) butyl] (methyl)amine
[Insert D4_Table10.skf here] D4: {3-[(2-amino-2-methylpropyl)amino] propyl}(methyl) amine	[Insert D5_Table10.skf here] D5: [3-amino-3-(methylamino) butyl](methyl)amine
[Insert D6_Table10.skf here] D6: (3-{[1-(dimethylamino)propan-2-yl] amino}propyl)dimethylamine	[Insert E7_Table10.skf here] E7: Bis-(3-dimethylaminopropanol) (TMBPA)
[Insert D2_Table10.skf here] D2: 4-amino-2,4-dimethylpentan-2-ol (CAS: 91875-44-4)	[Insert D7_Table10.skf here] D7: 3-amino-2,3-dimethylpentan-1-ol
[Insert D8_Table10.skf here] D8: 2-amino-1-(methylamino)propan-2-ol	[Insert D10_Table10.skf here] D10: 2-amino-4-(methylamino)butan-2-ol
[Insert D9_Table10.skf here] D9: 2-{[2-(methylamino)ethyl] amino} ethan-1-ol (CAS: 85771-07-9)	[Insert R10_Table10.skf here] R10: AEEA

4.2.3 Analysis of designed solvents

Among the solvents in Table 10, isomer D6 is clearly very similar to TMBPA, indicating that the CAMD method has delivered a potentially highly performing molecule despite its complex structure. Isomers D1, D3, D4 and D5 present a side chain with a longer carbon backbone. Structure-Property Relation (2) has been applied in the form of [NCCCN] for at least one pair of nitrogen atoms in each molecule. These molecules also contain a quaternary carbon atom bonded to a primary amine group, which introduces moderate steric hindrance to the molecule and resembles the structure of AMP. The existence of 3 amine groups with moderate steric hindrance is expected to increase the absorption capacity, while the appearance of at least one primary group and one secondary group is expected to ensure reasonable kinetic performance, assuming that a carbamate will always be formed.

Isomers D2, D7, D8, D9 and D10 all contain a hydroxyl. D2 is a commercially available structure with a CAS registry number of 91875-44-4 and is the solvent shared between the two cases investigated in Figure 5. It can be said to exhibit steric hindrance based on Structure-Property Relation (6) of Table 1. These molecules all contain one or two primary or secondary amine groups; they are combined with a quaternary carbon atom hence they also resemble AMP, except for D9 which is a simpler hydroxyl-diamine. In isomers D2, D7, D8 and D10, Structure-Property Relation (1) appears as [NCCO] or [NCCCO], while Structure-Property Relation (2) also appears in D10. D9 is a commercially available amine with a CAS number of 85771-07-9. The structure is very similar to AEEA (shown next to D9 in Table) which exhibits absorption capacity, CO₂ reactivity, and energy efficiency higher than those of MEA while it has also low vapour pressure¹²⁶.

4.3 SAFT-based evaluation of selected solvents (Step 2.2)

The sets of selected and designed solvents are now considered jointly for further analysis. The SAFT- γ SW EoS is used to predict the phase behaviour of all solvents in these sets that can be

modelled with the approach using the current set of group interactions available (Section D of Supporting Information). These are MMEA, EMEA, MPA, 1M2P, DBA and D9. Comparisons with experimental data, where available, and with the performance of MEA are provided.

Several authors have measured the solubility of CO₂ in aqueous mixtures of MMEA¹²⁷⁻¹²⁹ and EMEA¹¹⁰ at conditions of chemical and vapour-liquid equilibrium. In Figures 6a and 6b, we compare SAFT- γ SW predictions for MMEA mixtures and EMEA mixtures, respectively, with experimental data. The agreement with experiments is good considering these calculations are purely predictive and that the group parameters have been characterised based on data for other alkanolamines at lower loadings⁷¹. In particular, the relative solubility of CO₂ in the different solvents is found to be reliable. Furthermore, both model and experiments indicate that CO₂ loading (cf. Figure D1 in Supporting Information) decreases with solvent concentration at moderate to high pressures, but increases with solvent concentration at low pressures. The qualitative agreement achieved provides a useful basis for solvent selection. We note that the accuracy of the model could be improved by developing group interaction parameters that account for the proximity effects between the groups of relevance to MMEA and EMEA, as has been done successfully for other groups⁷⁰⁻⁷¹.

[Insert Figure6.tif here]

Figure 6- Partial pressure of CO₂ as a function of mole fraction of CO₂ in aqueous solutions of alkanolamine+ H₂O + CO₂. a) MMEA at 303.1 K. The symbols correspond to experimental data^{110, 127} for aqueous solutions of MMEA of various weight percentages in mass: 6.8 wt% (green), 11 wt% (red), 14 wt% (blue) and 19 wt% (orange). b) EMEA at 303.1 K. The symbols correspond to experimental data¹¹⁷ for aqueous solutions of EMEA of various weight percentages in mass: 6 wt% (green), 12 wt% (red), 18 wt% (blue) and 24 wt% (orange) and 30 wt% (purple). The solid curves correspond to SAFT- γ SW calculations for the same water:solvent ratios.

A further comparison with experimental data can be seen in Figure 7 for the solubility of CO₂ in aqueous solutions of MPA. Data for this mixture at two temperatures ($T=313.15$ K and 393.15 K) were used in the estimation of a number of group parameters⁷⁰ and therefore only the calculations at 343.15 K and 373.15 K can be seen as predictions. SAFT- γ is found to provide a very good description of the effect of temperature on the solubility of CO₂ in this case.

[Insert Figure7.tif here]

Figure 7- Solubility of CO₂ in a 4 M (~30 wt% in mass) aqueous solutions of MPA at $T = 313.15$ K (blue), 343.15 K (red), 373.15 K (purple) and 393.15 K (green) as a function of the partial pressure of CO₂ along the vapour-liquid equilibrium of the ternary mixture MPA + H₂O + CO₂. The solubility is represented as CO₂ loading θ_{CO_2} , defined as the moles of CO₂ absorbed in the liquid phase per mole of amine in the liquid. The symbols correspond to experimental data¹¹⁵ for the corresponding temperatures. The solid curves correspond to SAFT- γ SW calculations.

A comparison of the predicted performance of SAFT- γ for a number of high-performing solvents is presented in Figure 8. These solvents are compared here based on the solubility of CO₂ in aqueous solutions of a given mass percentage. At a low temperature ($T = 313.15$ K) typical of an absorption process, there is a range of pressures at which all solvents are predicted to reach higher solubility loadings than MEA for a given partial pressure of carbon dioxide. The range of pressures at which this occurs is wider for DBA and D9. Solvents such as MMEA, EMEA and 1M2P present very similar solubility curves; this can be attributed to the similarities in their molecular structures (see Table 4). At a high temperature ($T = 393.15$ K) typical of a desorption process, and at solubility loadings lower than 0.5, most of the solvents appear to outperform the regeneration capacity of MEA, as they lead to partial pressures of CO₂ higher than those obtained with MEA. At high loadings, MEA leads to only slightly higher partial pressures than MMEA, EMEA, 1M2P and MPA.

The performance predictions obtained for the designed or selected solvents with SAFT- γ SW therefore confirm that these solvents should be investigated further for their CO₂ capture potential.

[Insert Figure8.tif here]

Figure 8- Solubility of CO₂ in 30 wt% (in mass) aqueous solutions of MPA (blue), MMEA (red), EMEA (green), 1M2P (purple), MEA (black), DBA (orange) and D9 (grey) at a) $T = 313.15$ K and b) 393.15 K as a function of the partial pressure of CO₂ along the vapour-liquid equilibrium of each ternary mixture alkanolamine + H₂O + CO₂. The solubility is represented as CO₂ loading θ_{CO_2} , defined as the moles of CO₂ absorbed in the liquid phase per mole of amine in the liquid. The solid curves correspond to SAFT- γ SW calculations.

5. Conclusions

A systematic two-stage approach to molecular design has been proposed and applied to the discovery of CO₂ capture solvents. The approach is based on multiple criteria, is applicable to the identification of molecules from databases or functional groups, combines simple group contribution models with the molecular-based SAFT framework, and takes model uncertainty into account. It has been used to generate a prioritised list of candidate molecules from hundreds of thousands of possible compounds.

The proposed methodology has clearly led to the identification of important candidate solvents for CO₂ capture, including novel compounds. Several of these have been tested experimentally in published literature and recommended as highly performing alternatives to MEA, indicating that the proposed approach leads to the rapid identification of useful compounds. Others have received little attention or have yet to be considered as alternatives for CO₂ capture despite their commercial availability indicated by their CAS registry number, demonstrating that the proposed approach can help to generate novel leads. For example in the cases of 2A1B, 2A1PN and 2A1H

only a patent exists that mentions very good stability and reduced corrosiveness compared to MEA when they are used in mixtures with MDEA; however, no equilibrium or reaction kinetic data are reported. In several cases the structural and property resemblance of these candidates with state-of-the-art solvents recently considered as CO₂ capture options provides significant motivation to undertake further experimental testing. Without prior intention or use of any specific constraints the selection procedure also resulted in the identification of several thermomorphic biphasic solvents. Such solvents enable desorption at temperatures much lower than the 120°C employed in the MEA case, hence requiring a significantly lower regeneration load.

The novel solvents designed using the proposed CAMD approach combine structural characteristics that are known empirically to promote both high CO₂ absorption capacity and increased reaction rates. D6 is a novel structure proposed by CAMD which is very similar to TMBPA, a multi-functional amine identified in the published literature as a solvent of high CO₂ absorption capacity and reaction rate in mixtures with piperazine. The isomers corresponding to D2 and D9 are commercially available solvents. This is clear evidence that the systematic CAMD approach employed in this work points towards promising CO₂ capture options, despite the initial use of simple property prediction models. This is further confirmed by the phase and chemical equilibria predictions obtained with the SAFT- γ SW equation of state. For example, D9 exhibits very high performance as it actually outperforms all the other amines in the commercial set for a wide range of pressures. The fact that such an assessment of absorption and desorption potential can be made without recourse to any experimental data on the candidate solvents, and indeed without ever synthesizing the proposed molecules, is a significant benefit of the use of a predictive model such as SAFT- γ SW, relative to approaches based on more qualitative rules for solvent selection.

Thanks to the use of sustainability metrics in different stages of the methodology, a clear opportunity has been identified to develop less hazardous solvents than MEA. However, the LCA metrics do not always point in the same direction as safety metrics, and the improvement potential with respect to MEA is limited to specific solvents and metrics. Of course, it has to be noted that these LCA metrics refer only to the life cycle impact of solvent production and, therefore, the extent of solvent fugitive losses and degradation has been neglected in this analysis, despite its known importance in the case of MEA. Additionally, other more significant life cycle impacts (e.g., the solvent regeneration life cycle impact) are indirectly covered by the thermodynamic properties considered in CAMD.

In our current work, we have shown that considering multiple and diverse criteria in the design of solvents using a CAMD method can lead to novel molecules. Building on this approach, the space of molecules to be explored and the range of criteria for the *in silico* assessment of candidate compounds could both be extended. For instance, the consideration of higher order groups in the group contribution models employed could increase the accuracy of the property predictions and allow the more reliable identification of the most suitable isomers as CO₂ capture options. Predictive group contribution property models that provide accurate caloric information⁵⁶ could also be used to achieve a more direct assessment of energy consumption and operating cost. Furthermore, the proposed uncertainty quantification method could be expanded to account for additional properties, especially those associated with sustainability. Sustainability metrics in the form of group contribution models combined with uncertainty quantification could also be considered during CAMD to enable a more focused search as well as to increase the robustness of the obtained results. Since the results obtained from CAMD also provide insights into the structural combinations that favour CO₂ capture, this could be used to prioritize the efforts to develop new group contribution parameters for the SAFT- γ SW equation, in the form of new functional groups, or of combinations of existing groups to form second-order groups and

increase prediction accuracy. Such developments can provide further guidance for experimental studies and help to accelerate the search for better solvents.

Supporting Information

Details are provided as Supporting Information regarding all the molecular structures considered together with a discussion of the compounds in terms of their structural characteristics and likely CO₂ capture potential based on literature sources

Disclaimer

Certain commercial equipment, instruments, or materials (or suppliers, or software) are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Acknowledgements

The authors are grateful to the Commission of the European Union (project FP7-ENERGY-2011-282789) and the Engineering and Physical Sciences Research Council (EPSRC) of the UK (grants EP/E016340, EP/J014958/1 and EP/J003840/1) for financial support.

6. References

1. N. Mac Dowell, N. Florin, A. Buchard, J. Hallett, A. Galindo, G. Jackson, C.S. Adjiman, C.K. Williams, N. Shah and P. Fennell, *Energy Environ. Sci.*, 2010, **3**(11), 1645-1669.
2. F. Bougie and M.C. Iliuta, *J. Chem. Eng. Data*, 2012, **57**, 635-669.
3. G.F. Versteeg, L.A.J. Van Dijck, and W.P.M. Van Swaaij, *Chem. Eng. Commun.*, 1996, **144** (1), 113-158.
4. D.M. D'Alessandro, B. Smit and R.J. Long, *Angew. Chem., Int. Ed*, 2010, **49**, 6058-6082.

5. I. Kim and H.F. Svendsen, 2016, 6th Korea CCS Conference, Jan 27-29, Jeju-do, S. Korea, http://www.koreaccs.or.kr/esub02_2/data/down/year/2016/gubun/MOUTH/page/3/id/1093/nu/2 (accessed May 2016)
6. J. Oexmann and A. Kather, *Int. J. Greenhouse Gas Control*, 2010, 4 (1), 36-43.
7. P. Singh, J.P.M. Niederer and G.F. Versteeg, *Int. J. Greenhouse Gas Control*, 2007, 1, 5-10.
8. P. Singh, D.W.F. Brilman and M.J. Groeneveld, *Energy Procedia*, 2009, 1(1), 1257-1264.
9. P. Singh, J.P.M. Niederer and G.F. Versteeg, *Int. J. Greenhouse Gas Control*, 2009, 87, 135-144.
10. P. Singh and G.F. Versteeg, *Process Saf. Environ. Prot.*, 2008, 86, 347-359.
11. T. Sema, A. Naami, Z. Liang, R. Idem, P. Tontiwachwuthikul, H. Shi, P. Wattanaphan and A. Henni, *Chem. Eng. Sci.*, 2008, 81, 251-259.
12. P. Tontiwachwuthikul, A.G.H. Wee, R. Idem, K. Maneeintr, G.-J. Fan, A. Veawab, A. Henni, A. Aroonwilas and A. Chakma, *US Pat.*, 7 910 078, 2011.
13. T. Sema, A. Naami, Z. Liang, R. Idem and P. Tontiwachwuthikul, *Ind. Eng. Chem. Res.*, 2011, 50 (24), 14008-14015.
14. K. Maneeintr, R. Idem, P. Tontiwachwuthikul and A.G.H. Wee, *Energy Procedia*, 2009, 1, 1327-1334.
15. A. Naami, M. Edali, T. Sema, R. Idem and P. Tontiwachwuthikul, *Ind. Eng. Chem. Res.*, 2012, 51 (18), 6470-6479.
16. H. Shi, T. Sema, A. Naami, Z. Liang, R. Idem and P. Tontiwachwuthikul, *Ind. Eng. Chem. Res.*, 51 (25), 8608-8615.
17. L. Espinal, D.L. Poster, W. Wong-Ng, A. J. Allen and M. L. Green, *Environ. Sci. Technol.*, 2013, 47 (21), 11960-11975.
18. G. Puxty, R. Rowland, A. Allport, Q. Yang, M. Bown, R. Burns, M. Maeder and M. Attalla, *Environ. Sci. Technol.*, 2009, 43, 6427-6433.

19. H. Yamada, F.A. Chowdhury, Y. Matsuzaki, K. Goto, T. Higashii, S. Kazama, *Energy Procedia*, 2013, **37**, 499-504.
20. U.E. Aronu, H.F. Svendsen and K.A. Hoff, in *Proceedings of Distillation Absorption 2010*, A.B. de Haan, H. Kooijman H. and A. Górak, 2010, 151-156.
21. J. Zhang, Y. Qiao, D.W. Agar, *Energy Procedia*, 2012, **23**, 92-101.
22. G. Sartori, W.S. Ho, D.W. Savage, G.R. Chludzinski and S. Wlechert, *Sep. Purif. Rev.*, 1987, **16** (2), 171-200.
23. L.E.K. Achenie, R. Gani and V. Venkatasubramanian, *Computer Aided Chemical Engineering, Volume 12: Computer aided molecular design: Theory and practice*, Elsevier, 2002.
24. R. Gani, in *Chemical Product Design: Toward a Perspective through Case Studies*, ed. K.M. Ng, R. Gani and K. Dam-Johansen, Elsevier, 2007, 14, 435–458.
25. L.Y. Ng, F.K. Chong and N.G. Chemmangattuvalappil, *Comput. Chem. Eng.*, 2015, **81**, 115-129.
26. C.S. Adjiman, A. Galindo, E.N. Pistikopoulos, M.C. Georgiadis and V. Dua, *Process Systems Engineering: Volume 6: Molecular Systems Engineering*, Wiley-VCH, Germany, 2010.
27. R. Gani, *Comp. Chem. Eng.*, 2004, **28**, 2441–2457.
28. K.J. Kim and U.M. Diwekar, *Ind. Eng. Chem. Res.*, 2002, **41**(18), 4479-4488.
29. K.J. Kim and U.M. Diwekar, *Ind. Eng. Chem. Res.*, 2002, **41**(18), 1285–1296.
30. W. Xu and U.M. Diwekar, *Ind. Eng. Chem. Res.*, 2005, **44** (18), 7138–7146.
31. M. Folić, C.S. Adjiman and E.N. Pistikopoulos, *AIChE J.*, 2007, **53**, 1240–1256.
32. C. D. Maranas, *AIChE J.*, 1997, **43** (5), 1250–1264.
33. A.I. Papadopoulos, M. Stijepovic, P. Linke, P. Seferlis and S. Voutetakis, *Ind. Eng. Chem. Res.*, 2013, **52**, 12116-12133.

34. A.S. Hukkerikar, B. Sarup, A. Ten Kate, J. Abildskov, G. Sin and R. Gani, *Fluid Phase Equilib.*, 2012, **321**, 25–43.
35. A.S. Hukkerikar, S. Kalakul, B. Sarup, D.M. Young, G. Sin and R. Gani, *J. Chem. Inf. Model.*, 2012, **52**(11), 2823-2839.
36. C.C. Chen, J.F. Boston, H.I. Britt and L.B. Evans, *AIChE J.*, 1982, **23**, 588-596.
37. Y. Song and C.C. Chen, *Ind. Eng. Chem. Res.*, 2009, **48**, 7788–7797.
38. Y. Zhang, H. Que, and C.C. Chen, *Fluid Phase Equilib.*, 2011, **311**, 67-75
39. H. Lepaumier, D. Picq and P.L. Carrette, *Ind. Eng. Chem. Res.*, 2009, **48**, 9061-9067.
40. H. Lepaumier, D. Picq and P.L. Carrette, *Ind. Eng. Chem. Res.*, 2009, **48**, 9068-9075.
41. H. Lepaumier, S. Martin, D. Picq, D. Delfort and P.L. Carrette, *Ind. Eng. Chem. Res.*, 2010, **49**, 4553-4560.
42. F.J.G. Retief, MSc. Thesis, Stellenbosch University, South Africa, 2012.
43. F. Porcheron, A. Gibert, P. Mougin and A. Wender, *Environ. Sci. Technol.*, 2011, **45**(6), 2486-2492.
44. S. Bommareddy, N.G. Chemmangattuvalappil, C.C. Solvason and M.R. Eden, *Comput. Chem. Eng.*, 2010, **34** (9), 1481-1486.
45. N.G. Chemmangattuvalappil and M. R. Eden, *Ind. Eng. Chem. Res.*, 2013, **52** (22), 7090-7103.
46. J. Salazar, U. Diwekar, K. Joback, A.H. Berger and A.S. Bhow, *Energy Procedia*, 2013, **37**, 257-264.
47. ASPEN Plus®, 2015, <http://www.aspentech.com/products/engineering/aspens-plus/> (accessed December 2015).
48. H. Renon and J. M. Prausnitz, *AIChE J.*, 1968, **14**, 135–144.
49. A. Fredenslund, R. L. Jones and J. M. Prausnitz, *AIChE J.*, 1975, **21**, 1086–1099.
50. K.Z. Sumon, A. Henni and A.L.L. East, *Ind. Eng. Chem. Res.*, 2012, **51**, 11924-11930.

51. E.F. Da Silva and H.F. Svendsen, *Int. J. Greenhouse Gas Control*, 2007, **1** (2), 151-157,
52. I. Stanescu and L.E.K. Achenie, *Chem. Eng. Sci.*, 2006, **61**, 6199–6212.
53. H. Struebing, Z. Ganase, P.G. Karamertzanis, E. Sioukrou, P. Haycock, P. M. Piccione, Armstrong A., Galindo A. and Adjiman C.S., *Nat. Chem.*, 2013, **5**, 952–957.
54. A. Gil-Villegas, A. Galindo, P.J. Whitehead, S.J. Mills, G. Jackson and A.N. Burgess, *J. Chem. Phys.*, 1997, **106** (10), 4168-4186.
55. A. Galindo, L.A. Davies, A. Gil-Villegas and G. Jackson, *Mol. Phys.*, 1998, **93** (2), 241-252.
56. F.E. Pereira, E. Keskes, A. Galindo, G. Jackson and C.S. Adjiman, *Comput. Chem. Eng.*, 2011, **35** (3), 474-491.
57. V. Papaioannou, T. Lafitte, C. Avendaño, C.S. Adjiman, G. Jackson, E.A. Müller and A. Galindo, *J. Chem. Phys.*, 2014, **140**, 054107.
58. J. Burger, V. Papaioannou, S. Gopinath, G. Jackson, A. Galindo and C.S. Adjiman, *AIChE J.*, 2015, **61** (10), 3249–3269.
59. A. Bardow, K. Steur and J. Gross, *Ind. Eng. Chem. Res.*, 2010, **49**, 2834–2840.
60. B. Oyarzún, A. Bardow and J. Gross, *Energy Procedia*, 2011, **4**, 282-290.
61. M. Stavrou, M. Lampe, A. Bardow and J. Gross, *Ind. Eng. Chem. Res.*, 2014, **53** (46), 18029-18041.
62. M. Lampe, M. Stavrou, J. Schilling, E. Sauer, J. Gross and A. Bardow, *Comput. Chem. Eng.*, 2015, **81**, 278–287.
63. J. Gross and G. Sadowski, *Ind. Eng. Chem. Res.*, 2001, **40** (4), 1244-1260.
64. J. Gross, *AIChE J.*, 2005, **51** (9), 2556-2568.
65. J. Gross and J. Vrabec, *AIChE J.*, 2006, **52** (3), 1194-1204.
66. M. Kleiner and J. Gross, *AIChE J.*, 2006, **52** (5), 1951-1961.
67. A. Lympieriadis, C.S. Adjiman, G. Jackson and A. Galindo, *Fluid Phase Equilib.*, 2008, **274**, 85-104.

68. A. Lymeriadis, C.S. Adjiman, A. Galindo and G. Jackson, *J. Chem. Phys.*, 2007, **127** (23), 234903.
69. A. Chremos, E. Forte, V. Papaioannou, A. Galindo, G. Jackson and C.S. Adjiman, *Chem. Eng. Trans.*, 2013, **35**, 427-432.
70. A. Chremos, E. Forte, V. Papaioannou, A. Galindo, G. Jackson and C.S. Adjiman, *Fluid Phase Equilib.*, 2016, **407**, 280-297.
71. A. Chremos, A. Galindo, G. Jackson and C.S. Adjiman, Predictive SAFT- γ SW models for application in the development in absorption processes for carbon capture, Unpublished work, 2016.
72. W. Mendenhall, *Statistics for the engineering and computer sciences*, Dellen Pub. Co., San Francisco, 2nd edition, 1988.
73. A.I. Papadopoulos and P. Linke, *AIChE J.*, 2006, **52** (3), 1057-1069.
74. K.A. Hoff, E.F. Da Silva, I. Kim, A. Grimstvedt and S. Ma'mun, *Energy Procedia*, 2013, **37**, 292-299.
75. J.H. Hildebrand, *J. Am. Chem. Soc.*, 1929, **51**, 66-80.
76. C.M. Hansen, *J. Paint Technol.*, 1967, **39**, 104-117.
77. E. Stefanis and C. Panayiotou, *Int. J. Thermophys.*, 2008, **29** (2), 568-585.
78. C. M. Hansen, *Prog. Org. Coat.*, 2004, **51** (1), 77-84.
79. J. Oexmann, PhD Thesis, Technischen Universität Hamburg-Harburg, Germany, 2011.
80. A.V. Rayer, A. Henni and P. Tontiwachwuthikul, *Can. J. Chem. Eng.*, 2012, **90**, 367-376.
81. B.I. Dvorak, D.F. Lawler, J.R. Fair and N.E. Handler, *Environ. Sci. Technol.*, 1996, **30** (3), 945-953.
82. J. Marrero and R. Gani, *Fluid Phase Equilib.*, 2001, **183**, 183-208.
83. Y. Mergler, R. Rumley-van Gurp, P. Brassler, M. de Koning and E. Goetheer, *Energy Procedia*, 2011, **4**, 259-266.

84. U.E. Aronu, H.F. Svendsen, K.A. Hoff and O. Juliussen, *Energy Procedia*, 2009, **1**, 1051–1057.
85. D.D. Perrin, B. Dempsey and E.P. Sergeant, *pKa prediction for organic acids and bases*, Springer, Netherlands, 1981.
86. Marvin 6.0.5, <http://www.chemaxon.com>, (accessed April 2015).
87. J. Szegezdi and F. Csizmadia, presented at 27th ACS National Meeting, Anaheim, California, USA, March 28-April 1, 2004.
88. F. Csizmadia, A. Tsantili-Kakoulidou, I. Panderi and F. Darvas, *J. Pharm. Sci.*, 1997, **86** (7), 865-871.
89. Verein-Deutcher-Ingenieure, *Vdi-richtlinie 4600: Cumulative Energy Demand, Terms, Definitions, Methods of Calculation*, Beuth Verlag GmbH, Berlin, 1997.
90. Intergovernmental Panel on Climate Change (IPCC), 2007, Climate change 2007 synthesis report, https://www.ipcc.ch/pdf/assessment-report/ar4/syr/ar4_syr.pdf (accessed February 2016)
91. M. Goedkoop and R. Spriensma, The eco-indicator 99: a damage orientated method for life-cycle impact assessment, Methodology report 2000a, Pre-Consultants, https://www.pre-sustainability.com/download/EI99_annexe_v3.pdf (accessed February 2016).
92. Ecoinvent database, <http://www.ecoinvent.org/database/database.html> (accessed February 2016).
93. G. Wernet, S. Papadokonstantakis, S. Hellweg and K. Hungerbühler, *Green Chem.*, 2009, **11**, 1826-1831.
94. G. Koller, U. Fischer and K. Hungerbuehler, *Ind. Eng. Chem. Res.*, 2000, **39** (4), 960-972
95. W.G. Chapman, K.E. Gubbins, G. Jackson and M. Radosz, *Fluid Phase Equilib.*, 1989, **52**, 31-38.

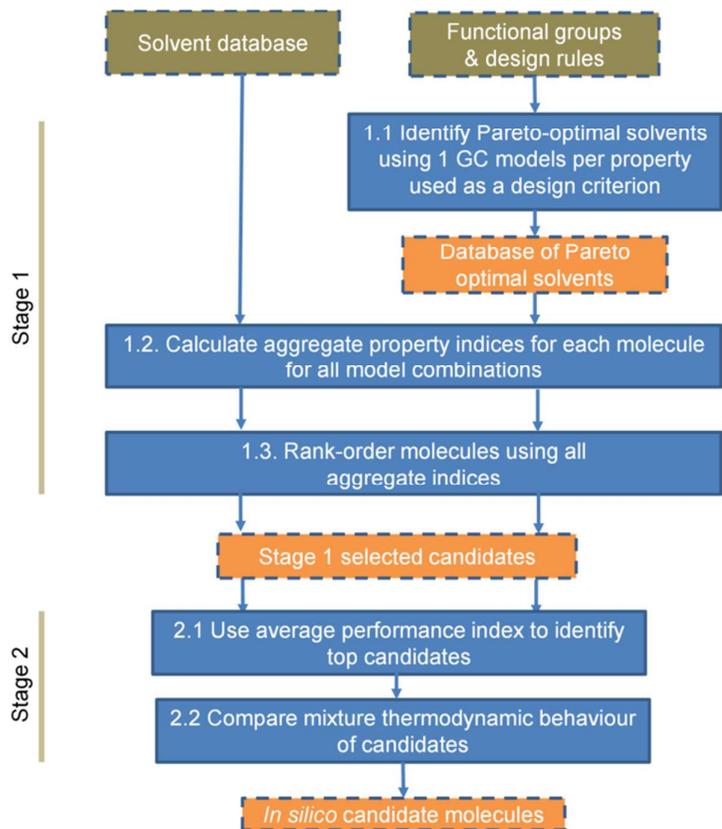
96. W.G. Chapman, K.E. Gubbins, G. Jackson and M. Radosz, *Ind. Eng. Chem. Res.*, 1990, **29** (8), 1709-1721.
97. N. Mac Dowell, F. Llovell, C.S. Adjiman, G. Jackson and A. Galindo, *Ind. Eng. Chem. Res.*, 2010, **49**, 1883-1899.
98. N. Mac Dowell, F.E. Pereira, F. Llovell, F.J. Blas, C.S. Adjiman, G. Jackson and A. Galindo, *J. Phys. Chem. B*, 2011, **115** (25), 8155-8168.
99. J. Rodriguez, N. Mac Dowell, F. Llovell, C.S. Adjiman, G. Jackson and A. Galindo, *Mol. Phys.*, 2012, **110**, 1325-1348.
100. NIST (National Institute of Standards) Chemistry Webbook <http://webbook.nist.gov/chemistry> (accessed February 2016).
101. EPA (Environmental Protection Agency), <http://www.epa.gov> (accessed February 2016).
102. Sigma-Aldrich (Commercial catalogue), <http://www.sigmaaldrich.com/catalog>, (accessed February 2016).
103. E.F. Da Silva, *Energy Procedia*, 2011, **4**, 164-170.
104. D. Fernandes, W. Conway, X. Wang, R. Burns, G. Lawrance, M. Maeder and G. Puxty, *J. Chem. Thermodyn.*, 2012, **51**, 97-102.
105. I. Harbou, H.P. Mangalapally and H. Hasse, *Int. J. Greenhouse Gas Control*, 2013, **18**, 305-314.
106. C. Zheng, B. Zhao, K. Wang and G. Luo, *AIChE J.*, 2015, **61**, 4358-4366.
107. S. Ma'mun, PhD Thesis, NTNU, Norway, 2005.
108. T. Suda, M. Iijima, H. Tanaka, S. Mitsuoka and T. Iwaki, *Environ. Prog.*, 1997, **16**, 200-207.
109. S.H. Ali, S.Q. Merchant, M.A. Fahim, *Sep. Purif. Technol.*, 2002, **27** (2), 121-136.
110. G. Kumar, PhD Thesis, National Institute of Technology, Rourkela, Odisha, India, 2013.
111. D. Tong, PhD Thesis, Imperial College London, UK, 2012.

112. P.N. Sutar, P.D. Vaidya and E.Y. Kenig, *Chem. Eng. Sci.*, 2013, **100**, 234–241
113. Z. Xu, S. Wang, G. Qi, A.A. Trollebø, H.F. Svendsen, C. Chen, *Int. J. Greenhouse Gas Control*, 2014, **29**, 92-103.
114. A. Henni, J. Li and P. Tontiwachwuthikul, *Ind. Eng. Chem. Res.*, 2008, **47**, 2213-2220.
115. L. Dong, J. Chen and G. Gao, *J. Chem. Eng. Data*, **55**, 2010, 1030–1034.
116. P.C. Rooney, *US Pat.* 6 165 432 A, 2000.
117. J. Zhang, Y. Qiao and D.W. Agar, *Chem. Eng. Res. Des.*, 2012, **90**, 743–749.
118. J. Zhang, PhD Thesis, Technical University of Dortmund, Germany, 2013.
119. X. Wang and B. Li, in *Novel Materials for Carbon Dioxide Mitigation Technology*, ed. F. Shi and B. Morreale, Elsevier, Amsterdam, 3-22. 2015.
120. H. Yamada, F.A. Chowdhury, K. Goto and T. Higashii, *Int. J. Greenhouse Gas Control*, 2013, **17**, 99-105
121. Z. Bouzina, A. Negadi, I. Mokbel, J. Jose and L. Negadi, presented at AICHE Annual Meeting, Pittsburgh, USA, November, 2012.
122. M.G. Rebolledo-Morales, M.E. Rebolledo-Libreros and A. Trejo, *J. Chem. Thermodyn.*, 2011, **43** (5), 690-695.
123. A.K. Hoff, T. Mejdell, I. Kim, A. Grimstvedt and E.F. Silva, *WIPO Pat.* WO 2014/086 988 A1, 2014.
124. S. Asad Rahman, M. Bashton, G. L. Holliday, R. Schrader and J. M. Thornton, *J. Cheminform.*, 2009, **1** (12), 1-13.
125. D.J. Cook and L.B. Holder, *Mining Graph Data*, Wiley & Sons Inc., New York, 2006.
126. M. Stec, A. Tatarczuk, D. Spiewak and A. Wilk, *J. Solution Chem.*, 2014, **43**, 959-971.
127. G. Kumar, M. Kundu, *Can. J. Chem. Eng.*, 2012, **90** (3), 627-630.
128. R. Pacheco, A. Sanchez, M.D. La Rubia, A.B. Lopez, S. Sanchez, F. Camacho, *Ind. Eng. Chem. Res.*, 2012, **51** (13), 4809-4818.

129. H.A.M. Haider, R. Yusoff and M.K. Aroura, *Fluid Phase Equilib.*, 2011, **303** (2), 162-167.

Table of contents entry

We propose to use Figure 1 of the paper the graphic on the table of contents entry.



Regarding the 20 words we propose the following:

“We propose an optimization-based Computer Aided Molecular Design method for chemisorption-based CO₂ capture and identify novel solvents”