Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information for

Charting the known chemical space for non-aqueous Lithium-air battery electrolyte solvents

Tamara Husch[1] and Martin Korth[1*]

[1] Institute for Theoretical Chemistry, Ulm University, Albert-Einstein-Allee 11, 89069 Ulm (Germany)

* Corresponding author e-mail: martin.korth@uni-ulm.de

Contents:

1. Comparison of donor numbers and chemical potential of Li^+

Johnson *et al.* as well as Aetukuri *et al.* proposed the donor number (DN) for the classification of Li-air battery electrolyte solvents in terms of their Li^+ solubility.[1, 2] COSMOtherm allows to compute the chemical potential of ions in bulk liquids, so that we we do not have to rely on acceptor and DNs, which apart from likely being less accurate are not readily available for large scale computational screenings. (For more details on the performance of COSMOtherm see section 2.) The chemical potential was calculated with COSMOtherm on top of BP86/TZVP-DFT calculations to compare the DN to the chemical potential of Li^+. The results can be seen in table 1.

Table 1: Comparison of experimentally derived donor numbers and the calculated chemical potential for Li^+ in the listed compound.

Molecule	DN	$\mu(Li^+) kcal/mol$
Acetonitrile	14.0	17.5
BN	11.9	19.0
DMC	17.2	17.9
DME	20.0	17.1
DMSO	30.0	9.3
EC	16.4	17.2
GBL	18.0	15.6
Me-Im	47.0	10.0
MeOAc	16.3	16.3
NMP	27.3	10.4
PC	15.1	17.2
SFL	14.8	15.4
TMP	23.0	15.3

Perusing this table we find a high correlation (R=0.85) between DN and $\mu(Li^+)$. COSMOtherm thus reproduces the trend that solvents with a higher DN are better at solving Li^+. We also find the nearly linear dependence in a DN-Range from 10-25 as reported by Aetukuri *et al.*.[2] The experimental finding is that for higher DNs a saturation of the Li^+ solubility is observed and increased DNs do not lead to better Li^+ solubility. This behavior is more correctly captured by $\mu(Li^+)$, which should therefore be more valuable for our purpose. As an alternative we also evaluated explicit Li^+ solvation energies at various theoretical levels by automatically constructing Li^+/solvent clusters for each candidate compound (including bulk solvation implicitly via COSMO), but correlation with DN and experiment was less good than for COSMOtherm chemical potentials (which unlike the cluster calculations include the correct statistical thermodynamics of the liquid phase).

The chemical potential in COSMO-RS theory (as computed by COSMOtherm) is therefore a very central quantity for our purpose, which is why further details are given for the information of the reader not familiar with the basics of COSMO-RS theory: COSMO-RS theory is based on the assumption of (independently) interacting molecular surface charges. It combines a QM-level electrostatic description of interacting molecules with the statistical thermodynamics of the liquid phase. The chemical potential for a compound in a system is derived from the integration of the σ-potential over the surface of the compound. The σ-potential measures the the affinity of the system to a surface of a certain polarity, and can be derived from a QM calculation with the COSMO implicit solvation model. COSMO-RS predictions thus make use of data from electronic structure theory calculations to allow for the description also of hitherto experimentally unknown species, which makes this approach superior to chemical engineering models for our purpose.

2. Further methodological considerations

Our multi-level screening strategy is based on the assumption that properties are treated acceptably well for pre-screening purposes already at lower theoretical levels. To evaluate this, we look at correlation values and deviation statistics between the predictions at different theoretical levels. Mean deviations (MDs) are a measure for how much values are shifted between methods on average, low mean average deviations (MADs) indicate similar absolute values, High Pearson R values indicate similar trends between methods. This analysis is sometimes swamped by a small number of compounds for which the methods to compare give completely different values, in some cases a reasonable and a totally unrealistic one. We have therefore calculated MD, MAD and R values after excluding such cases, also giving the percentage of compounds for which this was necessary, as well as the uncorrected values in parenthesis.

Perusing table 2, the following conclusions can be drawn: We use COSMOfrag for a very quick pre-screening with respect to the viscosity of candidate compounds. R, MD and MAD values in comparison to SQM predictions are 0.60, -0.5 cP and 1.24 cP. This seems acceptable given the fact that we use a much higher cutoff value of 10cP to select compounds at the COSMOfrag stage in comparison to 5cP at the SQM and 3cP at the DFT stages. We use SQM single points to pre-

Table 2: Correlations (Pearson R values) and deviation statistics (mean and mean absolute deviations, MD and MAD) for predictions at different theoretical levels.

	R	MD	MAD	exclusions
	COSMOfrag vs SQM			
η/cP	0.60(0.37)	-0.50(-0.81)	1.24(1.74)	20.6%
	SQM single points vs SQM optimizations			
μ(Li+)/(kcal/mol)	0.82	-0.17	0.95	0
	SQM vs DFT			
η/cP	0.71(0.51)	0.17(0.28)	0.73(0.85)	2.3%
T_f/K	0.83	-13.50	15.75	0
μ(Li+)/(kcal/mol)	0.60(0.39)	2.99(3.85)	3.09(3.95)	17.4%
μ(O_2-)/(kcal/mol)	0.84	-0.02	0.97	0

screen compounds with respect to the chemical potential for Li+. Compared to SQM optimizations a high R value (>0.8) and a low MAD is found (<10%), justifying this strategy. We use SQM-based property predictions to reduce the number of much more costly DFT calculations. The corresponding comparison of predicted values gives comparably high R values and acceptable MAD values, in the cases of η and μ(Li+) with some outliers. The positive MDs for η and μ(Li+) indicate an underestimation of these properties by SQM, in a rather pronounced way in the case of μ(Li+), but this is not harmful for our purpose of pre-screening compounds, as it only leads to false-positives, which are excluded at the higher level of theory.

COSMOtherm itself has known accuracy limitations for small ions, polar aprotic solvents and secondary and tertiary amines. Fortunately, trends for chemical potentials of a given small ion in a series of solvents are far less afflicted, and the performance for the combination of polar aprotic solvents and symmetric polar solutes like Li^+ and O_2^- is fine, as is the performance for secondary and tertiary amine solvents if only aprotic cases (without H-bonding effects) are investigated.[3]

3. Evaluation of COSMOtherm pK$_a$ predictions

To evaluate the performance of pK$_a$-calculations with COSMOtherm, we compared our results to data from Bryantsev.[4] The results can be seen in figure 1. Bryantsev evaluated pK$_a$ values in DMSO at MP2/CBS+δCCSD(T) level with an empirically adjusted SVPE2 solvent model and used the proton free energy of solvation as a parameter. He picked the most acidic protons based on chemical intuition, while we calculate pK$_a$ values for every hydrogen atom in a molecule, giving relevance only to the lowest value. After proton abstraction, anions are re-optimized at BP86/TZVP level. The pK_a is then calculated using COSMOtherm with DMSO as solvent. Not all structures are stable in the charged state and undergo severe structural changes. If the structure changes significantly, this is indicated by a "'-"' in the column "'stable charged structure"'. In our screening protocol, we take this as an indication of instability in an electrochemical setting.

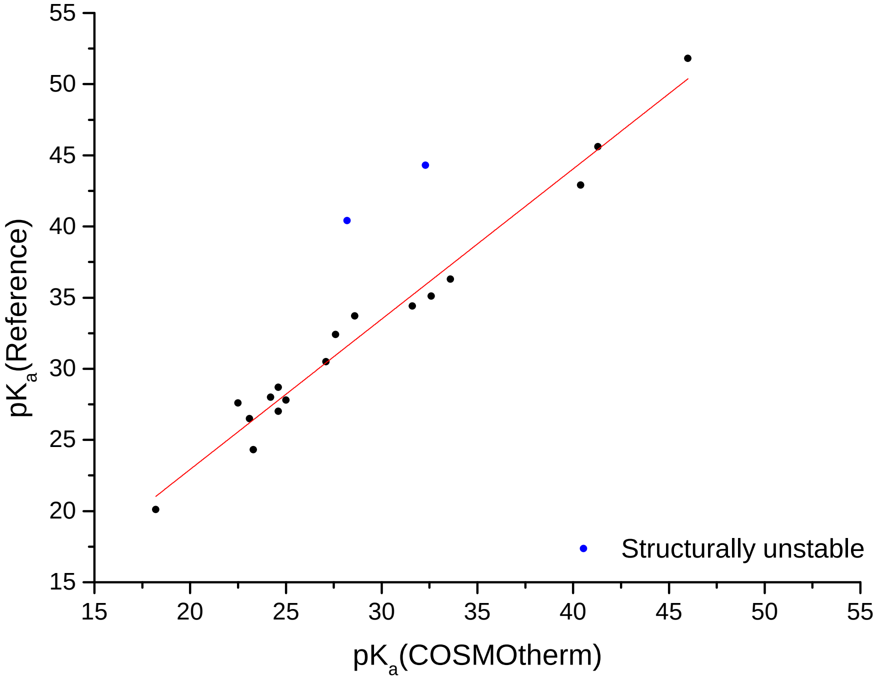


Figure 1: Comparison of calculated pK_a values with theoretical reference data from Ref.[4]

After excluding cases with significant structural changes (indicated in blue), we observe a very high correlation (R=0.99) for the comparison of COSMOtherm with higher-level computations. COSMOtherm pK_as are systematically shifted by 3.44 (MD=3.44, MAD=3.44). After adjusting this shift, the MAD is 1.05. Bryantsev empirically adjusted the solvation energy of the proton to improve agreement with experimental data. The COSMOtherm approach offers the advantage that no ad-hoc fitting is necessary. The observed high correlation with reference data furthermore shows the good predictive power of COSMOtherm, which is computationally inexpensive and suitable for high-throughput computations. The systematic shift can easily be taken into account by adjusting the cutoff-value for sorting out unstable structures. Bryantsev proposed a pK_a less than 35 as a criteria for unstable compounds, in our approach we are using 25. Similarly, the direct comparison with our reference compounds (DMSO, MeIm) is unproblematic.

Bryantsev *et al.* developed a protocol to predict the resistance of Li-air battery electrolyte solvents against nucleophilic attacks by the superoxide anion radical.[5] Later on Bryantsev as well as Khetan *et al.* proposed the pK_a as an estimator for this property, in addition to activation barriers for the nucleophilic attack $\Delta G_{act} \geq 25$.[4, 6, 7, 8] The computation of the explicit nucleophilic reaction barriers is not yet feasible for large-scale high-throughput calculation, but in table 3, we show that pK_a values are a reasonable first estimate for the stability against nucleophilic attacks, as measured by ΔG_{act}.

Perusing table 3, we observe that structural changes only occur for compounds which are predicted to be unstable against nucleophilic attacks. For all cases in which the activation energy is predicted to be higher than 25 kcal/mol, the pK_a is predicted as higher than 25. For cases that are predicted to be unstable against nucleophilic attacks, the pK_a is not a sufficient criteria. An example in this data set is DMC, which is predicted to be very susceptible to nucleophilic attacks, but has a pK_a of 42. From this we can conclude that compounds with a computed pK_a below 25 are likely to not stable in the setting of a Li-air battery. If the pK_a is high, this is an indication of stability, but has to be taken with care. We recommend a detailed investigation of final candidates that goes beyond the simple descriptor of pK_a, for example by following the suggestions of Bryantsev *et al.*.[5]

Table 3: Comparison of pK_a and ΔG values.

Compound	$\Delta G_{act}[kcal/mol]$*	$\Delta G_r[kcal/mol]$*	pK_a	stable charged structure
4,4-diF-GBL	27.41	-1.01	14,45	-
4,4-diMe-GBL	20.16	-1.15	18,02	-
Acetonitrile	24.92	18.96	27,45	
BN	25.31	21.57	37,71	
DMC	12.42	-11.39	42,40	
DME	31.56	19.88	45,46	
DMMP	12.37	-10.45	29,55	
EC	14.22	-6.94	10,26	-
EMS	22.79	2.79	18,66	-
GBL	16.52	-4.99	15,65	-
GVL	18.26	-3.84	17,10	-
MDMP	17.63	-6.55	31,93	
MeOAc	19.25	-1.30	27,27	
Me-Tos	10.17	-19.06	7,62	-
NMP	40.18	25.59	32,28	
PC	15.47	-5.64	10,77	-
PhB	14.45	0.91	38,78	
SFL	20.19	-2.98	15,43	-
TMP	12.55	-13.59	42,37	

(* Obtained from ref.[5])

4. Results for evaluation strategy 1

Table 4 presents the results for the first evaluation strategy, in which Pareto-optimal candidates from the whole final data set are picked. QSPR-estimated melting points are given as a first indication of whether compounds are liquid at room temperature, but the mean accuracy of these estimates is low and this data should thus be taken with care, especially if COSMOtherm predicts low viscosities.

Table 4: Strategy 1 - un-filtered Pareto-optimal candidates.

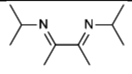
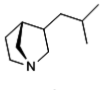
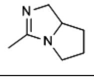
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ C$]	pK_a	Pic
00125001_00150000_13066	5.94	3.31	396.10	4.17	15.59	-2.17	37.3	26.49	
59700001_59725000_1543	5.65	2.07	335.96	4.65	17.87	-2.19	46.4	49.56	
20125001_20150000_6024	5.38	1.96	350.15	4.93	15.17	-2.13	68.5	34.45	

Table 4: Strategy 1 - un-filtered Pareto-optimal candidates.

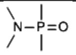
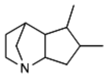
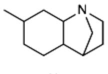
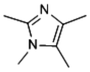
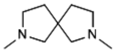
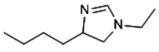
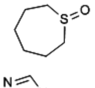
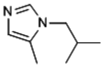
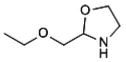
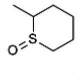
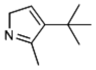
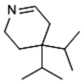
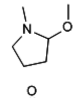
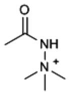
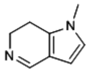
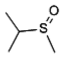
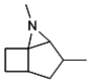
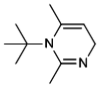
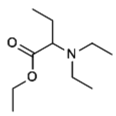
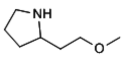
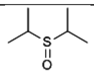
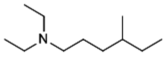
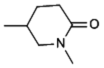
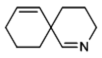
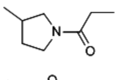
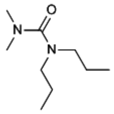
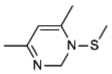
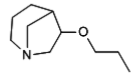
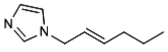
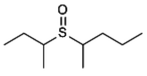
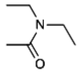
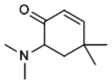
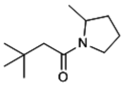
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^{-})$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ\mathrm{C}$]	pK$_a$	Pic
00500001_00525000_13753	6.92	1.79	349.89	5.44	12.63	-1.78	24.1	34.79	
54500001_54525000_13165	5.67	2.76	353.65	5.69	17.86	-2.20	105.7	49.34	
54500001_54525000_8047	5.67	3.08	358.98	6.19	17.99	-2.19	113.4	49.73	
00050001_00075000_16265	5.25	2.04	376.95	6.38	14.23	-2.03	69.3	37.25	
12675001_12700000_3024	5.08	2.93	352.71	6.56	16.54	-2.17	47.1	48.69	
57000001_57025000_14211	5.34	3.47	383.37	6.75	15.53	-2.18	23.6	40.21	
00525001_00550000_7313	6.18	2.88	406.61	6.86	13.50	-1.76	13.3	26.37	
58700001_58725000_15219	5.76	2.66	384.94	7.17	13.71	-2.05	38.9	38.90	
68175001_68200000_8242	6.63	2.86	370.29	7.24	13.33	-1.79	-19.9	31.41	
12200001_12225000_3956	6.18	2.41	399.77	7.29	13.89	-1.78	10.7	29.28	
58250001_58275000_10259	6.38	1.75	334.85	7.34	15.79	-2.16	4.2	16.20	
68250001_68275000_15315	6.02	2.19	349.74	7.41	17.03	-2.23	26.2	32.16	
69175001_69200000_7354	5.25	1.16	311.26	7.43	16.68	-1.97	-4.8	34.98	
10025001_10050000_6291	6.29	3.72	410.94	7.69	10.31	-1.62		40.69	
11750001_11775000_12592	5.35	3.75	402.42	7.69	12.59	-1.94	72.7	33.18	
00125001_00150000_10034	6.18	1.29	352.12	7.98	12.51	-1.68	-30.2	27.48	
59125001_59150000_664	5.47	1.69	325.78	8.00	17.89	-2.20	63.4	45.70	
67675001_67700000_5104	5.09	2.68	381.24	8.08	15.94	-2.18	52.2	24.48	
00125001_00150000_11962	5.23	2.71	372.12	8.20	16.39	-1.99	10.9	20.31	
59075001_59100000_9892	6.49	1.92	356.20	8.21	15.15	-2.01	-15.7	29.71	

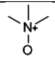
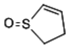
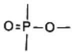
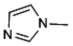
Table 4: Strategy 1 - un-filtered Pareto-optimal candidates.

ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^{-})$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
00000001_00025000_12645	6.09	1.38	367.16	8.24	14.80	-1.85	-0.5	26.96	
21950001_21975000_9158	4.91	1.74	332.63	8.42	19.20	-2.22	-40.2	49.03	
00150001_00175000_6040	6.34	1.89	373.60	8.47	15.21	-1.87	30.3	33.06	
00000001_00025000_516	6.29	1.24	343.99	8.50	9.12	-1.46	-45.2	32.88	
20400001_20425000_1834	5.84	2.96	363.49	8.52	16.41	-2.19	48.3	35.03	
54275001_54300000_4110	6.23	2.32	387.22	8.61	15.76	-1.94	27.1	33.87	
20050001_20075000_12073	5.81	1.92	375.41	8.67	16.64	-2.01	22.4	44.73	
18750001_18775000_3561	5.02	3.41	397.13	8.69	14.16	-2.06	92.0	2.96	
19825001_19850000_12260	5.32	3.62	374.74	8.70	17.10	-2.10	57.0	47.83	
54300001_54325000_8631	6.05	5.02	401.80	8.81	12.84	-2.04	30.6	32.28	
68150001_68175000_1461	6.04	2.77	400.51	8.83	15.86	-2.00		26.15	
00000001_00025000_9770	6.32	1.05	338.47	8.85	14.93	-1.83	-44.8	32.68	
18300001_18325000_5273	5.18	3.88	416.13	8.93	14.93	-1.95	57.3	26.49	
58350001_58375000_10825	6.27	3.44	407.00	8.97	16.56	-2.04	42.2	33.67	

5. Results for evaluation strategy 2

For table 5, Pareto-optimal candidates are picked after filtering out structures with one or more property much worse than DMSO and MeIm. Filter thresholds were: $IP \leq 6.0\ eV,\ Visc \leq 3.0\ cP,\ FP \geq 323.0\ K,\ \mu(Li^{+}) \leq 9.0\ kcal/mol,\ \mu(O_2^{-}) \leq 10.0\ kcal/mol,\ \mu(O_2) \leq -1.0\ kcal/mol)$. QSPR-estimated melting points are given as a first indication of whether compounds are liquid at room temperature, but the mean accuracy of these estimates is low and this data should thus be taken with care, especially if COSMOtherm predicts low viscosities.

Table 5: Strategy 2 - filtered Pareto-optimal candidates.

ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
00000001_00025000_868	6.09	1.70	364.75	1.33	9.38	-1.82		43.01	
12600001_12625000_1867	6.19	2.74	399.27	8.41	9.65	-1.46	-16.6	25.54	
11700001_11725000_294	7.31	2.01	392.91	8.43	9.75	-1.48	-21.7	31.99	
00000001_00025000_516	6.29	1.24	343.99	8.50	9.12	-1.46	-45.2	32.88	
00000001_00025000_1057	6.05	1.51	345.23	8.84	9.81	-1.66	3.4	38.24	

6. Results for evaluation strategy 3

For table 6, Pareto-optimal candidates are picked after filtering out structures which show significantly worse chemical potentials for Li^+ and O_2^-. Filter thresholds were the values of MeIm +10%, i.e. $\mu(Li^+) \leq 9.72kcal/mol,\ \mu(O_2^-) \leq 10.79kcal/mol$. QSPR-estimated melting points are given as a first indication of whether compounds are liquid at room temperature, but the mean accuracy of these estimates is low and this data should thus be taken with care, especially if COSMOtherm predicts low viscosities.

Table 6: μ-filtered Pareto-optimal candidates.

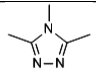
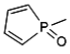
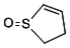
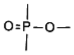
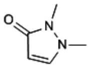
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
00550001_00575000_18589	6.41	2.81	449.34	7.34	10.59	-1.90	89.5	34.26	
20500001_20525000_7677	6.46	3.60	398.71	8.07	10.26	-1.49	18.5	29.66	
12600001_12625000_1867	6.19	2.74	399.27	8.41	9.65	-1.46	-16.6	25.54	
11700001_11725000_294	7.31	2.01	392.91	8.43	9.75	-1.48	-21.7	31.99	
12200001_12225000_1705	5.38	4.56	449.08	8.46	9.27	-1.58	55.2	34.02	
00000001_00025000_516	6.29	1.24	343.99	8.50	9.12	-1.46	-45.2	32.88	

Table 6: μ-filtered Pareto-optimal candidates.

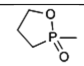
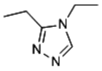
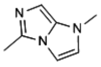
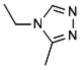
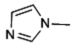
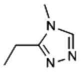
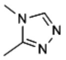
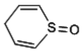
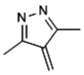
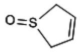
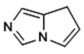
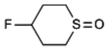
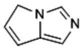
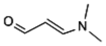
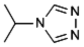
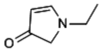
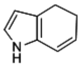
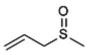
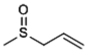
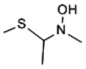
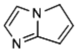
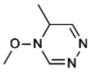
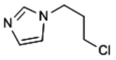
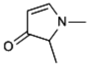
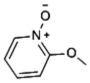
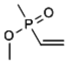
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
11875001_11900000_14792	7.34	2.94	412.32	8.67	10.69	-1.47	18.8	14.05	
20550001_20575000_677	6.71	5.12	450.56	8.75	10.21	-1.92	75.0	34.82	
68325001_68350000_3338	4.72	5.58	429.73	8.76	10.60	-1.86	83.9	34.95	
12275001_12300000_592	7.45	4.23	446.60	8.82	9.27	-1.81	69.0	33.86	
00000001_00025000_1057	6.05	1.51	345.23	8.84	9.81	-1.66	3.4	38.24	
18725001_18750000_1413	6.71	4.83	448.98	9.07	8.89	-1.79	68.0	34.95	
12200001_12225000_2961	6.97	3.81	445.80	9.10	7.92	-1.70	70.6	33.76	
23550001_23575000_15879	6.09	4.65	438.73	9.10	9.09	-1.40	3.3	11.05	
10325001_10350000_1650	6.82	2.89	425.12	9.21	10.58	-1.88	79.8	22.25	
12725001_12750000_2416	6.36	2.48	386.95	9.23	9.95	-1.43	-21.9	26.66	
20700001_20725000_15339	5.84	3.14	376.59	9.27	9.57	-1.70	70.7	20.33	
69625001_69650000_7810	6.33	3.90	418.83	9.28	10.45	-1.56	21.8	24.76	
12275001_12300000_10	5.47	3.13	386.31	9.33	9.65	-1.64	70.2	21.66	
00625001_00650000_8861	5.61	1.82	422.65	9.42	10.69	-1.60	-33.3	36.15	
05250001_05275000_3828	7.08	5.70	440.50	9.51	9.25	-1.89	80.1	22.59	
19850001_19875000_11134	5.53	3.72	427.28	9.56	10.45	-1.58	19.4	24.14	
20700001_20725000_7948	6.80	3.51	410.19	9.57	10.74	-1.74	28.3	-0.22	
10875001_10900000_9510	6.22	2.12	371.81	9.58	10.22	-1.51	-44.7	24.99	
11750001_11775000_12499	6.22	2.12	371.81	9.58	10.22	-1.51	-44.7	24.99	
13700001_13725000_2503	5.39	2.66	423.05	9.62	10.19	-1.60	16.5	28.84	
10975001_11000000_17384	5.49	3.06	393.28	9.64	9.75	-1.66	71.0	20.09	
53675001_53700000_11625	5.83	4.71	426.22	9.66	9.61	-1.71	80.0	-5.43	
15750001_15775000_544	6.11	4.08	398.79	9.67	10.76	-1.79	35.0	7.77	

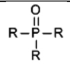
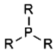
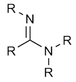
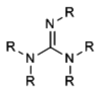
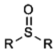
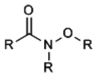
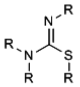
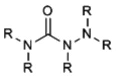
Table 6: μ-filtered Pareto-optimal candidates.

ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
68100001_68125000_17906	5.52	3.24	424.10	9.68	10.67	-1.59	57.1	24.75	
00125001_00150000_10898	6.80	5.76	443.34	9.71	9.43	-1.42	14.0	35.31	
06450001_06475000_2171	7.12	2.59	400.80	9.72	10.22	-1.43	-29.8	30.60	

7. Analysis of compound class performance

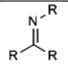
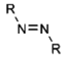
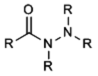
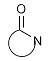
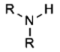
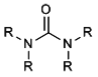
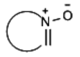
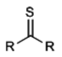
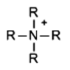
To analyze the average performance of the different compound classes in the PubChem database, functional groups of the whole final set of compounds were determined, and property values were averaged over all members of the same compounds class, defined by a functional group each. Table 7 presents the average value of all compounds that contain the specific functional groups. Compound classes are sorted by the chemical potential of Li^+ and listed if the average $\mu(Li^+)$ was lower than the MeIm value plus 1.

Table 7: Compound class analysis.

Compound Class*	Pic	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	pK$_a$
Phosphine Oxide		6.65	2.69	387.18	5.85	13.43	-1.78	33.17
Phosphine		6.56	2.78	386.57	6.34	13.55	-1.80	33.17
Carboxylic Acid Amidine		5.34	2.52	370.88	8.13	15.44	-2.13	33.42
Secondary Aliphaticaromatic Amine Alkylarylamine	N-Rings	-5.43	3.96	392.81	8.91	12.65	-1.97	25.61
Guanidine		5.16	1.99	373.97	9.16	16.32	-2.17	42.27
Sulfoxide		6.14	2.68	387.80	9.41	12.82	-1.74	26.25
Hydroxamic Acid		5.69	2.89	390.50	9.44	8.41	-1.28	-12.05
Isothiourea		5.26	2.97	405.83	9.53	13.98	-2.02	36.61
Semicarbazide		4.97	2.25	375.24	9.89	15.29	-1.96	27.31

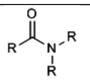
* According to Checkmol by Haider [9]

Table 7: Compound class analysis.

Compound Class*	Pic	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	pK$_a$
Imine		5.79	2.84	366.07	10.27	15.39	-2.12	26.05
Cation		5.32	4.25	397.34	10.27	12.44	-1.97	31.05
Azo Compound		5.18	2.70	383.29	10.32	16.56	-2.28	0.00
Carboxylic Acid Hydrazide		5.30	2.69	387.40	10.33	14.75	-1.90	24.29
Lactam		6.20	3.12	395.42	10.42	14.80	-1.88	33.46
Secondary Amine		5.90	2.97	365.11	10.46	15.29	-2.10	30.69
Urea		5.71	2.75	390.00	10.46	15.42	-1.98	44.05
Oxohetarene		5.94	3.98	409.76	10.48	12.96	-1.80	31.94
Secondary Aliphatic Amine Dialkylamine	N-Rings	5.94	2.88	362.58	10.61	15.53	-2.11	30.84
Tertiary Aliphaticaromatic Amine Alkylarylamine	N-Rings	5.32	3.30	399.17	10.65	13.76	-2.06	33.98
Azide	$R-N=\overset{+}{N}=\overset{-}{N}$	5.28	2.57	381.02	10.65	14.85	-1.98	24.79
Thioketone		6.68	3.48	410.36	10.66	13.50	-1.76	0.00
Quaternary Ammonium Salt		4.76	3.39	371.74	10.83	14.94	-2.09	18.18

* According to Checkmol by Haider [9]

Table 7: Compound class analysis.

Compound Class*	Pic	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	pK$_a$
Tertiary Carboxylic Acid Amide		6.21	3.02	391.93	10.84	14.98	-1.92	33.98

* According to Checkmol by Haider [9]

8. Phosphine oxides screening results

Phosphine oxides were identified as a promising compound class, but no structure in the PubChem database was found to be better than the reference molecules (DMSO,MeIm). We therefore investigated 362 systematically generated phosphine oxides and their sulfur analogs, as well as 962 cyclic phosphine oxides. See main text for more details. Our best suggestions are listed in table 8. QSPR-estimated melting points are given as a first indication of whether compounds are liquid at room temperature, but the mean accuracy of these estimates is low and this data should thus be taken with care, especially if COSMOtherm predicts low viscosities.

Table 8: μ-filtered Pareto-optimal candidates.

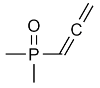
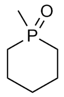
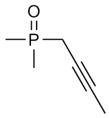
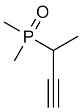
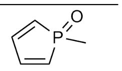
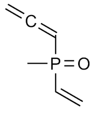
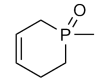
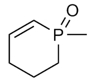
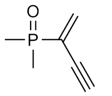
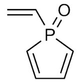
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
po911	6.62	3.27	386.42	6.90	10.60	-1.61	-4.0	24.17	
po138	5.92	4.70	423.24	7.23	10.72	-1.59	22.5	20.47	
po164	6.37	4.77	407.00	7.61	10.41	-1.61	4.4	24.90	
po110	6.75	5.78	372.49	7.83	10.07	-1.65	8.5	25.47	

Table 8: μ-filtered Pareto-optimal candidates.

ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
po922	6.46	3.63	399.10	8.05	10.27	-1.49	18.4	29.62	
po129	6.73	4.17	393.85	8.65	10.51	-1.62	6.0	23.16	
po175	6.31	5.04	398.43	8.88	10.07	-1.50	24.8	23.16	
po177	7.06	5.07	457.40	8.97	10.13	-1.52	22.8	22.45	
po196	6.85	5.99	375.76	9.03	9.30	-1.61	4.3	23.30	
po277	6.50	5.34	417.80	9.07	9.81	-1.46	33.1	24.35	

9. Heterocycles screening results

Heterocycles were identified as a promising compound class, and some structures in the PubChem database were found to be competitive to the reference molecules (DMSO, MeIm), though a trade-off between the relevant ion solubilities and viscosity could be identified. For futher investigations, all small- to medium-sized aromatic N-hetero-bi/mono-cycles, all small N-hetero-monocycles, and all N/O/S-hetero-cycles with simple aliphatic substituents were investigated (overall about 350.000 structures). See main text for more details. From all competitive candidates we list here the 31 Pareto-optimal ones which contain only one hetero-atom species in table 9 QSPR-estimated melting points are given as a first indication of whether compounds are liquid at room temperature, but the mean accuracy of these estimates is low and this data should thus be taken with care, especially if COSMOtherm predicts low viscosities.

Table 9: μ-filtered Pareto-optimal candidates.

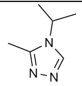
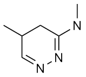
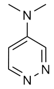
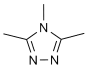
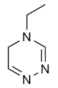
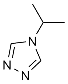
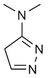
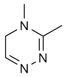
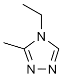
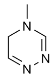
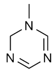
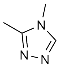
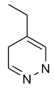
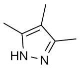
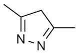
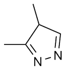
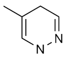
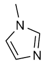
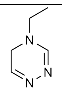
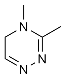
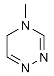
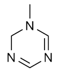
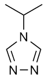
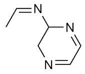
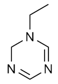
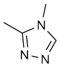
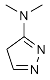
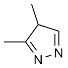
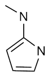
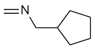
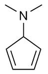
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^{-})$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
best 'simple' N-hetero-cycles (see main text for details)									
c1e29189	6.71	4.27	448.63	8.26	8.26	10.70	79.3	24.72	
c1e30459	5.91	4.94	469.24	8.53	8.53	9.76	63.9	5.21	
c1e31492	5.74	3.43	458.09	7.79	7.79	10.81	77.0	38.52	
c1e33518	6.41	2.81	448.89	7.38	7.38	10.56	89.5	34.31	
c1e33581	5.63	3.71	440.72	7.32	7.32	9.95	46.9	32.12	
c1e33608	7.08	5.16	439.50	9.69	9.69	8.46	73.5	23.34	
c1e33639	5.45	3.54	444.78	7.86	7.86	9.99	60.9	13.01	
c1e33648	5.51	2.81	437.15	6.65	6.65	10.73	59.0	28.22	
c1e33651	7.45	4.25	445.67	8.76	8.76	9.25	68.5	33.88	
c1e34374	5.63	3.06	436.27	7.66	7.66	8.72	56.6	31.39	
c1e34376	5.78	2.61	376.91	9.13	9.13	9.64	59.2	35.93	
c1e34388	6.97	3.86	445.41	9.09	9.09	7.89	70.6	33.78	
c1e78346	6.08	2.68	414.32	8.79	8.79	10.81	41.0	11.38	
c1e78454	6.81	2.94	425.94	9.25	9.25	10.56	79.6	22.26	
c1e79038	6.69	2.45	404.66	8.83	8.83	10.29	80.6	10.35	
c1e79088	6.97	2.09	390.50	9.62	9.62	10.32	64.4	8.06	
c1e79160	6.28	2.35	392.09	9.71	9.71	9.37	52.2	8.46	
c1e79277	6.05	1.51	343.88	8.84	8.84	9.74	3.4	38.28	

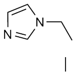
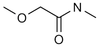
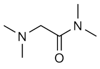
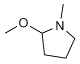
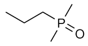
Table 9: μ-filtered Pareto-optimal candidates.

ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	$\mu(O_2)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
best 'simple' N/O/S-hetero-cycles (see main text for details)									
c1f88642078	5.63	3.66	440.70	7.35	7.35	9.95	46.9	30.72	
c1f88642106	5.51	2.80	437.16	6.64	6.64	10.73	59.0	28.23	
c1f88642376	5.63	3.01	436.30	7.68	7.68	8.79	56.6	31.39	
c1f88652377	5.78	2.61	376.91	9.11	9.11	9.65	59.2	35.93	
c1f7248105	7.08	5.15	439.53	9.71	9.71	8.46	73.05	23.10	
c1f8850623	6.13	4.12	386.63	8.49	8.49	10.56	59.8	12.33	
c1f88512079	5.78	3.11	382.02	8.88	8.88	10.68	46.4	35.54	
c1f7249574	7.43	3.76	445.09	9.09	9.09	7.92	70.6	33.81	
c1f7250123	5.46	3.50	444.60	7.92	7.92	9.98	60.9	13.04	
c1f7267966	6.97	2.08	390.36	9.64	9.64	10.32	64.4	8.09	
c1f7267979	6.66	2.06	368.84	9.58	9.58	9.98	21.1	8.74	
c1f7287380	4.72	5.62	402.59	-7.02	-7.02	10.74	-39.5	2.22	
c1f7287759	5.15	5.10	518.14	8.61	8.61	8.88	-23.0	18.18	

10. Final results

Finding Li-air battery electrolyte solvents better than DMSO and MeIm turns out to be surprisingly hard, if indeed both Li^+ and O_2^- solubilities play an important role, as suggested in recent experimental reports. The following compounds are our suggestions for experimental investigations into the optimal balance of Li^+/O_2^- solubilities with viscosity, see main text for details.

Table 10: Final results.

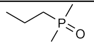
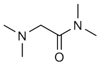
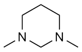
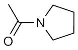
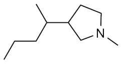
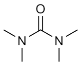
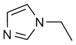
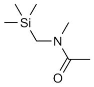
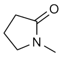
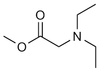
ID	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^+)$ [kcal/mol]	$\mu(O_2^-)$ [kcal/mol]	MP[$^\circ$C]	pK$_a$	Pic
18725001_18750000_2537	6.11	1.85	390.15	9.72	13.32	-16.5	32.74	$\mathrm{H_3C{-}C(=O){-}N({-}CH_3){-}CH_2{-}CH_2{-}O{-}CH_3}$
00075001_00100000_5084[1]	6.04	1.80	353.44	8.52	10.93	3.2	38.06	
11125001_11150000_287	6.42	1.74	393.28	7.30	12.49	-17.3	30.38	
00075001_00100000_6354	5.19	1.76	377.63	6.25	13.98	-11.6	31.33	
18700001_18725000_7879	5.24	1.28	320.59	5.97	16.01	-4.8	33.68	
20475001_20500000_3853	6.88	1.93	362.27	4.17	13.16	4.9	35.54	

[1] DMSO has the lowest $\mu(O_2^-)$ in the 8-9 kcal/mol interval for $\mu(Li^+)$; 00075001_00100000_5084 is listed as alternative.

11. Alternative results

Our approach and the computed data is of value for the development of Li-air battery electrolytes, even if the assumed hypothesis about the importance of the different ion solubilities turns out to be wrong, because our database can easily be re-evaluated with respect to other criteria. As an example we give here a list of the most promising motifs for the case that the solubility of the negative species is acutally of lesser importance, i.e. looking for compounds with a viscosity lower than 2 cP in combination with a Lithium cation chemical potential below 10 kcal/mol, as well as pK$_a$s higher than 25 and melting points lower than 10K. Very different compounds are found in this analysis in comparison to the previous ones, now with an emphasis on amide and amine motifs, thus nicely illustrating how important the choice of selection criteria (and therefore input from experiment) is.

Table 11: Alternative Results: Representative with the lowest $\mu(Li^{+})$ for each motif.

Motif	IP[eV]	Visc[cP]	FP[K]	$\mu(Li^{+})$ [kcal/mol]	$\mu(O_2^{-})$ [kcal/mol]	MP[$^\circ$C]	pK_a	Pics
Phosphine Oxide	6.88	1.93	362.27	4.17	13.16	4.90	35.54	
Tertiary Amine, Amide	5.19	1.76	377.63	6.25	13.98	-11.60	31.33	
Tertiary Diamine	4.97	1.24	305.59	7.10	16.60	9.00	47.98	
Amide	6.26	1.52	364.29	7.89	14.50	4.60	32.69	
Tertiary Amine	4.85	1.88	332.90	7.94	18.37	5.10	49.95	
Urea	5.97	0.90	327.42	8.51	15.33	-7.50	44.06	
Aromatic	6.04	1.80	353.44	8.52	10.93	3.20	38.06	
Silane, Amide	5.70	1.98	335.57	9.02	15.21	3.80	32.45	
Lactam	6.35	1.14	342.03	9.08	14.07	-16.50	32.52	
Tertiary Amine, Ester	5.24	1.64	331.27	9.24	15.11	-1.40	28.09	

References

[1] L. Johnson, C. Li, Z. Liu, Y. Chen, S. A. Freunberger, P. C. Ashok, B. B. Praveen, K. Dholakia, J.-M. Tarascon, P. G. Bruce, *Nature chemistry* **2014,** *6,* 1091–1099.

[2] N. B. Aetukuri, B. D. McCloskey, J. M. García, L. E. Krupp, V. Viswanathan, A. C. Luntz, *Nature chemistry* **2014,** *7,* 50-56.

[3] F. Eckert, A. Klamt, *Version C3.0* **2014,** *1,*.

[4] V. S. Bryantsev, J. Uddin, V. Giordani, W. Walker, D. Addison, G. V. Chase, *Journal of The Electrochemical Society* **2013,** *160,* A160-A171.

[5] V. S. Bryantsev, M. Blanco, *The Journal of Physical Chemistry Letters* **2011,** *2,* 379-383.

[6] V. S. Bryantsev, *Chemical Physics Letters* **2013,** *558,* 42 - 47.

[7] A. Khetan, H. Pitsch, V. Viswanathan, *The Journal of Physical Chemistry Letters* **2014,** *5,* 1318-1323.

[8] A. Khetan, H. Pitsch, V. Viswanathan, *The Journal of Physical Chemistry Letters* **2014,** *5,* 2419-2424.

[9] N. Haider, *Molecules* **2010,** *15,* 5079–5092.