



HAL
open science

In silico Design, Virtual Screening and Synthesis of Novel Electrolytic Solvents

G. Marcou, B. Flamme, G. Beck, A. Chagnes, O. Mokshyna, Dragos Horvath,
A. Varnek

► **To cite this version:**

G. Marcou, B. Flamme, G. Beck, A. Chagnes, O. Mokshyna, et al.. In silico Design, Virtual Screening and Synthesis of Novel Electrolytic Solvents. *Molecular Informatics*, 2019, 38 (10), pp.1900014. 10.1002/MINF.201900014 . hal-02950590

HAL Id: hal-02950590

<https://hal.science/hal-02950590>

Submitted on 13 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOI: 10.1002/minf.200((full DOI will be filled in by the editorial staff))

In Silico design, virtual screening and synthesis of novel electrolytic solvents

G. Marcou*^[a], B. Flamme^[b], G. Beck^[a], A. Chagnes^[c], O. Mokshyna^[a], D. Horvath^[a], A. Varnek^{[a]*}

Abstract: We report the building, validation and release of QSPR (Quantitative Structure Property Relationship) models aiming to guide design of new solvent for the next generation of Li-ion batteries. The dataset compiled from the literature data included oxidation potentials (E_{ox}), specific ionic conductivities (κ), melting points (T_m) and boiling points (T_b) for 103 electrolytes. Each of the resulting consensus models assembled 9-19 individual Support Vector Machine models built on different sets of ISIDA fragment descriptors^[1]. They were implemented in the ISIDA/Predictor software.

Developed models were used to screen a virtual library of 9965 esters and sulfones. The most promising compounds prioritized according to theoretically estimated properties were synthesized and experimentally tested.

Keywords: QSPR, Generative Topographic Mapping, electrolyte, oxidation potential, conductivity

Despite a tremendous success in the development of new materials for positive electrodes of Lithium-ion batteries, still low attention was paid to the design of suitable electrolytes. In principle, new electrodes support high voltage, but their application is limited by the poor electrolyte stability towards oxidation in high-voltage lithium-ion batteries. Besides, security issues have to be managed due to possible exothermic reactions between positive electrodes and electrolytes, especially at high-voltage. In particular, classical electrolytes used in Lithium-ion batteries such as ethylene carbonate (EC)-dimethyl carbonate (DMC)+lithium hexafluorophosphate (LiPF₆) can hardly be used at voltage greater than 4.2 V with appropriate cathodes. Higher voltage may lead to dramatic decrease of the cycle ability of the battery and increases a risk of explosion, fire and release of toxic substances^[2]. Thus, increase of the flash point, the thermal stability and the anodic stability of the solvent while keeping electrolyte viscosity as low as possible and electrolyte ionic conductivity as high as possible are important goals of electrolyte optimization.

In this work, QSPR approach has been used to guide the design of new generation of electrolytic solvents. Particular attention was paid to molecules belonging to two different chemotypes: sulfones and esters/ethers. Although, esters and ethers are largely investigated in the literature, their physicochemical and electrochemical properties still need to be optimized for their use in electrolyte for Lithium-ion batteries. Sulfones remain liquid within a large range of temperature, their thermal behavior is very interesting as they are generally non-flammable and have very high flash points (for instance, the flash point of dimethylsulfone is 145 °C whereas dimethyl carbonate and a mixture of ethylene carbonate:dimethylcarbonate (1:1) exhibit flash points equal to 16 and 25 °C, respectively^[3]).

In this work, 4 key properties^[3-4] were considered: ionic conductivities (κ) and oxidation potentials (E_{ox}) of lithium-based electrolytes as well as melting points (T_m) and boiling point (T_b) of dipolar aprotic organic solvents usually studied for Lithium-ion batteries. The electrolyte

should have advantageous transport properties in order to reduce the ohmic-drop caused by the internal resistance of the battery cell (high ionic conductivity). The electrolyte must be in liquid state for a large range of temperatures (ideally between -40 °C and 120 °C, considering an operating range between -30 °C and 60 °C). Simultaneously, the electrochemical window should be as wide as possible. Currently, oxidation potential of electrolytes reaches about 4-4.2 V at active cathode materials. At high voltage a massive oxidation of the electrolyte may occur, leading to a steep increase of the current density with a rapid loss of battery cycling performance. Thus, it is highly desirable to design new solvents exhibiting high oxidation potential in the presence of lithium salts.

At the first stage, a dataset of 103 dipolar aprotic organic solvents was collected from the literature. This included 155 oxidation potential values measured at 5 mV/s in sulfone or ester solvents in the presence of 1M LiPF₆ or Lithium bis(trifluoromethanesulfonyl)imide (LiTFSI) at platinum, glassy carbon, LiCoO₂, LiMn₂O₄ or LiNi_{1/3}Mn_{1/3}Co_{1/3}O₂ electrodes. Having inspected these data, we decided to select for the modeling 82 values measured at a Pt electrode in the presence of LiPF₆ or LiTFSI, which are the most commonly used salts in lithium-ion batteries. Collected dataset contained also

[a] Faculty of Chemistry - UMR7140, University of Strasbourg
4, rue Blaise Pascal, 67000 Strasbourg, France

e-mail: g.marcou@unistra.fr; varnek@unistra.fr

phone: +33(0)3 68 85 13 04

[b] Ecole Nationale Supérieure de Chimie de Paris,
11 Rue Pierre et Marie Curie, 75005 Paris, France

[c] Université de Lorraine, CNRS, GeoRessources, F-54000 Nancy, France

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/minf>. ((DOI will be filled in by the editorial staff)).

Running title

ionic conductivity values for 55 molecules measured at similar conditions, as well as boiling and melting points for 60 and 47 compounds, respectively.

Many QSPR models for boiling point of organic compounds are reported in the literature [5]. For the most performant models, the accuracy of prediction is no less than 20 °C [6]. Melting point models are also frequent but less precise the error of prediction is about 35 °C [7]. In this context, one may expect that a local model built on selected solvents will lead to better results than those already obtained from available models in the literature.

Only few QSPR studies of the oxidation potential were reported so far. For instance, some empirical relationships between E_{ox} and molecular structure for flavonoids [8] and anilines [9] were reported. Another approach is to approximate E_{ox} by ionization potential. In turn, the latter, according to the Koopman theorem [10] can be approximated by the HOMO energy. The Hammett-like equations for oxidation potential of phenoxide ions [11], anilines [12] and di-substituted benzenes [13] were established. There exist also atomistic physical models for solid state electrolytes [14], although they are yet in qualitative rather than quantitative agreement with the experimental data. However, up to our knowledge, no models of ionic conductivity in dipolar aprotic organic solvents in the presence of lithium salts were reported so far.

To resume, predictive models for ionic conductivity and oxidation potential are needed because existing approaches are either difficult to use or inaccurate. For the boiling and melting points, we expected to obtain more accurate models by focusing specifically on the specific types of solvents.

The models were built using the ISIDA fragment descriptors [1] and Support Vector Regression (SVR) method with linear kernel. Compared to non-linear models, the linear ones are mathematically simpler and their performances are similar. Several individual models based on the ISIDA fragments of particular type were combined in the consensus models, each associated with a fragment control applicability domain (AD) - see details in Computational Methods section. The model performances in cross-validation are given in Table 1. In this table, the R^2 values are given for the sake of completeness. However, since the property ranges in training and test sets differ, retrospective success of the models was assessed on the basis of RMSE values.

Developed models for all 4 properties were then used to annotate a virtual library of enumerated compounds. The library contained initially 9965 esters, ethers, carbonates and sulfones based on pre-defined molecular scaffolds. A first filter was used to remove problematic structures from a synthesis viewpoint and due to security concerns (for instance, peroxydes was removed from the list because of their explosive properties). This reduced the library to 8070 compounds to which the developed models were applied. Calculations shown for 3832 compounds computed properties satisfied required technological specifications: $E_{ox} > 4V$, $T_b > 50$ °C and $T_m < 20$ °C.

Final selection of a set of 35 compounds followed synthetic feasibility requirements. It contained a subset of 12 esters, ethers and carbonate containing compounds (Table S2 in Supplementary Information). These compounds were synthesized or acquired and their boiling points, ionic conductivities and the oxidation

potentials vs Li/Li^+ were measured. Their melting point was systematically observed below -70 °C, the limit of DSC equipment used in this work. During electrochemical experiments, four compounds bearing a gem-dimethyl group in alpha position to acetyl or carbonate function underwent a violent reaction when the lithium salt was added to the solvent. All other compounds were found out of the applicability domain AD of the boiling point and melting point models. Calculations without accounting for AD resulted in poor estimations (RMSE=49 °C for boiling point). Considering the ionic conductivity, only three compounds were found within AD (RMSE=1.0 mS/cm). When estimating the oxidation potential, 5 compounds passed the AD filter (RMSE=0.23 V).

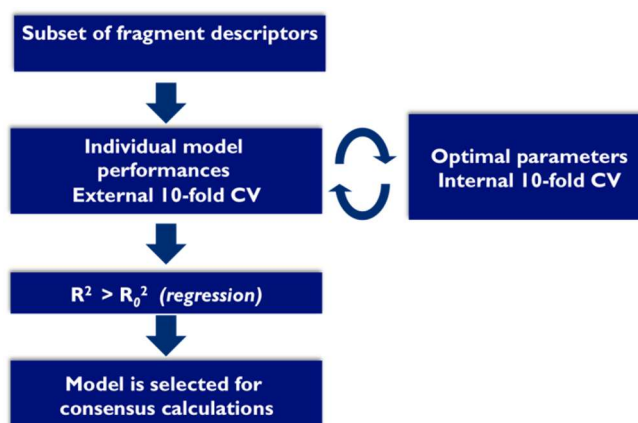


Figure 1. Modeling workflow.

Another subset contained 23 sulfones which were synthesized and tested (Table S1 in Supplementary Information). Most of these compounds were found within models AD. Comparison with the measured experimental data confirmed a good prediction performance of all models except of that for the melting point model. Thus, the prediction error (RMSE) was equal to 45 °C for T_m (for 15 compounds in AD), 14 °C for T_b (21 compounds), 14 °C for T_b (21 compounds), 0.49 mS/cm for κ (20 compounds) and 0.18 V for E_{ox} (20 compounds). Notice that publicly available OCHEM software [6c] predicts T_b of new compounds with much smaller accuracy (RMSE of about 50 °C).

Generally, model performances for new molecules (within AD) are similar or even better than those obtained for the modeling set in cross-validation (Table 1).

The combined dataset including both initial data and novel compounds was analyzed using the Generative Topographic Mapping (GTM) approach [15]. GTM is a dimensionality reduction method allowing to visualize the data distribution on 2-dimensional map. Figure 3 visualizes chemical space of electrolytes. One may see that molecules with different chemotypes form dense clusters: C=O containing compounds are located in the left part of the map, whereas sulfoxide-containing compounds occupy mostly several zones on the right and on the top of the map. One may see that the top/right corner is populated exclusively by sulfolane

Running title
derivatives which have hardly been studied as electrolytes.

The maps were colored according to the oxidation potential and the ionic conductivity (Figure 4). These “property landscapes” [16], help to monitor how the studied properties are distributed in the chemical space. For each property, two landscapes were considered: one corresponding to the modeling data set (on the left) and another one corresponding to the entire dataset including molecules designed in this project (on the right). One may see that new sulfones, in particular sulfolane derivatives, are **extending** previously explored domain regarding both the ionic conductivity and the oxidation potential. Additionally, new esters and ethers change the perception of these families. For instance, the ethers appear less stable to oxidation than it was initially expected.

Table 1. Model performances on the modeling set and on the subset of new molecules designed in this work. Statistical parameters for the modeling set were obtained in 10-fold cross-validation. N_m is the number of data points

Property	Modeling set			New molecules		
	RMSE	R ²	N _m	RMSE	R ²	N _m
Melting Point	32.5°C	0.73	47	45°C	<0	15
Boiling Point	37.0°C	0.77	60	14°C	0.84	21
Ionic conductivity	2.13 mS/cm	0.64	55	0.58 mS/cm	0.94	23
Oxidation Potential	0.45 V	0.68	82	0.18 V	0.91	25

Figure 2. Chemical structures of two best sulfone candidates as electrolytic component.

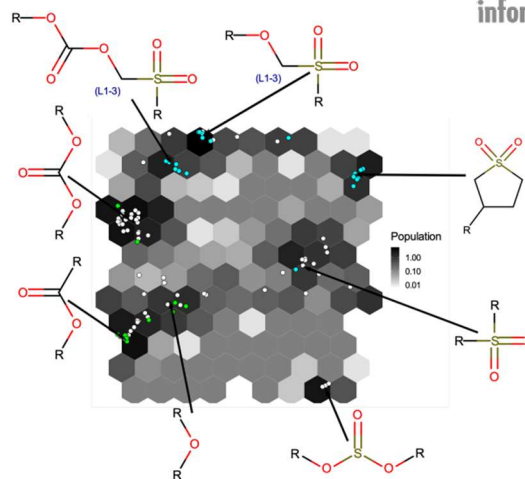


Figure 3. Generative Topographic Map of the electrolyte dataset including training set and novel structures. The grayscale bar indicates the density of compounds. The darker is an area, the more populated it is. White points correspond to compounds collected from literature whereas cyan and green points correspond, respectively, to new sulfones and new carbonates, esters and ethers. Arrows show the areas populated by molecules with particular core structures.

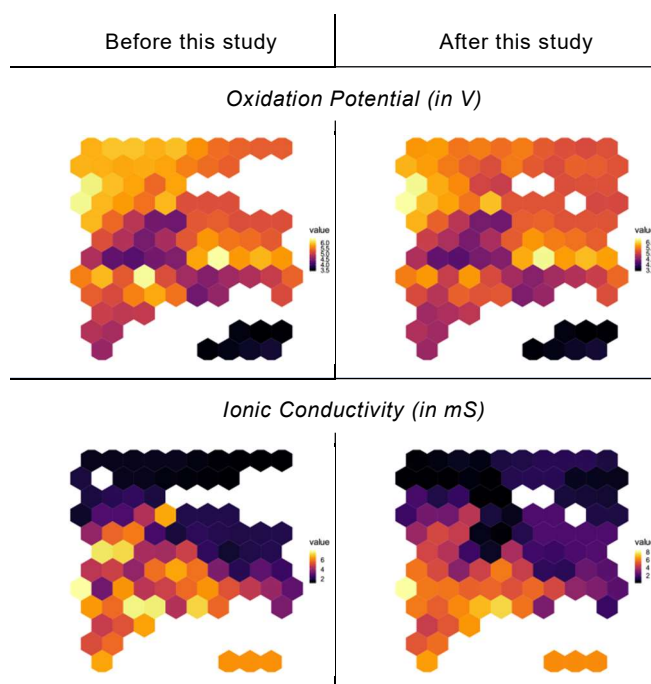


Figure 4. Evolution of the property landscape: oxidation potential (in V, top) and ionic conductivity (in mS, bottom). The color code indicates the considered property values. In white areas the property values are not available.

Optimization of the oxidation potential together with the ionic conductivity is delicate task because those two properties are generally anti-correlated. Yet, the ethers and esters provide with their best compromise which explains the preference for these compounds in current application. The sulfones display an interesting oxidation potential profile, the ionic conductivity is low, but higher than initially expected. However, the latter can be corrected at the formulation stage, with co-solvents.

To conclude, 35 theoretically designed ethers/esters and sulfones were synthesized and tested for a potential

Running title

application as solvent in Lithium-ion batteries. These compounds have been prioritized according to developed QSPR models and synthetic feasibility criteria. The most interesting compounds (**1** and **2**, Figure 2) were selected for the application in Lithium-ion batteries. Their electrolytic formulation is ongoing.

Computational Methods and Experimental Section

QSPR development and validation. Molecular structures were encoded by the ISIDA fragment descriptors^[1]. A pool of 175 types of descriptors differed by size and topology of fragments was systematically explored for the modeling using Support Vector Regression (SVR) with linear kernel. The modeling workflow included two nested 10-fold cross-validation procedures (Figure 1). For each set of descriptors the models were built and validated using two nested 10-fold cross-validation procedures. The inner one is used to tune the method parameters and the outer one is used to assess the performances of the resulting model. In the consensus models, each individual model is associated with its own applicability domain (AD). Here, the fragment control AD was used: compounds with possessing unseen chemical features in the training set were discarded. The optimized models, using the most relevant molecular descriptors were selected to enter a consensus model. The performances of the consensus and the individual models were assessed using the corresponding test sets. Finally, the entire dataset was used to build the final individual and consensus models, using the method parameters optimized at the cross-validation stage. The resulting consensus models have been implemented in the ISIDA/Predictor web service (<http://infochim.u-strasbg.fr/cgi-bin/predictor.cgi>). Notice that Ridge Regression and SVR (RBF kernel) models were also developed but not used because they performed less good compared to SVR (linear kernel) models.

Generative Topographic Mapping.

The maps have been build according to the protocol reported by Gaspar et al ^[1] using as descriptors atom-centered fragments accounting explicitly atoms and bonds up to 3 coordination spheres of the central atom. This type of descriptors was among those systematically retrieved during the QSPR modeling procedure. The number of RBF and their width (6 and 0.4, respectively) was tuned in order to optimize the likelihood of half the dataset, while fitting the GTM on the other half. The number of nodes and the regularization term were set to default: 150 and 1.0 respectively.

Synthesis. The sulfones syntheses were performed according to the methods reported in our previous work ^[17]. Appropriate amounts of lithium salts (LiPF₆ or LiTFSI) were added into the synthesized solvents in order to determine the physicochemical and electrochemical properties of the electrolytes.

Physicochemical and electrochemical properties. All molecules have been characterized using ¹H and ¹³C NMR as well as by GC-MS to check their chemical structure. They were furthermore characterized by DSC to obtain their melting and ebullition point. The lower range of DSC was -70°C. Ionic conductivity was acquired on 1M concentrated

solutions of compounds. Linear voltamperograms were acquired at 5 mV/s using platinum and vitreous carbon electrodes.

Acknowledgements

The authors thank the Agence Nationale de la Recherche (ANR DEVEGA, 2014-2019 for a financial support, B. F. and O. M.).

References

- [1] A. Varnek, D. Fourches, F. Hoonakker, V. P. Solov'ev, *J. Comput.-Aided Mol. Des.* **2005**, *19*, 693-703.
- [2] a) C. L. Champion, W. Li, B. L. Lucht, *Journal of The Electrochemical Society* **2005**, *152*, A2327-A2334; b) A. Hammami, N. Raymond, M. Armand, *Nature* **2003**, *424*, 635.
- [3] S. Hess, M. Wohlfahrt-Mehrens, M. Wachtler, *Journal of The Electrochemical Society* **2015**, *162*, A3084-A3097.
- [4] a) A. Chagnes, J. Swiatowska, *Lithium process chemistry: Resources, extraction, batteries, and recycling*, Elsevier, **2015**; b) K. Xu, *Chemical reviews* **2014**, *114*, 11503-11618; c) A. Chagnes, B. Carré, P. Willmann, D. Lemordant, *Journal of Power Sources* **2002**, *109*, 203-213; d) Y. Sasaki, *Electrochemistry* **2008**, *76*, 2-15; e) B. Flamme, PSL Research University **2017**.
- [5] J. C. Dearden, *Environmental Toxicology and Chemistry* **2003**, *22*, 1696-1709.
- [6] a) I. Oprisiu, G. Marcou, D. Horvath, D. B. Brunel, F. Rivollet, A. Varnek, *Thermochimica acta* **2013**, *553*, 60-67; b) K. Mansouri, C. M. Grulke, R. S. Judson, A. J. Williams, *Journal of cheminformatics* **2018**, *10*, 10; c) I. Sushko, S. Novotarskyi, R. Körner, A. K. Pandey, M. Rupp, W. Teetz, S. Brandmaier, A. Abdelaziz, V. V. Prokopenko, V. Y. Tanchuk, R. Todeschini, A. Varnek, G. Marcou, P. Ertl, V. Potemkin, M. Grishina, J. Gasteiger, C. Schwab, I. I. Baskin, V. A. Palyulin, E. V. Radchenko, W. J. Welsh, V. Kholodovych, D. Chekmarev, A. Cherkasov, J. Aires-de-Sousa, Q.-Y. Zhang, A. Bender, F. Nigsch, L. Patiny, A. Williams, V. Tkachenko, I. V. Tetko, *Journal of Computer-Aided Molecular Design* **2011**, *25*, 533-554.
- [7] I. V. Tetko, D. M. Lowe, A. J. Williams, *Journal of Cheminformatics* **2016**, *8*, 2.
- [8] O. Firuzi, A. Lacanna, R. Petrucci, G. Marrosu, L. Saso, *Biochimica et Biophysica Acta (BBA)-General Subjects* **2005**, *1721*, 174-184.
- [9] P. Winget, E. J. Weber, C. J. Cramer, D. G. Truhlar, *Physical Chemistry Chemical Physics* **2000**, *2*, 1231-1239.
- [10] R. C. Morrison, *The Journal of chemical physics* **1992**, *96*, 3718-3722.
- [11] a) J. Lind, X. Shen, T. Eriksen, G. Merenyi, *Journal of the American Chemical Society* **1990**, *112*, 479-482; b) M. Jonsson, J. Lind, T. E. Eriksen, G. Merényi, *Journal of the Chemical Society, Perkin Transactions 2* **1993**, 1567-1568; c) C. Li, M. Z. Hoffman, *The Journal of Physical Chemistry B* **1999**, *103*, 6653-6656.

Running title

- [12] M. Jonsson, J. Lind, T. E. Eriksen, G. Merenyi, *Journal of the American Chemical Society* **1994**, *116*, 1423-1427.
- [13] M. Jonsson, J. Lind, T. Reitberger, T. Eriksen, G. Merenyi, *The Journal of Physical Chemistry* **1993**, *97*, 11278-11282.
- [14] a) V. Polyakov, *Physics of the Solid State* **2001**, *43*, 655-662; b) S. Adams, J. Swenson, *Solid State Ionics* **2002**, *154*, 151-159; c) R. Xiao, H. Li, L. Chen, *Journal of Materiomics* **2015**, *1*, 325-332.
- [15] C. M. Bishop, M. Svensén, C. K. Williams, *Neural computation* **1998**, *10*, 215-234.
- [16] H. Gaspar, I. Baskin, G. Marcou, D. Horvath, A. Varnek, *Molecular informatics* **2015**, *34*, 348-356.
- [17] B. Flamme, M. Haddad, P. Phansavath, V. Vidal, A. Chagnes, *ChemElectroChem* **2018**, *5*, 2279-2287.

Received: ((will be filled in by the editorial staff))

Accepted: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

et al.
