European Symposium on Computer Aided Process Engineering – 15
L. Puigjaner and A. Espuña (Editors)

A computer-aided methodology for optimal solvent design for reactions with experimental verification

Milica Folić, Claire S. Adjiman* and Efstratios. N. Pistikopoulos

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

Abstract

An extension of a hybrid experimental / computer-aided methodology for the design of solvents for reactions is presented. Previous work (Folić *et al.*, 2004a,b) was based on the use of reaction rate measurements to build a reaction model, followed by the formulation and solution of an optimal computer-aided molecular design problem (CAMD). In this work, feedback is introduced in the methodology to verify the suitability of the solvent candidates identified in the CAMD step via experimentation and to assess the reliability of the model used in the CAMD step. When the reliability of the model is found to be insufficient, experimental data for the candidate solvents are added to the original data set to create an updated reaction model which can be used to find new candidate solvents. This methodology is illustrated through application to a solvolysis reaction and to a Menschutkin reaction.

Keywords: solvent design, optimisation, solvatochromic equation, group contribution methods, reaction rate

1. Introduction

Solvents are widely used as a reaction medium in the fine chemicals industry, where they serve to bring solid reactants together by dissolving them, to control temperature, and to enhance reaction rate. The effect of solvent choice on reaction rate can be dramatic. Reichardt (1988) reports that the solvolysis of 2-chloro-2-methylpropane is 335,000 times faster in water than in ethanol, while the reaction between trimethylamine and trymethylsulfonium ion is 119 times faster in nitromethane than in water. In spite of the importance of solvent choice on productivity, there has been little work on systematic approaches to the selection of solvents for reactions. Thus, industry currently relies mostly on experience and intuition to guide its choice during the development of new processes.

This situation is in striking contrast with the selection of solvents for separation, where several computer-aided design approaches have been proposed in the last two decades. Several of these methods are described in Achenie *et al.* (2003). These have been successfully applied to a variety of solvent-based separation problems, allowing a much larger number of solvent molecules to be considered during separation system design

* Author to whom correspondence should be adressed : c.adjiman@imperial.ac.uk

than is possible by experimentation alone. Based on these considerations, the goal of this work is develop a systematic approach to solvent design for reactions.
The basic idea behind the methodology presented here is that, in the context of reactions, it is especially important to rely on a combination of experiments and computer-aided molecular design (CAMD). The computations serve as a guide to the experiments, focussing the search on promising solvents, and the experiments allow a verification of the models used. Furthermore, the methodology is developed with a view to plant-wide solvent selection, where it is important to focus on overall performance rather than the performance of single process units. This motivates the use of an optimisation-based approach to CAMD, where trade-offs between different aspects of the process can be accounted for explicitly. The methodology is described in section 2, and is applied to two reactions in section 3.

2. Methodology

The overall methodology proposed in this work is illustrated in Figure 1. For a given reaction, eight initial solvents are chosen. These solvents should be selected to be

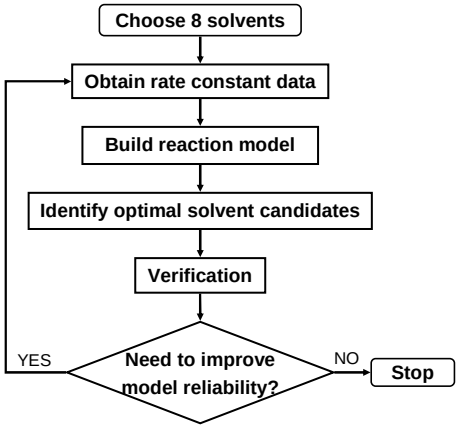

Figure 1. Overview of the solvent design methodology

diverse in terms of the types of interactions they can have with the species involved in the reaction. In general, the E_T^N solvent polarity scale (Reichardt and Harbusch-Görnert, 1983). In addition, solvents with different functional groups are typically chosen. Wherever possible, literature data should be used at this stage to minimise experimental costs. Experimental reaction rate constants for these eight solvents are obtained. This information is then used to build a reaction model that predicts the reaction rate constant in other solvents based solely on their molecular structure. Next, a computer-aided solvent design problem (CAMD) is formulated and solved. Here, the objective is to find candidate solvents which give high values of the reaction rate constant. In the verification step, the predicted rate constants for the best candidate solvents identified are compared to experimental rate constants, to determine whether the reaction model needs improvement. If so, the experimental rate constants for the candidate solvents are

added to set of initial solvents to build an updated reaction model. This procedure is repeated until the model reliability is sufficient. The computer-aided design step thus serves two purposes: it identifies promising solvents and it guides experiments. The model building and CAMD steps are briefly discussed in the next sections.

2.1 Building the reaction model

The reaction model, as illustrated in Figure 2, consists of a set of property estimation methods which relate solvent molecular structure to solvent properties, and a solvatochromic equation (Abraham *et al.*, 1987) which relates solvent properties to reaction rate constant for a given reaction.

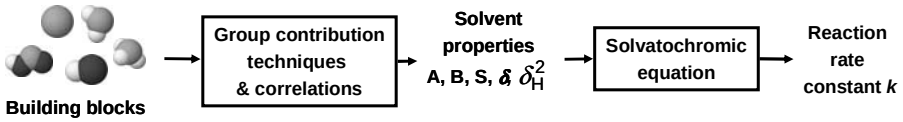

Figure 2. Schematic of the reaction model

Atom groups such as CH_2 and OH are used as building blocks. The solvent properties needed in the solvatochromic equation are the so-called solvatochromic parameters, *A*, *B* and *S*, a polarisability correction term, δ, and the cohesive energy density, which is the square of the Hildebrand solubility parameter, δ_H. The polarisability correction term can be calculated exactly based on molecular structure. The cohesive energy density is estimated through its relation with the molar volume, V_m, and the heat of vaporisation, H_v, as discussed in Sheldon *et al.* (2004). V_m and H_v are estimated using the first-order versions of the group contribution techniques of Constantinou and Gani (1994) and Marrero and Gani (2001) respectively. The group contribution techniques proposed in Sheldon *et al.* (2004) for the hydrogen-bond acidity *A* and the hydrogen-bond basicity *B* and the technique discussed in Folić *et al.* (2004a,b) for the dipolarity/polarisability *S* have been extended in this work. Group contribution coefficients are available for 43 groups, allowing a wider variety of solvent molecules to be represented. The regression has been based on a solvent database, which contains 350 solvents, giving increased confidence in the prediction techniques. The average absolute percentage error for each of the methods is reported in Table 1.

Table 1. Average absolute percentage error (AAPE) for the property estimation methods used to predict solvent properties.

Property	*A*	*B*	*S*	δ_H
AAPE	0.017	0.043	0.065	1.13

The solvent properties are used in the solvatochromic equation:

$$\log k = \log k_o + s(S + d\delta) + aA + bB + h\delta_H^2 \,/\, 100 \qquad (1)$$

where k is the reaction rate constant, and k_o, s, d, a, b and h are reaction-specific parameters. The values of these reaction parameters are obtained via linear regression, based on measurements of the rate constant in a number of different solvents. Here,

eight solvents are used to build the initial reaction model. Since the overall reaction model is based on predictions of the solvent properties, the *predicted* values of A, B, S and δ_H for the eight solvents are used in the regression.

2.2 The computer-aided molecular design problem

Once the reaction model has been developed, it is embedded within a CAMD optimisation problem. This is based on an MINLP formulation of the following form:

$$\begin{aligned} &\max_{n,y} \log k \\ &\text{s.t. } \log k = \log k_0 + s(S + d\delta) + aA + bB + h\delta_H^2 / 100 \\ &\quad \text{property estimation techniques for } A, B, S, \delta, V_m, H_v, T_m, \delta_H \\ &\quad \text{melting point constraint} \\ &\quad \text{molecular complexity constraints} \\ &\quad \text{definition of } n \text{ based on binary variables } y \end{aligned} \tag{2}$$

The constraint on the melting point T_m ensures the solvent designed is liquid at room temperature. The group contribution technique of Constantinou and Gani (1994) is used to estimate T_m. The molecular complexity constraints consist of the octet rule (Odele and Macchietto, 1993), the bonding rule (as modified by Buxton et al., 1999), limits on the combinations of functional groups that may appear in the final molecule, and limits on the total number of groups in the molecule. Finally, the continuous variables n_i which define the number of groups of type i in the optimal molecule are related to binary variables to make sure that they only take on integer values. Nonlinearities arise in the above formulation from the estimation of the cohesive energy density. As a result, this problem is an MINLP which is linear in the binary variables. It can be solved locally with the outer-approximation algorithm (Viswanathan and Grossmann, 1990).

3. Case studies

The case studies reported here are based on two reactions for which relatively large amounts of data are available in the literature. In such a case, it is desirable to complete the first iterations of the methodology using available data, in order to reduce process development time and cost. Such a procedure can then guide the choice of solvents in which to perform new measurements.

3.1 Solvolysis of *t*-butyl chloride

Reaction rate constant data for the solvolysis of *t*-butyl chloride ($CH_3CCl \rightarrow (CH_3)_3C^+$-$Cl^- \rightarrow (CH_3)_3C^+|Solv|Cl^- \rightarrow$ Products) is available in 41 solvents (Abraham, 1972, Abraham et al., 1981, Abraham et al., 1987, Gonçalves et al., 1992, Dvorko et al., 2002). The reaction rate constants reported vary by 11 orders of magnitude and the best experimental solvent is glycerol. The eight diverse solvents selected to build the reaction model are shown in Table 2 with their experimental ranking, where rank 1 denotes the solvent with the largest rate constant. A wide range of polarities and functional groups results in a set which contains both good and poor solvents. Good statistics are obtained for the solvatochromic equation regression: R^2 is 0.93 and the standard error is 1.9. The average absolute percentage error for all 41 solvents is 17%. A

comparison of solvent rankings using experimental data and predictions (Table 3), shows good overall agreement.

Table 2. Solvents for the solvolysis case study, with experimental rank.

Solvent	Rank	Solvent	Rank	Solvent	Rank	Solvent	Rank
1,2-ethanediol	2	2-methyl-2-propanol	4	Diethylene glycol	6	Acetic acid	14
Dimethyl acetamide	29	Chlorobenzene	36	Benzene	38	Pentane	41

Table 3- Comparison of solvent rankings: experiments and predictions.

Solvent	Exp	Pred	Solvent	Exp	Pred
Glycerol	1	1	1,2-ethanediol	2	3
Phenol	3	10	Propane-1,3-diol	4	5
Propane-1,2-diol	5	6	Diethylene glycol	6	4
Butane-1,4-diol	7	7	Triethylene glycol	8	2
Butane-1,2-diol	9	8	Aniline	10	15

The CAMD MINLP identifies glycerol as the best solvent, with a reaction rate constant three times larger than that of 1,2-ethanediol, the best solvent used in the regression. Verification against literature data shows that the rate constant in glycerol has been measured and that it is the best solvent known to date. Given the consistency between the computational and experimental results, the search is stopped.

3.2 A Menschutkin reaction

In this case study, the Menschutkin reaction between methyl iodide and tripropylamine is considered: $(n\text{-}C_3H_7)N + CH_3I \rightarrow (CH_3)(n\text{-}C_3H_7)_3N^+\text{-}I^-$. Reaction rate constant measurements in 59 different solvents can be found in Lassau and Jungers (1968). The range of rate constants reported covers five orders of magnitude and the best ranked solvent is benzyl cyanide. A set of eight diverse solvents for which experimental data are available is chosen: it consists of a cyanide, an alkyltetrahalide, a nitrate, a halosubstituted aromatic, an aromatic, an alcohol, an oxide and an alkane. The solvatochromic reaction parameters are regressed based on these data, giving an R^2 value of 0.85 and a standard error of 1.2. When the predictions for all 59 solvents are compared with the experimental data, the average absolute percentage error is found to be only 19%. A comparison of the experimental and predicted solvent rankings shows that 7 out of the top 10 experimental solvent are predicted in the top 10 and that 17 out of the top 20 experimental solvents are predicted in the top 20. The CAMD MINLP identifies chlorobenzylcyanide as the best solvent. Integer cuts are added to find the second and third best solvents, chlorobenzylnitrate and 1,7-heptanediol. The verification step is then performed. Although no experimental data are available for any of the candidate solvents, data are available for benzylcyanide and chlorobenzene, which have functional groups and structures similar to the top two solvent candidates. Benzylcyanide is already included in the set of eight solvents used to build the reaction model, but the error between prediction and measurement is large, indicating that the

model reliability could be improved. Since none of the eight initial solvents contains a chlorine group, the chlorobenzene data is added to the data set. A new reaction model is regressed based on these nine solvents. The overall statistics are similar to those obtained with eight solvents, but the qualitative ranking is slightly improved, with 8 of the top 10 experimental solvents predicted in the top 10. The CAMD MINLP is solved with the new reaction model and once again yields chlorobenzylcyanide as the top candidate. There is no further data in the data set which may be used to verify this prediction and the measurement of the rate of reaction in chlorobenzylcyanide is now needed.

4. Concluding remarks

A methodology for the systematic design of solvents for reactions has been proposed. It is based on an iterative approach which alternates between experiments and computer-aided molecular design. The reaction model at the core of the CAMD problem is based on the empirical solvatochromic equation, in which the solvent properties are obtained by group contribution techniques and the reaction parameters are regressed from experimental data. The CAMD results are verified against experimental data, and an improved reaction model is generated if needed. This is then used in an updated CAMD problem. The approach has been applied to a solvolysis reaction, in which only one reaction model was used, and to a Menschutkin reaction, in which two reaction models were used. Further verification of the results via experimentation is underway.

References

Abraham, M.H., 1972, J. Chem. Soc. – Perkin Trans. 2, 1343.

Abraham, M.H., R.W. Taft and M.J. Kamlet, 1981, J. Org. Chem. 46, 3053.

Abraham M.H., R.M. Doherty, M.J. Kamlet., J.M. Harris and R.W. Taft, 1987, J. Chem. Soc., Perkin Trans. 2, 913.

Achenie L.E.K., R. Gani and V. Venkatasubramanian, Eds., 2003, Computer Aided Molecular Design: Theory and Practice, Elsevier, Amsterdam.

Buxton A., A.G. Livingston and E.N. Pistikopoulos, 1999, AIChE J. 45, 817.

Constantinou L. and R. Gani, 1994, AIChE J. 40, 1697.

Dvorko, G.F., V.V. Zaliznyi and N.E. Ponomarev, 2002, Russian J. General Chemistry 72, 1549.

Folić M., C.S. Adjiman, E.N. Pistikopoulos, 2004a, Proceedings of ESCAPE-14, Elsevier.

Folić M., C.S. Adjiman, E.N. Pistikopoulos, 2004b, Proceedings of FOCAPD, *in press*.

Gonçalves R.M.C., A.N.M. Simões, R.A.S.E Leitão and L.M.P.C. Albuquerque, 1992, J. Chem. Research (S), 330.

Lassau C. and J.C. Jungers, 1968, Bull. Soc. Chim. Fr. 7, 2678.

Marrero J. and R. Gani, 2001, Fluid Phase Eq. 183-184, 183.

Odele O. and S. Macchietto, 1993, Fluid Phase Eq. 82, 47.

Reichardt C., 1988, Solvents and Solvent Effects in Organic Chemistry, VCM Publishers, UK.

Reichardt C. and E. Harbusch-Görnert, 1983, Liebigs. Ann. Chem., 721.

Sheldon T., C.S. Adjiman and J.L. Cordiner, 2004, Fluid Phase Eq. accepted for publication.

Viswanathan, J. and I.E. Grossmann, 1990, Computers chem. Engng. 14, 769.

Acknowledgements

Financial support from the ORS scheme and CPSE is gratefully acknowledged.