nature portfolio

Peer Review File

Active learning accelerates electrolyte solvent screening for anode-free lithium metal batteries

Corresponding Author: Professor Chibueze Amanchukwu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This study explores the selection of new solvents for metal lithium batteries using machine learning (ML) methods. While the research topic is compelling, there are several areas where the manuscript requires significant improvement. In particular, the description of how ML is applied to the experimental data is unclear, and the selection and quality of the training data are questionable. The paper appears to reflect a collaboration between experimental researchers and AI specialists; however, the integration of these two aspects is insufficient, resulting in a fragmented narrative. Detailed points are provided below:

1. Quality of Experimental Data
"As a starting point, a high-quality in-house dataset of 58 anode-free LMB cycling profiles" (page 3)
The definition of "high-quality data" is vague. It is critical to quantitatively assess the variability in battery performance when creating cells under identical conditions. Without this, the reliability and reproducibility of the dataset remain uncertain.

2. Standard Testing Configuration
"To reduce complexity and focus primarily on improving lithium metal cycling stability, Cu/LFP was selected as the standard testing configuration" (page 5)
Despite this claim, Table 2 indicates that NMC cathodes were used in some cells. The manuscript lacks a clear explanation of why multiple cathode types were employed and omits details regarding the experimental conditions for these cases.

3. Issues with Training Data
The selection of training data is crucial for constructing accurate ML prediction models. However, the dataset described in Table 2 is inadequate in several respects:
Many solvents, such as "Nonfluorinated ether A," are not sufficiently described.
Lithium salt concentrations vary between solvents, and the criteria for determining these concentrations are not explained.
Since both solvent type and salt concentration can independently affect battery performance, the dataset as presented is unsuitable for training a robust ML model.

4. Application of Machine Learning
The manuscript does not clearly describe how ML methods were applied to the dataset in Table 2. Specifically, the handling of varying lithium salt concentrations within the training process remains ambiguous. This raises concerns about the reliability of the ML analysis.

5. Interpretation of ML Results
The manuscript does not adequately explain how the solvents listed in Table S3 were identified as promising candidates based on the ML analysis. Detailed discussions on the model's interpretability and decision-making process are essential to validate these predictions.

Reviewer #2

(Remarks to the Author)
In this article, the authors demonstrate the use of active learning (AL) as an alternative approach to accelerate electrolyte

discovery for anode-free LMBs. The study employs sequential Bayesian experimental design with Bayesian model averaging (BMA) to efficiently identify optimal candidates in typical data-scarce and noisy label settings. However, the computational study could still be improved to better align with the experimental data. There are also some experimental data should be revised carefully.
1. Figure 1b fails to indicate the specific electrolyte used. Please label it or provide a brief description in the manuscript.
2. How to understand the virtual chemical search space (~380000 unique molecules) across different batches? What do the data points and regions mean in the Figure 2a? Can the authors compare the current AL suggestions results with other machine learning algorithms, such as Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Light Gradient Boosting Machine (LightGBM)?
3. Where are the codes available on GitHub (https://github.com/AmanchukwuLab/AL-anode-free)? It is surprising that the authors did not mention any active learning processes or algorithms in the Methods section, making it difficult for other researchers to reproduce the results.
4. The SEM images of lithium deposition for the four electrolytes illustrated in Figure 3c-f are not sufficient to demonstrate that all electrolytes form a tight and uniform stack of particles. To maintain experimental rigor, it is recommended to supplement the cross-sectional SEM images of these four electrolytes.
5. Figure 4c compares ionic conductivity of 1 M LiFSA in AL discovered solvents as a function of temperature. However, no description is found in the article. Also, Figure 4d compares diffusivity of lithium ion (DLi) and FSA anion (DFSA) and lithium transference number (tLi) of AL discovered electrolytes at 1 M LiFSA concentration, and again no description is found in the article. The authors should provide a brief description and analysis in the manuscript.
6. In the Supporting Information and Notes, the authors presented the Different kernels used in GPR and performance comparison, Bayesian model averaging (BMA), Oracle or acquisition functon(EI) and related terms, and so on. I would suggest the authors share the relevant database, codes and atomistic models online.

Reviewer #3

(Remarks to the Author)
The paper demonstrates an active learning approach to accelerate the discovery of optimal electrolyte solvents for anode-free lithium metal batteries (LMBs). Using Bayesian Model Averaging (BMA), the study efficiently explores a chemical space of over 1 million candidates, identifying four high-performing solvents through iterative Cu/LFP coin cell tests. The main work focuses on screening electrolyte solvents paired with a 1M LiFSI salt to optimize capacity retention and improve battery cycling performance. However, major revisions are required to improve the clarity of the AL workflow, particularly regarding the implementation details shown in Figure 1c. Providing the code on GitHub, if possible, would significantly enhance transparency and reproducibility.
The study focuses primarily on 1M LiFSI in Cu/LFP coin cells, and I did not notice a systematic exploration of other lithium salts, salt concentrations, or cathode materials in the main text or supplementary information. While this fixed condition is reasonable given the scope of the AL model, more detailed information on how the workflow operates is necessary. For example, when screening solvents from eMolecules and PubChem, the first priority should be liquid organic solvents besides the criteria outlined in Table S6. (in other words, it is not necessary to screen a huge number of molecules if we know they are solid at RT).
Additionally, the dataset used in this study primarily consists of ether derivatives. When the model suggests candidates with entirely different functional groups that are not commercially available (as explained in Figure S4a), it would be more beneficial to select a structurally similar, commercially available molecule to collect experimental data for model training. This would ensure more robust feedback to the AL framework.
Finally, while the focus on single-solvent electrolytes is a valuable starting point, exploring multi-solvent formulations to leverage synergistic effects could further enhance battery performance. If the AL workflow can be adapted to optimize solvent mixtures, it would significantly increase the impact and practical relevance of this study.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Part of the concerns have been solved by suitable revision of the manuscript. however, there is still technicall problesm in the manuscript. Alhtough the concept and strategy of the present work looks interesting, there is serious concerns for the way of accumulating the inital datasets for constructing AI model.

1: type of cathode material
“The decision to include NMC cathode data in the initial dataset was made to enlarge the number of initial training dataset, ensuring a more robust starting point for the active learning framework. “
The degradation mechanism of NMC/Li cells are quite different with that of LFP/Li cells. In general, even using the same electrolyte material and setting the same cycling condition, the cycle life is quite different by the choice of cathode materials and charging cutoff voltage condition. In particular, in NMC811/Li cell, in which cutoff voltage set to be 4.4V, the oxidative decomposition of electrolyte occurs, and this induces the complicated chemical crossover phenomenon. Thus, the investigation should be completed by focusing on sole cathode material. The reviewer recommends that the AI model

should be reconstructed focusing on only data sets of LFP/Li cells.

2: training dataset
“This is because in-house data were collected from different projects prior to this work and therefore naturally have slightly different testing conditions.”
This explanation is not a reasonable excuse. The arrangement of datasets is the most important point for constructing AI models. The author should handle well-arranged and organized data sets for constructing AI models. The reviewer considers it is inappropriate to use the data that is obtained by other projects with different purposes.

3: Details of solvent structure used in the electrolyte
“Regarding the use of nomenclature such as “nonfluorinated ether A,” their structures were withheld because they are involved in other unpublished work from our group. We have now updated Table S2 to disclose the names of all such solvents.”
Without showing the actual structure of the molecule, the readers cannot follow the details of this study. The reviewer recommends that this manuscript should be resubmitted after their unpublished work is published.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
The authors addressed my comments. I would recommend the acceptance.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
I appreciate the authors' thorough and thoughtful responses to my previous comments and concerns. The revised manuscript addresses all major points I raised. The only remaining revision I suggest is a final proofread to correct minor typographical and grammatical errors. With these updates, I find the manuscript suitable for publication and recommend acceptance after minor editorial revision.

(Remarks on code availability)
This is a valuable addition that significantly enhances the transparency and reproducibility of the work.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for your response to my previous comments. The manuscript was properly revised, and my previous concerns related to AI parts are mostly solved. Here, let me comment on the experimental analytical part. In particular, latter part of the manuscript, during figure 3 to 5. The details are described as follows.

1:
In the experiments of figure 3, the electrolyte of 1 M LiFSA in F5DEE was utilized as control. However, this electrolyte composition is not so common in the LMB community. Instead, 4M LiFSI in DME (1,2-Dimethoxyethane) is widely utilized as standard electrolyte for lithium metal electrode investigation. In addition, this study also uses this electrolyte for some cells for in-house anode-free cells. Thus, the reviewer suggests that the experiments for cycling test with the cell use of 4M LiFSI in DME also should be conducted. This result is very helpful for the readers in this community as a control experiment.

2:
The manuscript showed the analytical study for the four kinds of discovered electrolyte, by series of standard experiments for Li metal study, including, SEM, Raman, Li potential shift, measurement of conductivity and diffusivity, XPS. These studies provide detailed information that is necessary for answering the question, what is the origin of high-capacity retention of the discovered electrolyte. To deepen the discussion, the reviewer suggests that the data obtained from the cell with control electrolyte (1 M LiFSA in F5DEE and 4M LiFSI in DME) also shown and the detailed discussion for comparing the results is encouraged. Such data and discussion are very meaningful for readers and the community of LMB to get deep insight into the discovered electrolyte through the present study.

(Remarks on code availability)

Version 3:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The manuscript has been properly revised and all the issues have been resolved. Therefore, the reviewer suggests accepting the manuscript in its current form.

(Remarks on code availability)

Comments for Author

The paper demonstrates an active learning approach to accelerate the discovery of optimal electrolyte solvents for anode-free lithium metal batteries (LMBs). Using Bayesian Model Averaging (BMA), the study efficiently explores a chemical space of over 1 million candidates, identifying four high-performing solvents through iterative Cu/LFP coin cell tests. The main work focuses on screening electrolyte solvents paired with a 1M LiFSI salt to optimize capacity retention and improve battery cycling performance. **However, major revisions are required to** improve the clarity of the AL workflow, particularly regarding the implementation details shown in Figure 1c. Providing the code on GitHub, if possible, would significantly enhance transparency and reproducibility.

The study focuses primarily on 1M LiFSI in Cu/LFP coin cells, and I did not notice a systematic exploration of other lithium salts, salt concentrations, or cathode materials in the main text or supplementary information. While this fixed condition is reasonable given the scope of the AL model, more detailed information on how the workflow operates is necessary. For example, when screening solvents from eMolecules and PubChem, the first priority should be liquid organic solvents besides the criteria outlined in Table S6. (in other words, it is not necessary to screen a huge number of molecules if we know they are solid at RT).

Additionally, the dataset used in this study primarily consists of ether derivatives. When the model suggests candidates with entirely different functional groups that are not commercially available (as explained in Figure S4a), it would be more beneficial to select a structurally similar, commercially available molecule to collect experimental data for model training. This would ensure more robust feedback to the AL framework.

Finally, while the focus on single-solvent electrolytes is a valuable starting point, exploring multi-solvent formulations to leverage synergistic effects could further enhance battery performance. If the AL workflow can be adapted to optimize solvent mixtures, it would significantly increase the impact and practical relevance of this study.

---

1. What are the noteworthy results?

The study presents an AL framework for efficiently screening electrolyte solvents for anode-free lithium metal batteries under fixed conditions of 1M LiFSI in Cu/LFP coin cells. It explores a vast chemical space of over 1 million virtual candidates and identifies four promising electrolyte solvents with cycling performance comparable to state-of-the-art systems (need to show a comparison plot in SI). The integration of Bayesian Model Averaging (BMA) to manage noisy and data-scarce environments is a key highlight. The

experimental data reveals a consistent preference for ether-based solvents, and the work showcases a scalable strategy for electrolyte discovery using sequential optimization.

2. Will the work be of significance to the field and related fields? How does it compare to the established literature?

Yes, this work is significant in this field. It provides a data-efficient approach to optimize electrolyte formulations, which is crucial for improving anode-free lithium metal batteries. Compared to existing literature, this work differentiates itself by employing active learning rather than traditional high-throughput or brute-force experimentation. While similar approaches have been demonstrated in catalyst optimization and drug discovery, its application in battery research—particularly for liquid electrolytes under fixed conditions—is novel and valuable. However, the study's focus is limited to 1M LiFSI in Cu/LFP coin cells, with no exploration of other lithium salts, salt concentrations, or cathode materials, which should be clearly stated as a limitation.

3. Does the work support the conclusions and claims, or is additional evidence needed?

The work generally supports its conclusions, particularly in identifying high-performing solvent candidates through iterative experimental validation. However, the manuscript lacks clarity on how the AL workflow operates in practice, particularly the process described in Figure 1c. Providing code or pseudocode on GitHub would enhance transparency and reproducibility. Moreover, the decision to focus solely on single-solvent formulations is justified but could be complemented by exploring multi-solvent mixtures to potentially improve performance further. Including more experimental validation for different functional group classes would strengthen the claims.

4. Are there any flaws in the data analysis, interpretation, and conclusions? Do these prohibit publication or require revision?

The overall data analysis is sound, but there are some areas that need improvement:

Clarification of the AL workflow: The process in Figure 1c is not fully explained in the main text. Providing step-by-step details or code would significantly improve clarity.

Selection of promising candidates: The AL model suggests candidates with functional groups not present in the original dataset, which raises concerns about the generalizability of the model. It would be better to experimentally test structurally similar, commercially available molecules to improve model feedback. These issues do not prohibit publication but require major revisions to enhance the transparency and reproducibility of the study.

5. Is the methodology sound? Does the work meet the expected standards in your field?

The methodology is generally sound and meets the expected standards for data-driven electrolyte discovery. The use of Bayesian Model Averaging (BMA) over Gaussian Process (GP) models is appropriate for managing data-scarce environments. However, the study would benefit from more detailed explanations of the AL workflow and the decision criteria for selecting solvent candidates. Additionally, providing code for reproducibility is becoming a standard practice in machine learning-based research and should be considered here. A detailed discussion about the limitations should also be included.

6. Is there enough detail provided in the methods for the work to be reproduced?

- The methods section provides a general overview of the experimental procedures and the AL framework. However, there are missing details that hinder full reproducibility, such as: The exact implementation of the AL workflow in Figure 1c.
- How the virtual search space was filtered and prioritized when using eMolecules and PubChem databases.
- The criteria for selecting commercially available solvents.
- Providing code on GitHub or more detailed workflow steps would address this issue and make the study more transparent and reproducible.

Reply to Reviewers

Reviewer #1 (Remarks to the Author):

This study explores the selection of new solvents for metal lithium batteries using machine learning (ML) methods. While the research topic is compelling, there are several areas where the manuscript requires significant improvement. In particular, the description of how ML is applied to the experimental data is unclear, and the selection and quality of the training data are questionable. The paper appears to reflect a collaboration between experimental researchers and AI specialists; however, the integration of these two aspects is insufficient, resulting in a fragmented narrative. Detailed points are provided below:

Response: We appreciate the reviewer's feedback and recognize the importance of improving the clarity and integration of the machine learning (ML) and experimental aspects of our study. In this revision, we have made several key improvements to address these concerns. The point-by-point response to the reviewer's comments and concerns appear below.

1. Quality of Experimental Data
"As a starting point, a high-quality in-house dataset of 58 anode-free LMB cycling profiles" (page 3)
The definition of "high-quality data" is vague. It is critical to quantitatively assess the variability in battery performance when creating cells under identical conditions. Without this, the reliability and reproducibility of the dataset remain uncertain.

Response: We thank the reviewer for questions regarding the quality of initial training dataset. The in-house dataset contains replicated tests of the same electrolyte in identical conditions. For example, Table S2 show three testing data of 1,2-Dimethoxyethane with 2 M LiFSA in Cu/LFP cell. The standard deviation of the three replicates is 5.6%. Similarly, the standard deviation of four replicates of Fluorinated ether A with 1 M LiFSA in Cu/LFP cell is 8.9%. Such standard deviation is comparable with literature where multiple replicates are shown (https://doi.org/10.1038/s41560-020-0668-8, Fig. 7a, standard deviation of three replicates estimated to be ~2.5%; https://doi.org/10.1038/s41560-021-00962-y, Fig. 5j, standard deviation of four replicates estimated to be ~3.3%). The following discussion has been added to the "Design of experiments (DoE) for electrolyte optimization; datasets, and virtual search space" section in the manuscript:

"*The quality of in-house dataset is analyzed by standard deviation among replicated tests. For example, three replicated tests of 1,2-Dimethoxyethane with 2 M LiFSA in Cu/LFP cell show a standard deviation of 5.6% and four replicates of Fluorinated ether A with 1 M LiFSA in Cu/LFP cell show a standard deviation of 8.9% , which are comparable to literature.*"

2. Standard Testing Configuration

"To reduce complexity and focus primarily on improving lithium metal cycling stability, Cu/LFP was selected as the standard testing configuration" (page 5)

Despite this claim, Table 2 indicates that NMC cathodes were used in some cells. The manuscript lacks a clear explanation of why multiple cathode types were employed and omits details regarding the experimental conditions for these cases.

Response: We appreciate the reviewer's question about the presence of NMC data in Table S2. To clarify, data using NMC cathode are only presented in initial training dataset and were not used in subsequent optimization batches. The decision to include NMC cathode data in the initial dataset was made to enlarge the number of initial training dataset, ensuring a more robust starting point for the active learning framework. However, all electrolyte evaluations performed in later batches (1-7) were conducted exclusively in Cu/LFP cells as stated in the manuscript.

Furthermore, to account for differences in cathode selection and maintain consistency in model training, the discharge capacities were normalized to the full capacity of the respective cathodes. This approach ensured that the surrogate models could distinguish battery chemistry variations while mitigating potential biases arising from different cathode types. Importantly, the full capacity of each cathode was explicitly included as an input feature in the surrogate models to preserve battery chemistry distinguishability within the optimization process (discussed in Supplementary Note 3 in the SI). We have added the following clarification to the discussion of initial training dataset in the manuscript:

"*However, unlike later data (optimization batch 1 to 7) that were collected in standard condition, we should note that the in-house dataset do contain varied conditions, such as salt concentration of 2 M, NMC cathode and higher temperature. To account for such heterogeneity, the in-house dataset was featurized (details in Supplementary Note 3) before being used as initial training data for the AL workflow.*"

We have also added the cell configuration and testing conditions of Cu/NMC cells into experimental section:

"***Coin cell preparation:*** *... Cu/NMC (NMC811 or NMC111) cells were prepared in the following configuration: negative case || conical spring || spacer || spacer || Cu || 25 μL electrolyte || separator || 25 μL electrolyte || NMC || Al foil || Al-coated positive case.*

Electrochemical tests: *... Cu/NMC111 cells were cycled in a voltage window of 3-4.3 V. Cu/NMC811 cells were cycled in a voltage window of 3-4.4 V. Two formation cycle were performed at a current rate of C/10 before long term cycling at C/3.*"

3. Issues with Training Data

The selection of training data is crucial for constructing accurate ML prediction models. However, the dataset described in Table 2 is inadequate in several respects:
Many solvents, such as "Nonfluorinated ether A," are not sufficiently described.
Lithium salt concentrations vary between solvents, and the criteria for determining these concentrations are not explained. Since both solvent type and salt concentration can independently affect battery performance, the dataset as presented is unsuitable for training a robust ML model.

Response: We thank the reviewer for questions regarding training data. Indeed, the initial training dataset contains data points that were tested at conditions different from the standard testing condition used in the later batches (including differences in salt concentration, cathode selection and temperature). This is because in-house data were collected from different projects prior to this work and therefore naturally have slightly different testing conditions. For example, most electrolytes in Table S2 use salt concentrations of 1 M or 2 M because such salt concentrations are most commonly used in electrolyte literature, especially for anode-free lithium metal batteries (Table S1). Such compromise partially reflects the data scarcity dilemma in anode free battery research. Furthermore, rather than being a limitation, we believe this diversity in experimental conditions is beneficial for training a generalizable ML model. By exposing the model to a richer variety of experimental conditions, it learns to generalize better across different electrolyte compositions rather than being overfitted to a narrow set of conditions. If we had strictly limited the testing condition to be the same as later batches, many in-house data would be excluded, and the scale of in-house dataset would be reduced to 31 data points. This would risk diminishing the model's ability to capture trends across different electrolytes and testing conditions.

The data scarcity issue motivated us to use an active learning (AL; sequential Bayesian experimental design) approach instead of conventional ML approaches. As we discussed in introduction, the AL approach is better equipped to handle small and noisy dataset. Despite the variations in initial training dataset, the surrogate models used in our AL framework were specifically designed to accommodate such heterogeneity. As explained in more detail in the responses to later points, both solvent identity and salt concentration were explicitly provided as input features to the Gaussian process regression (GPR) surrogate models, ensuring that the models could learn dependencies between these factors and the target property. Moreover, in the AL workflow, the model is not trained in a single step as in some conventional ML approaches. As more and more data were fed in later batches, the model's accuracy improved. Our results demonstrate that the surrogate models successfully guided electrolyte selection despite the inherent variability in the training data, validating the robustness of the approach.

Regarding the use of nomenclature such as “nonfluorinated ether A,” their structures were withheld because they are involved in other unpublished work from our group. We have now updated Table S2 to disclose the names of all such solvents.

4. Application of Machine Learning

The manuscript does not clearly describe how ML methods were applied to the dataset in Table 2. Specifically, the handling of varying lithium salt concentrations within the training process remains ambiguous. This raises concerns about the reliability of the ML analysis.

Response: We appreciate the reviewer’s concern regarding the application of machine learning (ML) and the handling of varying lithium salt concentrations in the training process. The details of how all features, including solvent identity, salt concentration, and cathode capacity, were featurized and incorporated into the surrogate models are provided in Supplementary Note 3. To clarify, salt concentration was explicitly included as an independent input feature in the ML framework, alongside solvent molecular descriptors and cathode capacity. Before training, the input data were standardized using the Standard Scaler function from the scikit-learn Python library, ensuring that variations in salt concentration were properly accounted for without biasing the model’s predictions. This step is critical for allowing the model to learn the independent effects of salt concentration on battery performance. The machine learning approach used in this work is based on GPR with Bayesian model averaging (BMA) technique. This method integrates multiple covariance kernels, enabling the model to capture the effects of electrolyte composition, including both solvent identity and salt concentration, on battery performance. The model does not assume a fixed salt concentration but rather learns from the distribution of different salt concentrations across the dataset. Additionally, to improve clarity, the main text has been revised to explicitly state that salt concentration was incorporated as a feature during featurization. We have also expanded the Methods section in the manuscript to provide a detailed discussion on data processing and ML development part. Furthermore, to ensure full transparency and reproducibility, we have released our codes in this round of revision so that reviewers and readers can directly examine the implementation of the ML framework.

Modified text in manuscript (Page 6-7, ‘Design of experiments (DoE) for electrolyte optimization; datasets, and virtual search space’ section):

“*However, unlike later data (optimization batch 1 to 7) that were collected in standard condition, we should note that the in-house dataset do contain varied conditions, such as salt concentration of 2 M, NMC cathode and higher temperature. To account for such heterogeneity, the in-house dataset was featurized (details in Supplementary Note 3) before being used as initial training data for the AL workflow.*”

Text in SI (Supplementary Note 3: Featurization and data transformation):

"*The heterogeneous input data, consisting of solvent and salt molecule identities, salt concentration, amount of electrolyte, and theoretical capacity of cathodes was featurized into suitable numerical form and the target property was normalized before feeding to the surrogate models. The molecules are represented using the popular Morgan fingerprints, specifically a 1024-dimensional binary vector with radius cutoff of 2 (also called ECFP4 variant), generated by feeding SMILES text representation of molecules into the RDKit python library. However, since Gaussian processes (GPs) tend to suffer from high computational costs when dealing with such high-dimensional data, the 1024-dimensional Morgan fingerprints for solvents and salts were reduced to 10 dimensions each using the principal component analysis (PCA) analysis technique. The other inputs — salt concentrations, amount of electrolyte, and theoretical capacity of cathodes (See note b of Table S2), were directly used and concatenated with reduced molecular representation resulting in a tabular input data. Therefore, the input data consists of a combination of continuous (composition and cycling condition) and discrete variables (molecular fingerprints). These input features corresponding to both labeled and unlabeled datasets were then standardized using the Standard Scaler function (transformed features have mean = 0 and unit s.d.) available in the scikit-learn python library by first fitting with respect to the labeled data and then transforming with respect to it. Since, the objective of this work is to optimize (maximize) based on discharge capacity that differs for different cathodes, the output was normalized with respect to the theoretical capacity.*"

5. Interpretation of ML Results

The manuscript does not adequately explain how the solvents listed in Table S3 were identified as promising candidates based on the ML analysis. Detailed discussions on the model's interpretability and decision-making process are essential to validate these predictions.

Response: We appreciate the reviewer's request for further clarification on how the ML model identified promising solvent candidates and how its decision-making process was validated. To clarify, the final solvent molecules were selected using the following multi-criteria decision process, as discussed in the manuscript: 1) suggestions from AL framework — the selection process from AL balanced high-confidence predictions (exploitation of already promising chemical motifs) and molecules from uncertain chemical regions (exploration of new solvent classes); 2) ranking within the ML-predicted top candidates — the AL framework produced a ranked list of solvents based on their predicted performance in anode-free lithium metal batteries and associated uncertainties. We considered candidates from top 5000 or top 10000 ranked molecules in the list; 3) cost and availability constraints — we prioritized compounds that were commercially available or synthesizable at a reasonable cost, ensuring that they could be experimentally tested without excessive delays; 4) filtering based on molecular diversity — to avoid redundancy, we selected structurally diverse molecules rather than picking multiple

solvents with highly similar molecular structures. If two molecules were chemically similar, only one was retained.

Moreover, while a model interpretability technique such as SHapley Additive exPlanations (SHAP) analysis was not directly used to guide the selection process, it played a crucial role in interpreting the ML model's decision-making. The SHAP analysis (detailed on Page 15 of the main text and Supplementary Note 10 in the SI) helped identify key molecular features that were correlated with improved electrolyte performance. Notably, the solvent molecular weight (MW_{solv}) was found to be the most influential descriptor and three specific molecular substructures — methoxy groups (-OCH_3), trifunctional carbon centers, and tetrafunctional carbon centers — were identified as strongly associated with improved performance. Figure 2i in the main text quantifies the fraction of selected molecules that contain these SHAP-identified features, showing strong alignment between ML-driven predictions and experimental performance. To improve clarity, we have slightly expanded the discussion in the manuscript to explicitly connect SHAP findings with the experimental validation of ML-selected candidates.

Text in manuscript (Page 8, 'Progress of sequential optimization using AL framework'):

"*To evaluate how the AL framework transitions from exploration to exploitation behavior, distribution of aggregated uncertainty (σ^{aggr}) values (defined in equation S11) for suggested compounds in each batch are shown using violin plots in Fig. S5a. As expected of performing AL using the EI acquisition function, the initial batches prioritize exploration, while later stages favor exploitation. The median of σ^{aggr} values (depicted by circles in each violin) decrease steadily across successive batches and eventually plateau between batches 5 and 6. During the first batch, the aggregated uncertainties are higher, thereby resulting in high EI^{aggr} (violin plot in Fig. S5b for all batches) and higher tendency for exploration. The decrease in σ^{aggr} in subsequent batches contribute to reduction in EI^{aggr} and preference for exploitation over exploration. The trend can be further corroborated through the t-SNE plots depicting exploitation ratio ($r_{exploit}$) defined in equation S13 in the SI, as shown in Fig. S6. These plots show that, in subsequent batches, points with higher exploitation ratios tend to cluster closer to the training data. This indicates that the AL framework is successfully focusing on regions where higher target properties are found, highlighting the promising areas for anode-free battery electrolytes.*"

Text in manuscript (Page 11, 'Progress of acquisition using AL suggestions'):

"*From the list of solvent molecules suggested by the AL framework obtained through trade-offs in exploitation and exploration strategies, about 10 compounds are selected and tested based on following criteria: 1) commercially purchasable (available in the United States with price < $1000 /g and lead time < 2 weeks; further details in Supplementary Note 2) solvent molecules are prioritized in the first 6 batches to accelerate progress; 2) when multiple similar compounds*

are recommended (such as methyl propyl ether, methyl butyl ether, methyl pentyl ether and methyl hexyl ether), only half of them with higher ranking are tested to promote diversity in the labeled dataset; 3) screening is performed sequentially, starting with the highest-ranked molecules and continuing until the desired number of purchasable candidates is selected, typically from the top 5,000 but not exceeding the top 10,000 ranked molecules; 4) chlorine containing compounds are purposefully omitted after batch 5 because all chlorinated solvents tested in previous batches show poor electrochemical profiles indicating continuous electrolyte degradation (Fig. S7). For each batch, selected compounds are purchased, paired with 1 M LiFSA in Cu/LFP cells and their discharge capacity data at 20th cycle is measured and fed to the surrogate model in the next AL campaign."

Modified text in manuscript (Page 15, 'Progress of acquisition using AL suggestions' section):

"*Moreover, we applied SHAP (SHapley Additive exPlanations) analysis to evaluate whether the surrogate model in the AL framework learns meaningful, physically consistent trends. Post batch 7 training, SHAP analysis identified solvent molecular weight (MW_{solv}) as the most significant feature (Fig. 2g and S14a), alongside solvent fingerprints (f_{solv}^{4}, f_{solv}^{5}, and f_{solv}^{6}; Fig. 2h) linked to molecular substructures methoxy, trifunctional, and tetrafunctional carbon centers, respectively, which are prevalent in high-performing electrolytes. Importantly, the SHAP feature rankings align with trends observed in experimental performance, suggesting the model captures genuine feature-property relationships rather than dataset frequency effects (Fig. 2i). For instance, although the tetrafunctional molecular substructure is present in only about 20% of all tested electrolyte solvent molecules (listed in Table S3), it leads to the highest average normalized discharge capacity among the three identified substructures. The presence of these substructures is also indicated in the molecules shown in the t-SNE plot in Fig, 2f through tick marks in different colors. Further SHAP analyses for the relevant molecular fingerprints are shown in Fig. S14 and discussed in detail in Supplementary Note 10…*"

Reviewer #2 (Remarks to the Author):

In this article, the authors demonstrate the use of active learning (AL) as an alternative approach to accelerate electrolyte discovery for anode-free LMBs. The study employs sequential Bayesian experimental design with Bayesian model averaging (BMA) to efficiently identify optimal candidates in typical data-scarce and noisy label settings. However, the computational study could still be improved to better align with the experimental data. There are also some experimental data should be revised carefully.

Response: We sincerely appreciate the reviewer's feedback and acknowledgment of our work in employing active learning (AL) for electrolyte discovery in anode-free lithium metal batteries. Regarding the alignment between computational and experimental data, we have carefully refined the discussion to ensure seamless integration of our Bayesian model averaging (BMA) approach with the experimental outcomes in the revised manuscript. For the experimental data, we have revised a few figures to incorporate relevant data. All necessary corrections or clarifications have been incorporated into the revised manuscript. The point-by-point response to the reviewer's comments appear below.

1. Figure 1b fails to indicate the specific electrolyte used. Please label it or provide a brief description in the manuscript.

Response: We thank the reviewer for question regarding Figure 1b. The curves shown in Figure 1b were not labeled because they are just for illustrative purposes and do not represent real electrolyte performance. We have modified the caption to clarify:

"***Figure 1****. … **b)** Illustrative examples for anode-free LMB capacity retention curves. The curves shown do not represent real data. Discharge capacity at 20th cycle normalized with respect to cathode theoretical capacity is selected as target property to balance the effects of initial Coulombic efficiency and long-term stability. …*"

2. How to understand the virtual chemical search space (~380000 unique molecules) across different batches? What do the data points and regions mean in the Figure 2a? Can the authors compare the current AL suggestions results with other machine learning algorithms, such as Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Light Gradient Boosting Machine (LightGBM)?

Response: We appreciate the reviewer's insightful question regarding the interpretation of the virtual chemical search space. We would like to clarify that the chemical search space (~380,000 unique molecules) was not static across different batches. Once a solvent molecule was tested, it was removed from the search space in subsequent batches to ensure that the AL framework did

not reselect previously evaluated candidates, which is a standard practice in AL-based optimization (https://doi.org/10.1016/j.cels.2020.09.007; https://doi.org/10.1002/adma.201702884). Moreover, the data points in Fig. 2a depict each of the unique solvent molecules in the search space and the different colors in this plot simply refer to the functional groups to which they belong. This aspect was inadvertently omitted in the original manuscript, and we have clarified in the revised version how the search space evolved during sequential optimization.

Regarding the suggested comparison with other machine learning algorithms, our AL framework was designed specifically for data-scarce and noisy-label settings using Bayesian model averaging (BMA) over Gaussian processes (GPs). This approach provides reliable uncertainty quantification, which is crucial for balancing exploration and exploitation in sequential experimental design. In contrast, frequentist models such as decision trees, RF, KNN, and LightGBM do not provide accurate aleatoric uncertainties for noisy experimental labels (https://doi.org/10.1039/C9SC00616H) and lack the posterior distributions necessary for BMA and expected improvement (EI)-based acquisition. This makes them unsuitable for direct integration into our AL workflow. However, in response to the reviewer's request, we compared the performance of these frequentist models with the averaged predictions from our ensemble of four surrogate models, which has been now included in the revised draft (Supplementary Note 2). The ensemble approach consistently outperformed the frequentist models on the in-house dataset, likely due to the data-hungry nature of models like LightGBM (Table 1 below).

Table 1: Comparison of performance of frequentist models with the GP surrogate models on in-house data (included in modified discussion in Supplementary Note 2):

Model	**5-fold CV RMSE**
GP (average of 4 surrogate models)	0.159 ± 0.040
Decision tree	0.234 ± 0.059
Random forest	0.208 ± 0.041
K-nearest neighbor (KNN)	0.271 ± 0.077
LightGBM	0.269 ± 0.063

Modified Fig. 2a caption:

"*2D t-SNE plots depicting location of suggested molecules (denoted by different scatter shapes) on the virtual chemical search space (~380000 unique molecules corresponding to each point in this reduced embedding space) across different batches. The virtual chemical space in the background has been colored according to different functional group classes (only top 7 populated classes shown here for clarity).*"

New text in manuscript (Page 7, 'Design of experiments (DoE) for electrolyte optimization; datasets, and virtual search space' section):

"*The tested molecules were removed from the virtual search space in subsequent batches to prevent reselection of such molecules.*"

Modified text in SI (Supplementary Note 2):

"*We also evaluated the performance of common frequentist ML models, including K-nearest neighbor regressor (5-fold CV RMSE: 0.271 ± 0.077), decision trees (0.234 ± 0.059), random forests (0.208 ± 0.041), and LightGBM (0.269 ± 0.063). However, these models underperformed compared to the Gaussian Process Regression (GPR) surrogate models, where the ensemble of four surrogate models achieved a lower 5-fold CV RMSE of 0.159 ± 0.040 on the in-house dataset (Fig. S2a). This discrepancy may be attributed to the data-intensive nature of these frequentist models. Consequently, they were not incorporated as surrogate models in our active learning (AL) workflow.*"

Modified text in SI (Supplementary Note 2):

"*The tested solvent molecules were removed from the virtual search space in subsequent batches as per standard practice followed in an AL-based optimization framework.*"

3. Where are the codes available on GitHub (https://github.com/AmanchukwuLab/AL-anode-free)? It is surprising that the authors did not mention any active learning processes or algorithms in the Methods section, making it difficult for other researchers to reproduce the results.

Response: We thank the reviewer for highlighting the importance of code availability and reproducibility. We acknowledge that the AL workflow, including its implementation details, was not explicitly mentioned in the Methods section, and so we have expanded this section in the revised manuscript to ensure that it contains a clear description of the AL processes and algorithms.

Regarding the code repository, we have uploaded all relevant datasets and codes to GitHub at https://github.com/AmanchukwuLab/AL-anode-free as part of this revision to ensure full reproducibility.

Modified Methods section in revised manuscript ('Data analysis and AL setup' subsection):

"*To accelerate electrolyte discovery for anode-free lithium metal batteries (LMBs), we employed an AL framework that iteratively refines predictions based on sequential experimental feedback. The AL workflow consists of the following steps:*

Creation of virtual search space*: The initial chemical search space (~1 million electrolytes) was constructed using small organic molecules filtered from eMolecules and PubChem databases. Molecules with undesirable functional groups (e.g., radicals, nitriles, aromatics) or low synthetic accessibility (RAscore < 0.01) were removed, resulting in ~380,000 unique solvent molecules for AL-driven optimization. Further details are provided in Supplementary Note 1.*

Feature representation and model training*: Molecular structures were represented using 1024-bit Morgan fingerprints (ECFP4), which were dimensionally reduced to 10 components via principal component analysis (PCA) to enhance computational efficiency for GP models. Other input features included salt identity, salt concentration, and electrolyte composition. The surrogate models were trained using GPR with four different covariance kernels: Radial Basis Function (RBF) + Exponential Sine Squared (ESS), Matern-3/2, Rational Quadratic (RQ), and Pairwise Kernels. Bayesian model averaging (BMA) was applied to integrate predictions from all four GP models, balancing their strengths and improving uncertainty estimation. Further details are provided in Supplementary Notes 2, 3, and 5.*

Acquisition strategy using expected improvement (EI)*: The trained surrogate models generated predicted means and uncertainties for the unlabeled search space, and molecules were ranked based on their EI scores. The aggregated EI (EI^{aggr}) was computed using weighted model predictions from BMA, guiding the selection of molecules for experimental testing. Further details are provided in Supplementary Note 6.*

Sequential batch selection and search space evolution*: In each batch, top-ranked solvent molecules (based on EI^{aggr}) were experimentally tested in Cu/LFP coin cells with 1 M LiFSA. Tested molecules were permanently removed from the search space before the next iteration to ensure non-redundant selection. Initial batches prioritized exploration (diverse functional groups such as amides, amines, imides), while later batches shifted toward exploitation (predominantly ether-based solvents) as the model refined its predictions.*

Experimental feedback and model refinement*: Battery cycling performance (normalized discharge capacity at 20th cycle, C_{norm}^{20}) was measured and added to the training dataset. The AL loop was repeated across six batches, refining predictions in each iteration until model convergence was observed. Further details are provided in Supplementary Note 2.*

Computational implementation: *All codes utilized in this study were built using various libraries available in the Python (version 3.9.0) programming language. Morgan fingerprints used for representing molecules for training surrogate models and for t-SNE visualization were generated using RDKit library. Scikit-learn (version 0.23) library was employed for standardizing datasets, training the GPs using different covariance kernels (detailed explanation in Supplementary Note 5), calculating error metrics. The reduced embeddings of chemical space for the t-SNE plots were generated using OpenTSNE library. The numpy and scipy libraries were*

used for performing most numerical analyses such as Wasserstein distances, cumulative and probability distribution functions (for calculating EI). All plots in the manuscript were made using matplotlib, seaborn, or plotly libraries in python or using Origin."

4. The SEM images of lithium deposition for the four electrolytes illustrated in Figure 3c-f are not sufficient to demonstrate that all electrolytes form a tight and uniform stack of particles. To maintain experimental rigor, it is recommended to supplement the cross-sectional SEM images of these four electrolytes.

Response: We appreciate the reviewer for recommendation regarding cross sectional SEM characterization of lithium deposition morphology. We have updated Figure 3c-f to include cross section SEM images as inset as shown below. The cross section SEM images support the conclusion that AL-discovered electrolytes enable compact and uniform lithium deposition morphology.

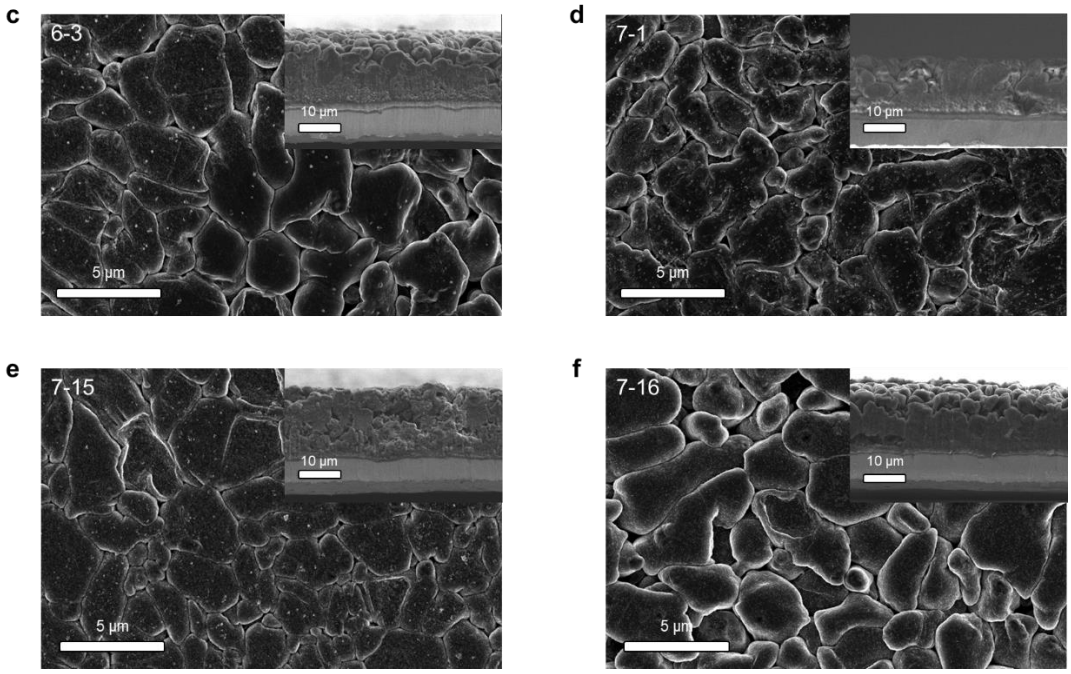


Figure 3 | Performance of AL-discovered electrolytes. *… c-f) SEM images of lithium metal deposited in (c) 1 M LiFSA in 6-3, (d) 1 M LiFSA in 7-1, (e) 1 M LiFSA in 7-15 and (f) 1 M LiFSA in 7-16 using a Li/Cu cell with a current density of 1 mA/cm^2 to a capacity of 1.5 mAh/cm^2. The inset shows corresponding cross section SEM images of lithium metal sample.*

5. Figure 4c compares ionic conductivity of 1 M LiFSA in AL discovered solvents as a function of temperature. However, no description is found in the article. Also, Figure 4d compares diffusivity of lithium ion (DLi) and FSA anion (DFSA) and lithium transference

number (tLi) of AL discovered electrolytes at 1 M LiFSA concentration, and again no description is found in the article. The authors should provide a brief description and analysis in the manuscript.

Response: We thank the reviewer for pointing out the lack of discussion on Figures 4c and 4d. The following discussions have been added to the manuscript.

"*The ion transport properties of AL-discovered electrolytes were also studied. Figure 4a shows the ionic conductivity of 6-3, 7-1, 7-15 and 7-16 with 1 M LiFSA measured by electrochemical impedance spectroscopy (EIS). At 20 °C, 7-1 electrolyte has the highest ionic conductivity of 3.5 mS/cm and the other three electrolytes show acceptable conductivities of 0.7~1.0 mS/cm. Figure 4d shows the ion diffusivity and lithium transference number of AL-discovered electrolytes measured by PFG NMR, where the ion diffusivity in general decreases with increasing viscosity (Table S5) from 6-3 to 7-16. All four electrolytes show high lithium transference number close or higher than 0.5, which agrees with the solvation structure rich in ion pairing as discussed earlier.*"

6. In the Supporting Information and Notes, the authors presented the Different kernels used in GPR and performance comparison, Bayesian model averaging (BMA), Oracle or acquisition functon(EI) and related terms, and so on. I would suggest the authors share the relevant database, codes and atomistic models online.

Response: We appreciate the reviewer's suggestion regarding data and code sharing to enhance reproducibility and transparency. In this revision, we have ensured that all relevant datasets, codes, and models are made publicly available (https://github.com/AmanchukwuLab/AL-anode-free).

Reviewer #3 (Remarks to the Author):

The paper demonstrates an active learning approach to accelerate the discovery of optimal electrolyte solvents for anode-free lithium metal batteries (LMBs). Using Bayesian Model Averaging (BMA), the study efficiently explores a chemical space of over 1 million candidates, identifying four high-performing solvents through iterative Cu/LFP coin cell tests. The main work focuses on screening electrolyte solvents paired with a 1M LiFSI salt to optimize capacity retention and improve battery cycling performance. However, major revisions are required to improve the clarity of the AL workflow, particularly regarding the implementation details shown in Figure 1c. Providing the code on GitHub, if possible, would significantly enhance transparency and reproducibility.

Response: We thank the reviewer for their constructive feedback and for recognizing the significance of our AL approach in electrolyte discovery for anode-free lithium metal batteries. We appreciate the emphasis on improving the clarity of the AL workflow, particularly the implementation details in Fig. 1c-d, and we have revised the manuscript to provide a more detailed explanation. Additionally, we acknowledge the importance of transparency and reproducibility in ML-based research and have made the relevant codes available on GitHub in this revision to support further validation. The point-by-point response to the reviewer's comments and concerns appear below.

1. The study focuses primarily on 1M LiFSI in Cu/LFP coin cells, and I did not notice a systematic exploration of other lithium salts, salt concentrations, or cathode materials in the main text or supplementary information. While this fixed condition is reasonable given the scope of the AL model, more detailed information on how the workflow operates is necessary. For example, when screening solvents from eMolecules and PubChem, the first priority should be liquid organic solvents besides the criteria outlined in Table S6. (in other words, it is not necessary to screen a huge number of molecules if we know they are solid at RT).

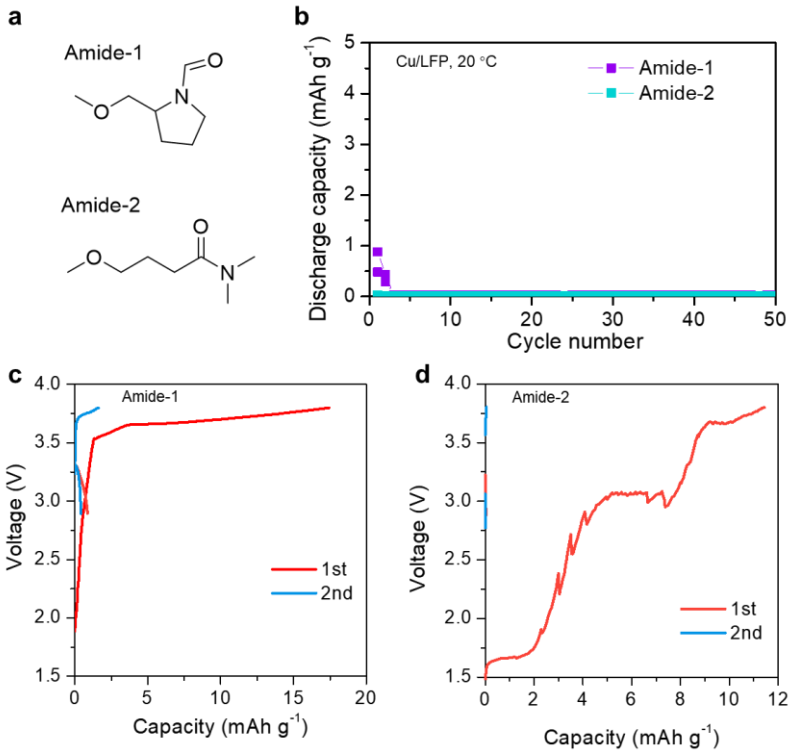
Response: We thank the reviewer for their comments and for recognizing the rationale behind fixing the electrolyte composition to 1 M LiFSA in Cu/LFP cells in our active learning (AL) workflow. We agree that systematically exploring different lithium salts, salt concentrations, and cathode materials would provide valuable insights. However, given the experimental constraints and the need to establish a well-defined optimization framework, we focused on single-solvent electrolyte screening under fixed conditions (detailed response in queries #3 and #5). We have clarified this limitation in the revised manuscript.

Regarding the molecule screening criteria, we acknowledge the importance of prioritizing liquid organic solvents. But strict pre-screening based on melting point or boiling point was not

enforced mainly because such data for electrolytes are not widely available. Instead, we employed a computational screening based on ionic conductivity predictions when constructing search space. Utilizing a deep-learning-based predictive model developed in our previous work (doi.org/10.26434/chemrxiv-2024-vqtc7), only candidates with high enough predicted ionic conductivity (top 1 million electrolytes with predicted conductivity > 0.8 mS/cm containing ~380000 unique molecules) were included in the search space, which should effectively filter out solid molecules that are unlikely to support ion transport. Such filtration seems to be successful as 14 out of 65 tested candidates fail to dissolve 1 M LiFSA but only 1 of them is a solid compound.

2. Additionally, the dataset used in this study primarily consists of ether derivatives. When the model suggests candidates with entirely different functional groups that are not commercially available (as explained in Figure S4a), it would be more beneficial to select a structurally similar, commercially available molecule to collect experimental data for model training. This would ensure more robust feedback to the AL framework.

Response: We thank the reviewer for suggestions to verify candidates with distinct functional groups. The major outstanding functionality that has not been explored is amide, which takes ~40% of the search space (Figure S4a) but no amide compound was tested due to their limited availability. Following the reviewer's recommendation, we reinvestigated the top suggestions from batch 1 to 7 and filtered out 2 amide compounds that are purchasable but at high prices. As the figure shown below, these two amide compounds (named as Amide-1 and Amide-2) were purchased and tested in the standard Cu/LFP cells. However, both of them show very poor cycling performance with almost no capacity left after the first cycle. In fact, there is no report of amides as main electrolyte solvent for anode free lithium metal batteries in literature (Table S1), which also suggests amide compounds might not be promising electrolyte solvent candidates. While we cannot eliminate the possibility of finding promising candidates from other solvent classes, it seems reasonable for the AL model to build the preference towards ether solvents.



Performance of amide compounds. a) Chemical structure of two purchasable amide compounds. b) Capacity retention of amide compounds with 1 M LiFSA in Cu/LFP cells. c) Voltage profiles of the cell using Amide-1 electrolyte. d) Voltage profiles of the cell using Amide-2 electrolyte.

3. Finally, while the focus on single-solvent electrolytes is a valuable starting point, exploring multi-solvent formulations to leverage synergistic effects could further enhance battery performance. If the AL workflow can be adapted to optimize solvent mixtures, it would significantly increase the impact and practical relevance of this study.

Response: We thank the reviewer for this insightful suggestion regarding the optimization of multi-solvent electrolyte formulations. We agree that leveraging synergistic effects through solvent mixtures could further enhance battery performance. However, adapting our AL framework to optimize multi-component formulations presents significant challenges that are beyond the scope of this study. Our current framework was designed specifically for single-solvent electrolytes, where solvent identity serves as a direct input feature using Morgan fingerprints. However, unlike single-solvent formulations, multi-solvent electrolytes introduce nonlinear interactions between different solvents, affecting properties such as solubility, viscosity, ionic conductivity, and electrochemical stability. Standard molecular featurization (i.e., Morgan fingerprints) used in our AL framework does not capture solvent interactions explicitly — which would require developing new featurization strategies and custom ML models for capturing (e.g., mixture descriptors, interaction terms) before we can begin training a reliable model. Another major technical hurdle in optimizing multi-component formulations is the near-

infinite number of possible solvent combinations, making it impractical to enumerate a comprehensive search space. For instance, even with a modest selection of 50 solvent candidates, a binary mixture search would involve 2,450 unique combinations (choosing 2 out of 50), and a ternary mixture search would involve 196,000 combinations (choosing 3 out of 50). Adding ten fixed m ratio for each combination would have already expand search space to be larger than the current work. To fully realize the advantage of formulation, additives and diluents should also be included (https://doi.org/10.1002/adma.201706102; https://doi.org/10.1002/adfm.202310516), which would require a solvent scope much greater than 50 candidates. Therefore, a more feasible and cost-effective alternative for future work is inverse design, where an optimal formulation is directly generated based on desired electrolyte properties. This is an avenue we plan to explore in our future research. Additionally, adapting the current AL framework to solvent mixtures would also require a new training dataset to address the new dimensions in the electrolyte formulation. But there is a lack of data on multi-solvent formulations for anode-free battery in our lab or in the literature (only ten in Table S1). Considering the much complicated search space, collecting a comprehensive in-hose dataset for formulation would take significantly more time and efforts. To address this important point, we have added a discussion on this limitation in the revised manuscript.

4. What are the noteworthy results? The study presents an AL framework for eAiciently screening electrolyte solvents for anode-free lithium metal batteries under fixed conditions of 1M LiFSI in Cu/LFP coin cells. It explores a vast chemical space of over 1 million virtual candidates and identifies four promising electrolyte solvents with cycling performance comparable to state-of-the-art systems (need to show a comparison plot in SI). The integration of Bayesian Model Averaging (BMA) to manage noisy and data-scarce environments is a key highlight. The experimental data reveals a consistent preference for ether-based solvents, and the work showcases a scalable strategy for electrolyte discovery using sequential optimization.

Response: We thank the reviewer for highlighting the key contributions of our study. We appreciate the recognition of our active learning (AL) framework and its ability to efficiently navigate a vast chemical space for electrolyte discovery. As noted, the integration of Bayesian Model Averaging (BMA) plays a crucial role in managing noisy, data-scarce environments, and our experimental findings consistently point to the preference for ether-based solvents. Due to the varied testing condition in anode free LMB literature, we performed an in-house comparison to a state-of-the-art fluorinated ether control (F5DEE) to benchmark our work, which is shown in Fig. 3 in the manuscript.

5. Will the work be of significance to the field and related fields? How does it compare to the established literature?

Yes, this work is significant in this field. It provides a data-eAicient approach to optimize electrolyte formulations, which is crucial for improving anode-free lithium metal batteries. Compared to existing literature, this work diAerentiates itself by employing active learning rather than traditional high-throughput or brute-force experimentation. While similar approaches have been demonstrated in catalyst optimization and drug discovery, its application in battery research—particularly for liquid electrolytes under fixed conditions—is novel and valuable. However, the study's focus is limited to 1M LiFSI in Cu/LFP coin cells, with no exploration of other lithium salts, salt concentrations, or cathode materials, which should be clearly stated as a limitation.

Response: We thank the reviewer for acknowledging the significance of our work and its novelty in applying active learning for electrolyte optimization in anode-free lithium metal batteries. To address the reviewer's suggestion, we have added a dedicated discussion section in the manuscript explicitly stating the limitations of our study, including the focus on 1 M LiFSI in Cu/LFP coin cells without exploring alternative lithium salts, salt concentrations, or cathode materials. This addition provides a clearer perspective on the scope of our approach and potential future directions.

New 'Discussion' section in the revised manuscript (Page 20-21):
"However, several limitations must be acknowledged. First, all experimental validations for AL suggestions were conducted using Cu/LFP cells, which, while widely used for laboratory-scale evaluation, may not accurately represent electrolyte performance with high-voltage cathodes (such as LiNixMnyCo1−x−yO2) that are proposed for commercial LMBs. Future studies should assess the compatibility of these electrolytes with high-voltage cathodes. Additionally, our study was limited to single solvent and single lithium salt (LiFSA) at a fixed 1 M concentration, which does not make use of formulation strategies well developed in electrolyte literature. Therefore, the use of alternative salts, higher salt concentration additives and diluents may further enhance interphase stability and cycle life. Another limitation is the practical accessibility of identified solvents. The AL framework prioritized purchasable molecules to accelerate the discovery process. However, many high-ranking candidates were unavailable commercially. While we successfully synthesized several promising candidates, future studies could integrate predictive models for solvent synthesizability and salt solubility to further refine selection criteria. Additionally, our study optimized capacity retention as a single target. While capacity retention might be the bottle neck for commercializing anode-free LMBs, practical battery applications also demand on other properties such as rate capability, working temperature window and safety. Expanding the AL framework to a multi-objective optimization paradigm would make electrolyte optimization more relevant to real-world applications."

6. Does the work support the conclusions and claims, or is additional evidence needed?

The work generally supports its conclusions, particularly in identifying high-performing solvent candidates through iterative experimental validation. However, the manuscript lacks clarity on how the AL workflow operates in practice, particularly the process described in Figure 1c. Providing code or pseudocode on GitHub would enhance transparency and reproducibility. Moreover, the decision to focus solely on single-solvent formulations is justified but could be complemented by exploring multi-solvent mixtures to potentially improve performance further. Including more experimental validation for diAerent functional group classes would strengthen the claims.

Response: We thank the reviewer for their thoughtful comments and for highlighting areas where additional clarification would strengthen the manuscript. In response to the concern regarding the AL workflow (Fig. 1d), we have expanded the Methods section to provide a clearer explanation of how the workflow operates in practice. Additionally, we have made the code available on GitHub to enhance transparency and reproducibility, as recommended by the reviewers (https://github.com/AmanchukwuLab/AL-anode-free).

Regarding the focus on single-solvent formulations, we acknowledge that exploring multi-solvent mixtures could potentially improve performance further. However, as noted in our previous responses, the AL framework was designed specifically for single-solvent systems and adapting it to multi-solvent mixtures would require developing specialized ML architectures for capturing non-linear solvent interactions in the formulations as well as optimizing on practically infinite search space. This is beyond the scope of the current study, but we have highlighted this limitation and our future outlook on inverse design for multi-component systems in the revised manuscript.

Finally, we have performed additional experimental validation for two amides based on the reviewer's earlier suggestion (query #2). However, both amides performed poorly in our tests, which agrees with literature trend that amides are likely not good electrolyte solvent candidates for anode-free LMB.

7. Are there any flaws in the data analysis, interpretation, and conclusions? Do these prohibit publication or require revision?
The overall data analysis is sound, but there are some areas that need improvement: Clarification of the AL workflow: The process in Figure 1c is not fully explained in the main text. Providing step-by-step details or code would significantly improve clarity. Selection of promising candidates: The AL model suggests candidates with functional groups not present in the original dataset, which raises concerns about the generalizability of the model. It would be better to experimentally test structurally similar, commercially available molecules to improve model feedback. These issues do not prohibit publication but require major revisions to enhance the transparency and reproducibility of the study.

Response: We thank the reviewer for their constructive feedback and appreciate their acknowledgment that the overall data analysis is sound. We have made several revisions to enhance the clarity, transparency, and reproducibility of our study, addressing the concerns raised.

1. **Clarification of the AL workflow (Fig. 1d):** We have expanded the Methods section to provide a more detailed step-by-step explanation of how the AL workflow operates in practice. Additionally, to further enhance transparency, we have made the relevant code available on GitHub, allowing for reproducibility into our framework.
2. **Selection of promising candidates and model generalizability:** Our AL framework inherently balances exploration (selecting structurally novel candidates) and exploitation (selecting candidates similar to high-performing molecules from previous batches). This transition towards exploitation in later batches has been discussed in the main text and Supplementary Note 8. To quantitatively assess this behavior, we calculated the average of maximum Tanimoto coefficient values for each batch relative to the in-house dataset. The values show that the framework initially explores more diverse molecules (batch 1: 0.12, batch 2: 0.38) but gradually converges toward structurally similar compounds in later batches (~0.5-0.6). This new analysis has been incorporated into the revised manuscript and presented in Figure S11(b) (attached below).

We have added a discussion section in the revised manuscript that contains detailed limitations of our work for enhanced transparency. We believe these revisions and explanations address the reviewer's concerns and significantly enhance the clarity and robustness of the study.

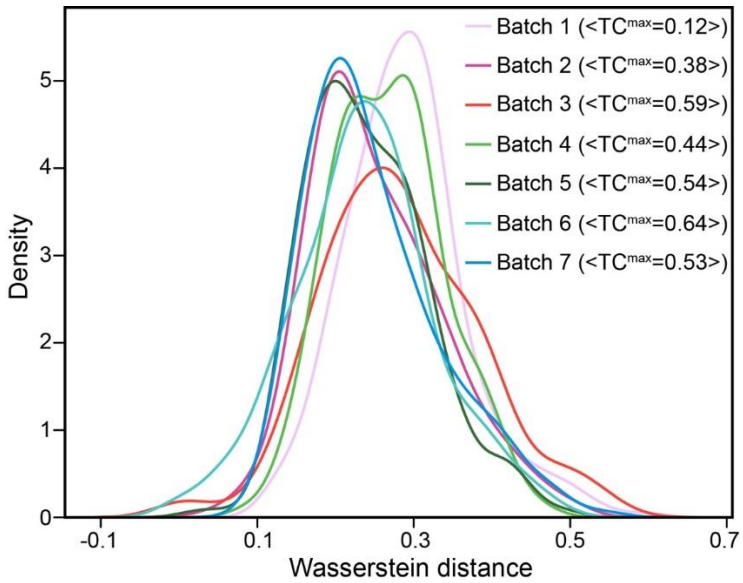


Figure S11*. Density of Wasserstein distance distribution for … **b)** solvent molecules in different batches with respect to in-house dataset.* $< TC^{max} >$ *= average of maximum Tanimoto coefficient values with respect to electrolyte solvent molecules in the in-house dataset.*

8. Is the methodology sound? Does the work meet the expected standards in your field? The methodology is generally sound and meets the expected standards for data-driven electrolyte discovery. The use of Bayesian Model Averaging (BMA) over Gaussian Process (GP) models is appropriate for managing data-scarce environments. However, the study would benefit from more detailed explanations of the AL workflow and the decision criteria for selecting solvent candidates. Additionally, providing code for reproducibility is becoming a standard practice in machine learning-based research and should be considered here. A detailed discussion about the limitations should also be included.

Response: We thank the reviewer for recognizing the soundness of our methodology and its alignment with the standards in data-driven electrolyte discovery. We appreciate the suggestion to provide a more detailed explanation of the AL workflow, and the criteria used for selecting solvent candidates. We have expanded on these aspects in the revised manuscript to enhance clarity (expanded Methods section in the revised manuscript). Additionally, we acknowledge the increasing importance of code availability for reproducibility in machine learning-based research and hence have released all relevant codes to support transparency and further validation on the GitHub repository. A discussion on the limitations of our approach has also been incorporated in the revised manuscript to provide a balanced perspective.

9. Is there enough detail provided in the methods for the work to be reproduced?
• The methods section provides a general overview of the experimental procedures and the AL framework. However, there are missing details that hinder full reproducibility, such as: The exact implementation of the AL workflow in Figure 1c.
• How the virtual search space was filtered and prioritized when using eMolecules and PubChem databases.
• The criteria for selecting commercially available solvents.
• Providing code on GitHub or more detailed workflow steps would address this issue and make the study more transparent and reproducible.

Response: We appreciate the reviewer's concerns regarding the level of detail in the methods section and have taken steps to improve transparency and reproducibility.

To clarify the AL workflow depicted in Fig. 1d, we have expanded the Methods section in the revised manuscript, providing additional details on how the iterative selection of candidates was performed using ML model predictions and experimental feedback. This includes specifying how new experimental data were incorporated into retraining cycles to refine model predictions.

Regarding the filtering and prioritization of the virtual search space, we have elaborated on the screening criteria applied to the eMolecules and PubChem databases. The details on removing

structurally incompatible compounds (e.g., highly reactive moieties, charged species, transition metals) and prioritizing molecules based on chemical feasibility are now explicitly stated in Supplementary Note 10. We emphasize that this curation process was necessary to ensure that the final candidates were realistic for electrolyte applications.

Regarding the selection of commercially available solvents, the manuscript (Page 10) and the SI (Supplementary Note 2) contained detailed discussion as to how solvents were selected for battery testing in the present study (please refer to the excerpts from manuscript and SI below).

Finally, to enhance reproducibility, we have ensured that all relevant ML models, datasets, and codes are provided in a public GitHub repository (link included in the revised manuscript). The repository includes trained models, Jupyter notebooks containing code snippets and stepwise instructions to allow researchers to reproduce the key steps in our workflow.

Text in manuscript (Page 10, 'Progress of acquisition using AL suggestions' section):
"*From the list of solvent molecules suggested by the AL framework, about 10 compounds are selected and tested based on following criteria: 1) commercially purchasable (available in the United States with price < $1000 /g and lead time < 2 weeks; further details in Supplementary Note 2) solvent molecules are prioritized in the first 6 batches to accelerate progress; 2) when multiple similar compounds are recommended (such as methyl propyl ether, methyl butyl ether, methyl pentyl ether and methyl hexyl ether), only half of them with higher ranking are tested to promote diversity in the labeled dataset; 3) screening is performed sequentially, starting with the highest-ranked molecules and continuing until the desired number of purchasable candidates is selected, typically from the top 5,000 but not exceeding the top 10,000 ranked molecules; 4) chlorine containing compounds are purposefully omitted after batch 5 because all chlorinated solvents tested in previous batches show poor electrochemical profiles indicating continuous electrolyte degradation (Fig. S7)…*"

Text in SI (Supplementary Note 2):
"***5. Feedback and other considerations in acquisition of suggested candidates:*** *It was found that most of the suggested molecules are not commercially purchasable within our budget constraints despite the search space being derived from well-known eMolecules and PubChem molecular repositories and removing difficult to synthesize molecules containing moieties such as radicals. Therefore, purchasable candidates were selected from a pool of molecules with top 5,000 (or even top 10,000 for batches 3 and 4) rank according to* EI^{aggr}*. It was performed by providing this list of solvent SMILES into the eMolecules search engine (https://search.emolecules.com/), in the 'List Search' tab and selecting 'SMILES' under 'Data Type/Search By' filter criteria. This provides a list of only purchasable molecules with all possible vendors it can find. For such molecules, we manually went through each of the vendors listed and selected molecules with price < $1000/g and lead time < 2 weeks for battery testing*

(procedure discussed in the Materials and Electrochemical characterizations subsections under the Methods section in the main text)…"

Reviewer #1 (Remarks to the Author):

Part of the concerns have been solved by suitable revision of the manuscript. however, there is still technicall problesm in the manuscript. Alhtough the concept and strategy of the present work looks interesting, there is serious concerns for the way of accumulating the inital datasets for constructing AI model.

1: type of cathode material

"The decision to include NMC cathode data in the initial dataset was made to enlarge the number of initial training dataset, ensuring a more robust starting point for the active learning framework. "

The degradation mechanism of NMC/Li cells are quite different with that of LFP/Li cells. In general, even using the same electrolyte material and setting the same cycling condition, the cycle life is quite different by the choice of cathode materials and charging cutoff voltage condition. In particular, in NMC811/Li cell, in which cutoff voltage set to be 4.4V, the oxidative decomposition of electrolyte occurs, and this induces the complicated chemical crossover phenomenon. Thus, the investigation should be completed by focusing on sole cathode material. The reviewer recommends that the AI model should be reconstructed focusing on only data sets of LFP/Li cells.

Response: We thank the reviewer for the insightful recommendation regarding the specificity of our dataset towards LFP/Li cells. As requested, we conducted an additional active learning (AL) campaign by excluding all NMC data from the initial dataset. We observed minimal degradation in the surrogate model performance for out-of-batch data upon removing the NMC data, indicating limited influence of the NMC data inclusion, e.g., average 5-fold CV RMSE of 0.2558 and 0.2614 for batch 6 of NMC-included and NMC-excluded AL campaigns, respectively. We have now added these new data as Figure S20 in the revised SI (attached below). Finally, the total number of high-performing electrolytes discovered from this new AL campaign remained unchanged at seven – compounds with IDs 5-7, 6-3, 7-1, 7-12, 7-14, 7-15, and 7-16. We have incorporated this new discussion in Supplementary Note 11.

We have publicly shared the datasets, Jupyter notebooks and surrogate model checkpoints obtained from both the original and the NMC-excluded AL campaigns via our GitHub repository. Users can thereby select models and datasets based on their specific application needs.

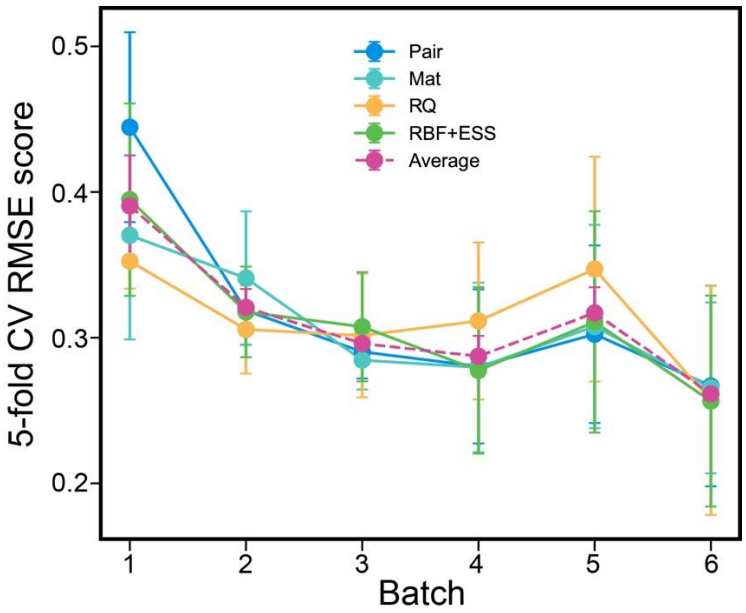


Figure S20*. 5-fold cross-validation (CV) out-of-batch performance of GP surrogate models as a function of batches. The error bars for each kernel (solid lines) denote standard deviation (s.d.) in root mean squared error (RMSE) of model predictions on 5-fold CV splits, and error bars for the average performance (dashed purple line) denote s.d. in RMSE of all 5-fold CV models. Pair = pairwise kernel, Mat = matern-3/2 kernel, RQ = rational quadratic kernel, RBF = radial basis function, ESS = exponential-sine squared kernel.*

2: training dataset

"This is because in-house data were collected from different projects prior to this work and therefore naturally have slightly different testing conditions."

This explanation is not a reasonable excuse. The arrangement of datasets is the most important point for constructing AI models. The author should handle well-arranged and organized data sets for constructing AI models. The reviewer considers it is inappropriate to use the data that is obtained by other projects with different purposes.

Response: We thank the reviewer for their comment. We would like to note that there is no single way to build a dataset for machine learning applications. There are pros and cons to each approach. In fact, it is very common practice among the machine learning community even while using active learning techniques to successfully leverage open source or literature data that are obtained from heterogeneous sources: https://doi.org/10.1016/j.cels.2020.09.007; https://doi.org/10.1039/D4DD00225C; https://doi.org/10.1021/acscentsci.2c01123; https://doi.org/10.1073/pnas.2214357120. Particularly, in the regime of AI assisted electrolyte

development, works by Cui *et al.* (https://doi.org/10.1073/pnas.221435712) and Sharma *et al.* (https://doi.org/10.1021/acs.jcim.3c01030), used Coulombic efficiency (CE) data obtained from a wide range of literature in order to build dataset large enough to train forward design model. While the heterogeneity of literature dataset (different testing conditions, different labs, different cell type, different current density, different capacity, different pressure, etc.) would diminish accuracy of ML model, their approach has been successful in developing novel electrolytes. In our work, we did curate anode-free data from the literature (Table S1). However, we recognized the heterogeneity in literature data and decided to focus on in-house data. While our in-house dataset is relatively small (which is overcome by using active learning approach), they were tested in the same lab under the same coin cell setup and same cycling protocol to ensure consistency. Hence, we agree with the reviewer that the dataset should be well-arranged and organized. That is the approach we took with our data selection.

3: Details of solvent structure used in the electrolyte

"Regarding the use of nomenclature such as "nonfluorinated ether A," their structures were withheld because they are involved in other unpublished work from our group. We have now updated Table S2 to disclose the names of all such solvents."

Without showing the actual structure of the molecule, the readers cannot follow the details of this study. The reviewer recommends that this manuscript should be resubmitted after their unpublished work is published.

Response: We have updated Table S2 with complete disclosure of all solvent compounds in the initial dataset. The SMILES structures for all the compounds are now shown.

Table S2 is attached below (References in Table S2 can be found in supporting information file):

***Table S2**. In-house anode free cell cycling data set*

Solvent	**Salt**	**Cell type** [a]	**Discharge capacity at 20th cycle (mAh/g)** [b]
1,2-Dimethoxyethane (SMILES: COCCOC)	1 M LiFSA	Cu/LFP	41.49
iPrME[29] (SMILES: CC(C)OCCOC)	1 M LiFSA	Cu/LFP	99.22
iBME[29] (SMILES: CC(C)COCCOC)	1 M LiFSA	Cu/LFP	26.76
iBME[29] (SMILES:CC(C)COCCOC)	1 M LiFSA	Cu/LFP	32.54
tBEE[29] (SMILES: CC(C)(C)OCCOCC)	2.2 M LiFSA	Cu/LFP	3.93
iPrEE[29] (SMILES: CC(C)OCCOCC)	2 M LiFSA	Cu/LFP	93.07
iPrEE[29] (SMILES: CC(C)OCCOCC)	2 M LiFSA	Cu/LFP	96.56
nBME[29] (SMILES: CCCCOCCOC)	1 M LiFSA	Cu/LFP	42.75
cPME[29] (SMILES: C1CCCC1OCCOC)	1 M LiFSA	Cu/LFP	69.89
iPrBE[29] (SMILES: CC(C)OCCOCCCC)	2 M LiFSA	Cu/LFP	84.9
Diglyme (SMILES: COCCOCCOC)	1 M LiFSA	Cu/LFP	16.80
E3F1[30] (SMILES: FC(F)(F)COCCOCCOCC(F)(F)F)	1 M LiFSA	Cu/LFP	55.04
E3CH2F[31] (SMILES: FCCOCCOCCOCCF)	1 M LiFSA	Cu/LFP	1.40
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	97.92
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	115.58
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	105.13
E4F1[30] (SMILES: FC(F)(F)COCCOCCOCCOCC(F)(F)F)	2 M LiFSA	Cu/LFP	71.67
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	112.76
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	102.73
E2CH2F[31] (SMILES: FCCOCCOCCF)	1 M LiFSA	Cu/LFP	36.525

1,1,2,2-Tetrafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(F)(F)C(F)F)	1 M LiFSA	Cu/LFP	122.52
1,1,2,2-Tetrafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(F)(F)C(F)F)	1 M LiFSA	Cu/LFP	103.74
FDMB[32] (SMILES: COCC(F)(F)C(F)(F)COC)	1 M LiFSA	Cu/LFP	104.91
FDMB[32] (SMILES: COCC(F)(F)C(F)(F)COC)	1 M LiFSA	Cu/LFP	101.62
FDMB[32] (SMILES: COCC(F)(F)C(F)(F)COC)	1 M LiFSA	Cu/LFP	88.11
FDMB[32] (SMILES: COCC(F)(F)C(F)(F)COC)	1 M LiFSA	Cu/LFP	85.60
TMEB[33] (SMILES: B(OCCOC)(OCCOC)OCCOC)	1 M LiFSA	Cu/LFP	0.112
Methyl 2,2,2-trifluoroethyl carbonate (FEMC, SMILES: COC(=O)OCC(F)(F)F)	1 M LiFSA	Cu/LFP	7.617
Trifluoro propylene carbonate (TFPC, SMILE: C1OC(=O)OC1C(F)(F)F)	1 M LiFSA	Cu/LFP	0.744
TFEB[33] (SMILES: B(OCCF)(OCCF)OCCF)	1 M LiFSA	Cu/LFP	58.13
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/LFP	116.46
2-(2-Methoxyethoxymethyl) furan (SMILES: COCCOCC1=CC=CO1)	1 M LiFSA	Cu/LFP	0.031
t-Butyl 2-methoxyethyl ether (SMILES: CC(C)(C)OCCOC)	1 M LiFSA	Cu/LFP [c]	41.42
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/NMC811	149.42
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA, 0.2 M LiDFOB	Cu/NMC811	0.11
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA	Cu/NMC811	150.84
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M LiFSA, 0.2 M $LiBF_4$	Cu/NMC811	19.80
1,1,1,2,2-Pentafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(C(F)(F)F)(F)F)	1 M $LiBF_4$	Cu/NMC811	0.002
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	1 M LiFSA	Cu/NMC811	127.22

1,1,2,2-Tetrafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(F)(F)C(F)F)	1 M LiFSA	Cu/NMC811	124.24
1,1,2,2-Tetrafluoro-3-(2-methoxyethoxy)propane (SMILES: COCCOCC(F)(F)C(F)F)	1 M LiFSA	Cu/NMC811	130.43
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	1 M LiFSA	Cu/NMC111	90.206
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	1 M LiFSA	Cu/NMC111	90.21
1,1,1,2,2,3,3-heptafluoro-4-(2-methoxyethoxy)butane (SMILES: C(COCCOC)(C(F)(F)C(F)(F)F)(F)F)	2 M LiFSA	Cu/NMC111	99.47
1,2-Diethoxyethane (SMILES: CCOCCOCC)	1 M LiFSA	Cu/LFP	115.52
1,2-Dimethoxyethane (SMILES: COCCOC)	1 M LiFSA	Cu/LFP	55.63
iPrME[29] (SMILES: CC(C)OCCOC)	2 M LiFSA	Cu/LFP	109.53
iPrME[29] (SMILES: CC(C)OCCOC)	2 M LiFSA	Cu/LFP	111.86
1,2-Dimethoxyethane (SMILES: COCCOC)	2 M LiFSA	Cu/LFP	102.06
1,2-Dimethoxyethane (SMILES: COCCOC)	2 M LiFSA	Cu/LFP	106.45
1,2-Dimethoxyethane (SMILES: COCCOC)	2 M LiFSA	Cu/LFP	95.3
iPrEE[29] (SMILES: CC(C)OCCOCC)	2 M LiFSA	Cu/LFP	40.21
EESF[2] (SMILES: CCOCCS(=O)(=O)F)	1 M LiFSA	Cu/LFP	3.55
EESF[2] (SMILES: CCOCCS(=O)(=O)F)	2 M LiFSA	Cu/LFP	11.63
ESF[2] (SMILES: CCS(=O)(=O)F)	1 M LiFSA	Cu/LFP	74.23
ESF[2] (SMILES: CCS(=O)(=O)F)	2 M LiFSA	Cu/LFP	114.49
ESF[2] (SMILES: CCS(=O)(=O)F)	4 M LiFSA	Cu/NMC811	145.87
1,2-Dimethoxyethane (SMILES: COCCOC)	4 M LiFSA	Cu/NMC811	154.52

Reviewer #2 (Remarks to the Author):

The authors addressed my comments. I would recommend the acceptance.

Response: We thank the reviewer for recommending acceptance of our manuscript.

Reviewer #3 (Remarks to the Author):

I appreciate the authors' thorough and thoughtful responses to my previous comments and concerns. The revised manuscript addresses all major points I raised. The only remaining revision I suggest is a final proofread to correct minor typographical and grammatical errors. With these updates, I find the manuscript suitable for publication and recommend acceptance after minor editorial revision.

Response: We thank the reviewer for finding our manuscript suitable for publication. We have proofread the manuscript to remove any typographical and grammatical errors.

Reviewer #3 (Remarks on code availability):

This is a valuable addition that significantly enhances the transparency and reproducibility of the work.

Response: We thank the reviewer for recognizing our efforts of enhancing transparency.

Reviewer #1 (Remarks to the Author):

Thank you for your response to my previous comments. The manuscript was properly revised, and my previous concerns related to AI parts are mostly solved. Here, let me comment on the experimental analytical part. In particular, latter part of the manuscript, during figure 3 to 5. The details are described as follows.

Response: We thank the reviewer for the feedback regarding the AI part and the new comments regarding experimental part. We have addressed them in the revised drafts which appear as point-by-point responses below.

1:
In the experiments of figure 3, the electrolyte of 1 M LiFSA in F5DEE was utilized as control. However, this electrolyte composition is not so common in the LMB community. Instead, 4M LiFSI in DME (1,2-Dimethoxyethane) is widely utilized as standard electrolyte for lithium metal electrode investigation. In addition, this study also uses this electrolyte for some cells for in-house anode-free cells. Thus, the reviewer suggests that the experiments for cycling test with the cell use of 4M LiFSI in DME also should be conducted. This result is very helpful for the readers in this community as a control experiment.

Response: We thank the reviewer for suggesting 4 M LiFSA in DME as an additional control electrolyte. We investigated the performance of 4 M LiFSA in DME in standard Cu/LFP cell cycling test. As shown in the attached figure below, 4 M LiFSA in DME electrolyte was found to have slightly worse capacity retention in comparison to F5DEE or the electrolytes discovered in this work. In fact, 4 M LiFSA in DME has been previously studied as electrolyte for Cu/LFP anode free cell (https://doi.org/10.1002/adfm.201602353), and the reported performance is not as good as F5DEE (Table S1). Figure 3 in the manuscript has been updated to include 4 M LiFSA in DME as an additional control.

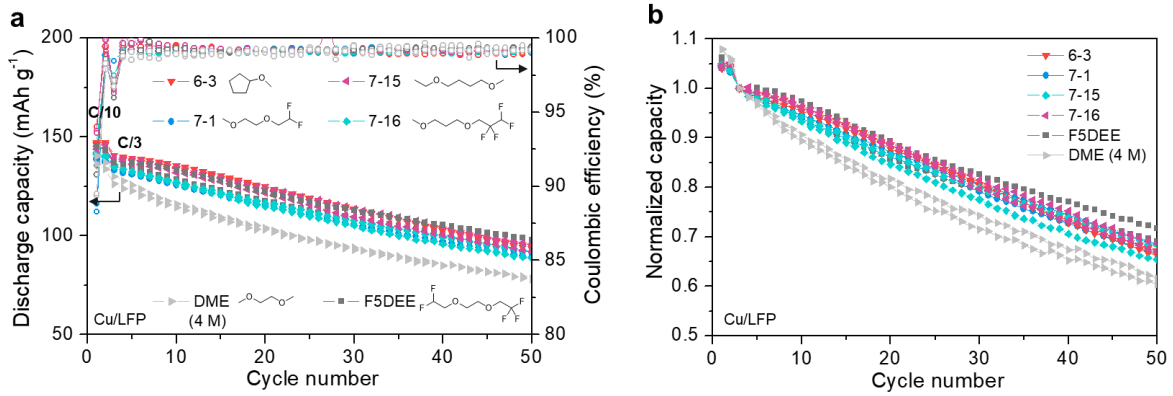


Figure 3 | Performance of AL-discovered electrolytes. a) *Cycling performance of Cu/LFP cells using AL discovered electrolytes with 1 M LiFSA in F5DEE and 4 M LiFSA in DME as controls.*

1 C ≈ 2 mAh/cm^2. ***b)*** *Normalized discharge capacity of Cu/LFP cells using AL discovered electrolytes with 1 M LiFSA in F5DEE and 4 M LiFSA in DME as controls. Two replicated cells are shown for each electrolyte.*

2:

The manuscript showed the analytical study for the four kinds of discovered electrolyte, by series of standard experiments for Li metal study, including, SEM, Raman, Li potential shift, measurement of conductivity and diffusivity, XPS. These studies provide detailed information that is necessary for answering the question, what is the origin of high-capacity retention of the discovered electrolyte. To deepen the discussion, the reviewer suggests that the data obtained from the cell with control electrolyte (1 M LiFSA in F5DEE and 4M LiFSI in DME) also shown and the detailed discussion for comparing the results is encouraged. Such data and discussion are very meaningful for readers and the community of LMB to get deep insight into the discovered electrolyte through the present study.

Response: We appreciate the reviewer's suggestion regarding showing the standard Li metal study data for control electrolytes. The corresponding data and discussions have been added to the manuscript or supporting information as listed below.

1) SEM images of lithium deposited from F5DEE and DME (4 M) control has been added as Figures 3g and 3h:

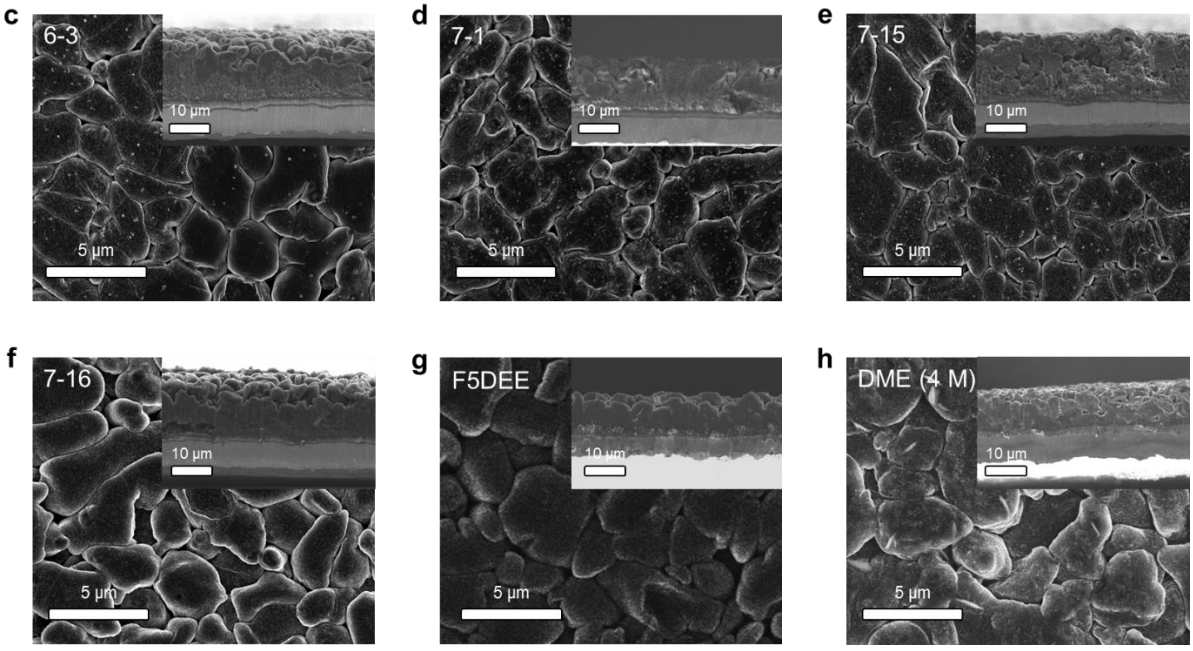


Figure 3 \| Performance of AL-discovered electrolytes. *... **c-f)** SEM images of lithium metal deposited in (**c**) 1 M LiFSA in 6-3, (**d**) 1 M LiFSA in 7-1, (**e**) 1 M LiFSA in 7-15, (**f**) 1 M LiFSA in 7-16, (**g**) 1 M LiFSA in F5DEE and (**h**) 4 M LiFSA in DME using Li/Cu cells with a current*

density of 1 mA/cm^2 to a capacity of 1.5 mAh/cm^2. The inset shows corresponding cross section SEM images of lithium metal sample.

Corresponding discussions:

"Fig. 3c-h show SEM images of lithium metal deposited in Li/Cu cell at 1 mA/cm^2 to 1.5 mAh/cm^2, where all four AL-discovered electrolytes lead to similar morphology with micron scale granules in compact and uniform stacking. Such morphology is similar to the F5DEE and DME (4 M) control electrolytes and has been widely reported as evidence of good lithium metal compatibility and high lithium cycling efficiency in literature."

2) The solvation structure of AL-discovered and F5DEE control electrolytes were originally compared in Fig. S16. The data for DME (4 M) control has been added to this figure and corresponding discussions have been modified in the manuscript.

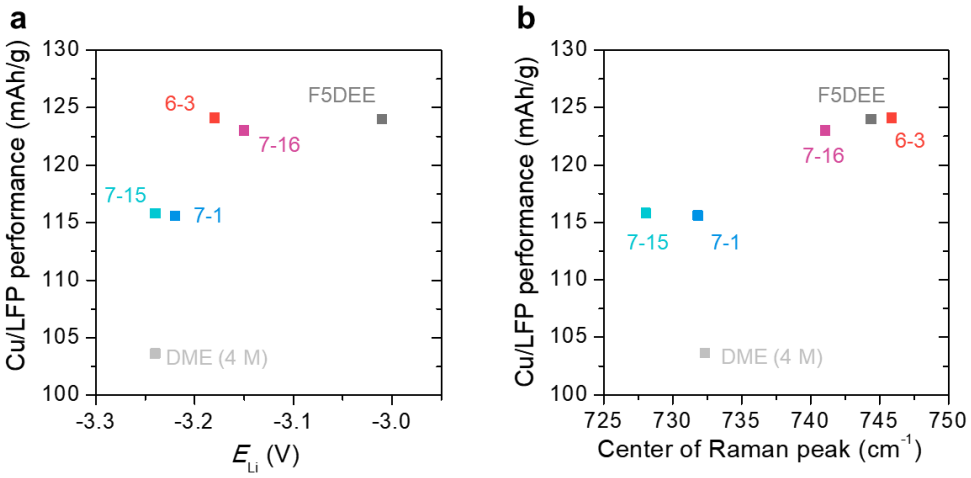


***Figure S16**. Correlation of ion solvation structure to anode free LMB cycling performance. **a)** Discharge capacity at 20th cycle as a function of lithium redox potential (E_{Li}). **b)** Discharge capacity at 20th cycle as a function of gravity center of FSA Raman peak. Ion solvation structure does not show clear correlation among the best performing electrolytes especially for 6-3, 7-16 and F5DEE.*

Corresponding discussions:

"As summarized in Fig. S16, the degree of ion pairing in 6-3 and 7-16 is comparable to F5DEE while 7-1, 7-15 and DME (4 M) electrolytes show relatively less ion pairing. However, the correlation between solvation structure descriptors and anode-free LMB performance remains unclear among the electrolytes under investigation. This is likely due to the subtle performance differences and suggests that increasing ion pairing beyond a certain threshold does not necessarily lead to better lithium metal cycling performance.[13,36,55]"

3) The ion transport properties of AL-discovered electrolytes have been compared to control electrolytes and discussed in the manuscript:

"At 20 °C, 7-1 electrolyte has the highest ionic conductivity of 3.5 mS/cm and the other three electrolytes show acceptable conductivities of 0.7~1.0 mS/cm. While the room temperature conductivity is lower than those reported for control electrolytes (5.0 mS/cm for 1.2 M LiFSA in F5DEE[13] and 5.7 mS/cm for 4 M LiFSA in DME[56]), some of AL-discovered electrolytes do show appreciable rate performance as discussed earlier (Fig. S15).

... ...

While the AL-discovered electrolytes show lower ionic conductivity, the lithium ion diffusivity (D_{Li}) of 6-3, 7-1 and 7-15 is higher than F5DEE ($D_{Li}=9.28\times10^{-11}$ m^2/s) or DME (4 M) ($D_{Li}=3.96\times10^{-11}$ m^2/s) controls, which indicates ionic conductivity is also affected by concentration of free charge carriers. All four AL-discovered electrolytes show high lithium transference number ranging from 0.47 to 0.53, which is similar to the control electrolytes (0.49 for F5DEE and 0.47 for DME (4 M)) and agrees with the solvation structure rich in ion pairing as discussed earlier."

4) The XPS data of lithium metal SEI in F5DEE and DME (4 M) control electrolytes have been added to Figure 5 and Figure S17. The updated figures and discussions are attached below:

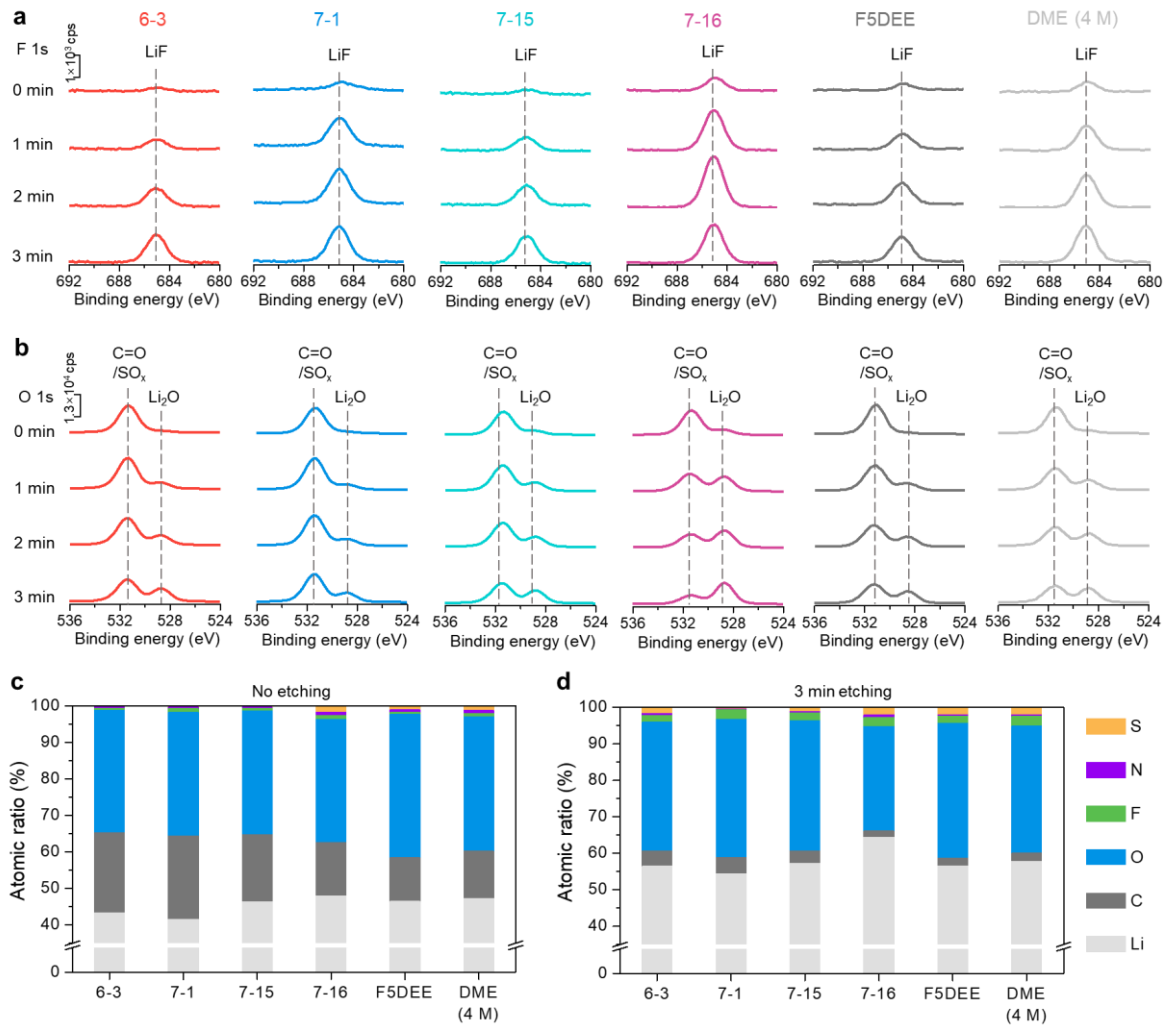


Figure 5 | SEI composition. ***a)*** *F 1s XPS spectra of lithium metal SEI at different etching time.* ***b)*** *O 1s XPS spectra of lithium metal SEI at different etching time.* ***c-d)*** *Atomic ratio quantified from XPS with* ***(c)*** *no etching and* ***(d)*** *3 minutes of etching.*

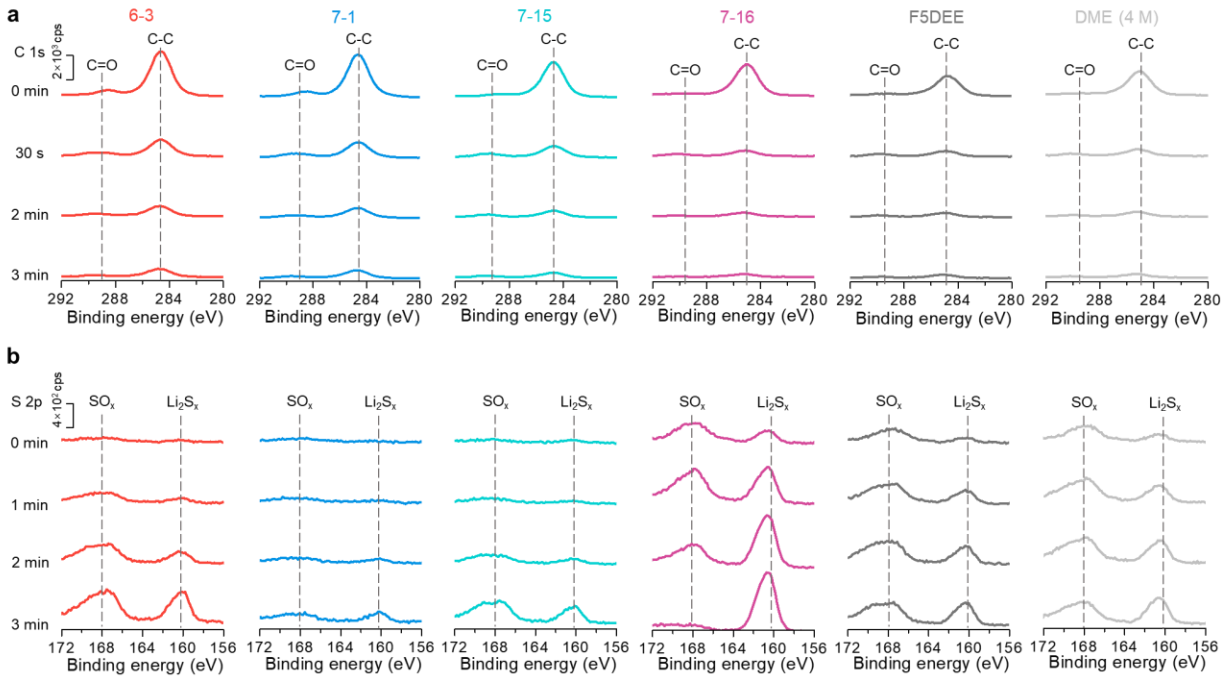


Figure S17*. XPS spectra of lithium metal SEI at different etching time.* ***a****) C 1s spectra.* ***b****) S 2p spectra.*

Corresponding discussions:

"Compared to F5DEE and DME (4 M) control electrolytes, AL-discovered electrolytes show similar SEI compositions and depth profiles. Such SEI structure resembles the anion derived inorganic rich SEI widely reported for electrolytes with superior lithium metal compatibility.[13,51,57]*"*