1. INTRODUCTION

Out-of-equilibrium interfacial reactions which result in functional, transport-selective surface films play a crucial role in materials-based technologies, enabling applications from batteries, capacitors, water splitting, protective coatings, and alloy design. An exemplar in the family of spontaneously formed, enabling interfaces is the passivation layer which results from the reductive decomposition of electrolytes on the anode of lithium-ion batteries (LIBs) during the first few charging cycles. This solid electrolyte interphase (SEI) is considered "the most important but the least understood in rechargeable Li-ion batteries," because of the complex relationships between battery conditions, SEI composition/morphology, and SEI properties. It is generally believed that an ideal SEI should be selectively permeable to the electrochemically active charge carrier (i.e., Li⁺) but also electron-insulating to prevent continuous reduction of the electrolyte species.

Over the last decades, careful bulk and surface characterization techniques, such as X-ray photoelectron spectroscopy (XPS), Fourier transform infrared (FTIR) spectroscopy, nuclear magnetic resonance spectroscopy (NMR), cryo-scanning transmission electron microscopy (cryo-STEM), and so on, have been employed to elucidate the chemical composition of the SEI on both Li metal and graphite. Critical understanding has been gained as to early SEI formation as well as its aging. However, there are still numerous uncertainties as to specific reaction mechanisms, influence of the electrode chemistry and potential on competing reactions, and even the characterization of fundamental SEI components such as, for example, lithium ethylene dicarbonate (LEDC). LEDC was first proposed to be the major organic constituent of the SEI from the EC electrolyte by Aurbach et al. and later investigated extensively by numerous studies including synthetic efforts of the LEDC standards and comparisons with...
those synthetic standards based on FTIR, XPS, and NMR techniques.\textsuperscript{11,20–26} Recent work by Wang et al.\textsuperscript{27} cast doubt on the LEDC characterization: it was suggested that common LEDC standards\textsuperscript{21,23,25} are composed of lithium ethylene monocarbonate (LEMC). The authors further suggested that LEMC degrades to ethylene glycol (EG) under D$_2$O extraction, thus rendering the $^1$HNMR spectra for LEDC and EG indistinguishable. However, while characterization of LEDC is under debate, questions also remain as to how LEMC forms under typical battery operation conditions and if it truly replaces LEDC as a major early SEI component.

Computational modeling of SEI-relevant reactions has also provided crucial insights into specific reaction mechanisms, including short-lived intermediate products.\textsuperscript{28,29} Prominent examples include the early pioneering work from Balbuena and co-workers based on static DFT calculations to clarify the reduction mechanisms of common organic solvents and additives,\textsuperscript{30--32} surface calculations,\textsuperscript{33--35} \textit{ab initio} molecular dynamics (AIMD) simulations,\textsuperscript{36,37} and others. However, currently employed methodologies such as static DFT, \textit{ab initio}, and classical molecular dynamics are inherently limited either by chemical intuition, simulation time scales, and/or sampling. In particular, more complex, rare, co-operative, and combined reaction pathways as well as the competition between different reaction mechanisms are difficult to capture. However, progress has been made in the automated searching for reaction pathways, traversing the potential energy surface (PES) using combinations of weighted sampling, heuristic rules, and graph theory.\textsuperscript{38--57} Recently, a new data-driven methodology was developed to analyze complex electrochemical reactive systems with minimal chemical intuitive bias through the use of computational reaction networks.\textsuperscript{58} Using this methodology, pathways from Li$^+$ and EC to LEDC under reductive conditions corresponding to Li metal were explored. It was demonstrated that the reaction network automatically recovers previously suggested reaction mechanisms for the formation of LEDC as well as proposed novel, thermodynamically feasible mechanisms that had not previously been reported. Here, we utilize the same framework to identify possible reaction pathways leading to LEMC based on automated pathfinding in a landscape built on \textit{ab initio} thermochemistry. We compare the proposed pathways on the basis of their thermodynamics, kinetics, and byproducts. We note that thermodynamic analysis alone is inadequate for assessing the energetic landscape, as kinetics also play a pivotal role in the SEI formation.\textsuperscript{36,59} In the present study, we also conduct static transition state (TS) calculations for the promising paths from the path-finding algorithm using a total thermodynamic cost cutoff, so as to refine the reaction network predicted paths and assess their kinetic feasibility. The purpose of our work is 2-fold:

(1) We implement and demonstrate a systematic and minimally biased methodology for the generation of relevant species and reactions. Subsequently, reaction pathways to a certain product of interest can be obtained in an automated fashion through pathfinding from a reaction network containing these species and reactions. All relevant molecules were generated through a systematic fragmentation/recombination procedure such that all reactions that satisfy a certain constraint (involving $\leq 5$ bond changes here) were included. The effectiveness of this approach is tested on the formation of the new SEI component of interest, LEMC. We demonstrate that our data-driven methodology not only

![Figure 1. Workflow for generating relevant molecules.](https://doi.org/10.1021/jacs.1c05807)
(2) We assess the reaction network-predicted pathways both in the context of experimental findings and from a comparison with DFT-refined pathways. We note that the advantage of the automatically generated pathways lies in the rapid evaluation of thermodynamically plausible reactions covering a vast electrochemical space and provides a thermochemically informed starting point for more bespoke computational and experimental explorations.

2. METHODS

2.1. Generation of Candidate Molecules and Molecular Intermediates through Fragmentation and Recombination.

To generate the constituents of the reaction network, we employ heuristic rules based on thermodynamic data and associated molecule-learned bond dissociation and formation, which we denote as the fragmentation and recombination procedure. This approach enables the generation of a large number of candidate molecular structures that represent the out-of-equilibrium, transient landscape of possible molecular fragments that form under electrochemical electrolyte decomposition. These candidate structures are subsequently optimized by electronic-structure methods to obtain minimum-energy structures. Molecular species are treated as graphs based on the connectivity of atoms. Fragmentation is defined as breaking one edge in a molecular graph and generating one or two molecular fragments; recombination is defined as forming one edge between two specified atoms on two molecular graphs and forming one recombinant molecular graph. In a one-step fragmentation, fragment molecular graphs are generated from breaking any one bond in the molecule. An n-step fragmentation involves multiple sequential one-step fragmentation processes; at each stage, the fragments from the previous stage undergo a one-step fragmentation, and the fragments from each such one-step procedure are collected. Similar to the fragmentation procedure, a general one-step recombination is conducted by iterating over all the molecular pairs in a given set of species, iterating over all atoms in both molecules, and generating all recombinant molecules by connecting one bond. During the process, a graph isomorphism comparison is applied at each step to check for structural duplicates, which are then removed. To compare and contrast LEDC with LEMC formation pathways as a result of electrochemical dissociation energies (BDEs; the recombination free energy would be the negative number of the corresponding BDE). The higher MAE than that reported in the BondNet paper is likely to originate from (1) examples far away from the regime of the training set, for example, forming a bond between sterically hindered oxygens and (or) carbons; and (2) the larger conformational space of large flexible recombinant molecules, as currently BonDNet is not taking the energies of different conformers into consideration. The performance of BonDNet can be iteratively improved by adding more representative data into the training set: by adding these recombinants back into the training, we observed a reduction of MAE from 0.496 to 0.355 eV when testing on a new test set consisting of a subset of the recombinants we previously eliminated (details given in SI Sec VIII). Out of 382 newly calculated recombinants, we only excluded 7 molecules (less than 2% error) due to the BonDNet prediction error, which represents a low percentage loss and is unlikely to change our conclusions significantly. Detailed error analysis on both models is given in SI Sec VIII.

(4) The geometries of the resulting molecules were optimized in DFT. The initial guesses for the 3D geometries of the recombinant molecules were generated on the basis of the 2D molecular connectivities from the OPLS_2005 molecular force field as implemented in the Schrödinger Python Suite. To reduce the computational burden of generating the recombinant molecules, no conformational sampling was conducted. During geometry optimization, a portion of the recombinant molecules changed their connectivities; those molecules were removed from consideration as the results of BonDNet were no longer applicable, and they were no longer in the regime of one-step recombination.

We obtained 570 candidate molecules in total through this procedure (45 from fragmentation and 525 from recombination). We note that the generation of the network species is self-contained, such that all fragments and recombinant molecules were generated from the procedures (1)–(4). Specifically we emphasize that no molecules were added from previous knowledge of Li electrolyte decomposition and that the target molecules LEDC and LEMC were obtained in this automated fashion. All the molecules used in this work are given in the SI; a more comprehensive data set which contains, for example, fragments for various salts and other electrolytes, is released in another work.

2.2. Construction of Reaction Network and Identification of Reaction Pathways.

The reaction network was constructed as a connected graph using the same graph representation as described in previous work. Allowing a broad range of reactions between the molecular species, we chose to include “concerted” reactions (reactions that involve multiple, simultaneous bonds breaking or forming) with up to 5 bond changes. Similar constraints (usually ≤2 bonds breaking and ≤2 bonds forming) were imposed in other works. We emphasize that metal coordination reactions, such as Li bonding to a heteroatom, is treated in the same manner regardless of their inductive nature, which motivates an inclusive, upper limit of allowed concerted bond changes. We note, however, the possibility that reactions considered here as concerted, may in reality proceed through multiple elementary steps. Hence, we consider reaction pathways as reasonable guesses for a collection of events that should be refined using quantum chemical calculations.

To efficiently include concerted reactions, a mixed-integer linear programming (MILP) scheme modified from a previous work for atom-mapping was utilized to calculate the minimum chemical energy of allowed concerted reactions. The BonDNet predicted vs DFT computed bond dissociation energies (BDEs; the recombination free energy would be the negative number of the corresponding BDE). The higher MAE than that reported in the BondNet paper is likely to originate from (1) examples far away from the regime of the training set, for example, forming a bond between sterically hindered oxygens and (or) carbons; and (2) the larger conformational space of large flexible recombinant molecules, as currently BonDNet is not taking the energies of different conformers into consideration. The performance of BonDNet can be iteratively improved by adding more representative data into the training set: by adding these recombinants back into the training, we observed a reduction of MAE from 0.496 to 0.355 eV when testing on a new test set consisting of a subset of the recombinants we previously eliminated (details given in SI Sec VIII). Out of 382 newly calculated recombinants, we only excluded 7 molecules (less than 2% error) due to the BonDNet prediction error, which represents a low percentage loss and is unlikely to change our conclusions significantly. Detailed error analysis on both models is given in SI Sec VIII.
distance (CD), that is, the minimum number of bond changes needed to transform reactants to products, for all possible combinations of (up to two) reactants and products (for details, see SI Sec I). All molecular pairs that satisfy a given stoichiometry were subsequently evaluated on the basis of the minimum number of bond changes required to convert one molecule pair to another. This procedure resulted in a collection of all reactions that can be reached within 5 bond changes, totalizing 570 species and 8973154 reactions.

To calculate the free energy for redox reactions relevant in Li-ion batteries, we employed a potential of $U = 0$ vs Li/Li$. Hence, to represent the reaction free energy for an arbitrary redox reaction, we employ the full-cell reaction $A + Li^+ \rightarrow A^- + Li$. The absolute electrode potential for the lithium electrode is calculated to be $\sim 1.40$ V by adding the Li/Li$^+$ potential relative to standard hydrogen electrode (SHE) to the absolute electrode potential of hydrogen electrode (4.44 V).

We note that this is equivalent to a chemical potential of $-1.40$ eV for a Li metal electron. The effect of the electrode potential (or any applied external potential) can be easily accounted for, by shifting the free energy of the electron by $-E_U$, where $U$ is the potential relative to Li/Li$^+$. For instance, the graphite anode in LiBs operates at a potential close to $U = 0.2$ V vs Li/Li$^+$ which is equivalent to an electron chemical potential of $-1.60$ eV.

To identify promising thermodynamic pathways from the nearly 9 million hypothetical reactions, we adopted the same procedure for pathfinding as described in the previous work. This approach provides us with a ranked list of reaction sequences, and each reaction sequence connects the starting molecules to the target molecule (LEMC), within a grand canonical thermodynamic ensemble, allowing for a variable number of each specie. A softmax function, that is, $ln(1 + \exp(x))$, of reaction free energy is used to determine the cost for traversing a reaction in the reaction network based on computational efficiency and the tolerance for slightly endergonic reactions. However, we note that the resulting ranking from the cost function is not used to exclude pathways from consideration. Indeed, after obtaining the representative reaction pathways from the reaction network, we subsequently subject the proposed pathways to a detailed quantum mechanical study to obtain the transition state barriers ($\Delta G^*$) as well as the reaction energies of the complexes ($\Delta G$) to verify their kinetic/thermodynamic feasibility. This information is intended to refine the pathways obtained from the reaction network which currently does not include any explicit kinetic information. Furthermore, concerted mechanisms are refined in this more thorough exploration.

### 2.3. Computational Methods.
DFT calculations were performed using the QChem$^99$ electronic structure code. The geometry optimizations for the species included in the reaction network used the $\omega$B97X-V$^{100}$ functional with the def2-TZVPDD basis set.$^{72}$ Solvent effects were taken into account by SMD implicit solvation models$^{73}$ with user-defined parameters representing the Li-ion electrolyte. Specifically, an experimentally measured dielectric constant (for 3:7 EC/EMC binary mixture$^{64}$) of 18.5 was used; for all other parameters, we used the values for pure EC.$^{72}$ For all stationary point geometries, frequency calculations were conducted to ensure they are true local minima on the PES for minimum stuctures or first-order saddle points for TS structures. Gibbs free energies based on harmonic approximations were used throughout to determine the energy landscape in the reaction pathways.

Geometry optimization and reaction path analysis for the predominant paths were performed at the $\omega$B97X-D$^{77}$/def2-SVPD$^3$ level of theory with the C-PCM implicit solvation model$^{101}$ (dielectric constant 18.5, optical dielectric 2.0, which is consistent with the SMD parameters), while single-point energies at the stationary points were obtained using the $\omega$B97X-V/def2-TZVPPD/$\omega$SMD level of theory using the same SMD parameters as above. These two levels of theory have been shown to perform well for metal organic systems.$^{78,79}$ We used the lower level of theory for computational efficiency in exploring the PES; in our experience, these two levels of theories give good agreements in both molecular geometries and relative energies. Vibrational frequencies were computed using the smaller basis, ensuring they are true local minima on the PES for mininum stuctures and frequency calculations were conducted to confirm the TSs as minima on the PES.

### 3. RESULTS AND DISCUSSION
Herein we present and analyze the formation pathways to a recently identified, potential key component of the Li-ion SEI, LEMC.$^{27}$ Analyzing the top-ranked 30000 pathways, first of all, we found that many of them were in fact the same, except for a difference in the order of reactions. These “duplicate” paths exhibit the same cost, and eliminating the paths with identical costs to threshold $10^{-3}$ resulted in 434 unique paths (the unique paths are included in the SI). We further grouped the paths roughly based on the main molecules involved in the paths and report the representative paths below. While this elimination may exclude subtle mechanistic differences in the paths, we believe that the main features are preserved.

Analyzing the participant molecules in the reaction pathways leading to LEMC formation, interestingly, we found that only 8 paths (among the 434 unique ones) proceeded without water. Furthermore, DFT refinement of these paths suggests none of these pathways are kinetically feasible, with the rate-limiting barriers for the key reactions in the paths exceeding 2 eV (details shown in SI Sec V). This important result indicates that water is likely to be critically important in the formation of LEMC. We note that the presence of kinetically infeasible paths is inevitable: this is a direct consequence of including a broader scope of reaction possibilities. Indeed, the identified infeasible mechanisms were found to include overcomplicated concerted mechanisms connecting two unstable open-shell species (e.g., carbene). Since this is an artifact that may result in missing important nonwater mechanisms, we constructed another reaction network with more limited concerted reactions, and the details will be given in Sec 3.3 below. An overview of the two reaction networks considered in this work is given in Figure 2.

The two major pathways that have been proposed in the literature, namely, the EC hydrolysis and LEDC hydrolysis reactions, were successfully identified by the reaction network, validating our data-driven approach. Additionally, other possibly interesting pathways were discovered and will be
discussed below, such as a pathway that generates lithium hydride; pathways involving new, recombinant compounds that have not been conventionally considered to be important in the SEI formation; and a pathway involving the breakage of the (CH₂)O−C(=O)OLi bond in LEDC.

In the following sections, we divide the reaction pathways to LEMC into three categories: (i) direct paths from EC, Li⁺, and H₂O, where only the fragments of LiEC serve as intermediates in the paths; (ii) paths involving recombinant molecules; and (iii) paths without water participation. We present, for each category, both the representative paths predicted from the reaction network and the DFT validated/redefined free energy diagrams, with some discussion about the similarities/differences between them. We also discuss the physical implications of those paths in the context of experimental findings.

3.1. Direct Reaction Paths to LEMC from EC, Li⁺, and H₂O. Figure 3 shows a representative set of “direct” paths leading to the formation of LEMC, utilizing EC, Li⁺, and H₂O as the starting molecules. Although differing in the detailed steps, the two major mechanisms identified were (i) LiOH or OH⁻ attacking an EC or LiEC⁺ to form LEMC (path A, B, C, E); and (ii) LiEC, water, and a second electron react to form LEMC and hydride (path D). Figure 3 presents the energy profiles from DFT (ignoring explicit solvation and surface effects) corresponding to these two mechanisms. We note that the generation of LiOH and or OH⁻ species can be corroborated by the known instability of water against Li metal,81 which is also captured within our level of theory, when an explicit surface is included (see SI Sec III). Since the reaction network does not have any knowledge of water splitting on the Li surface, it used indirect mechanisms for generating LiOH and or OH⁻.

The reaction of LiOH attacking EC to form LEMC was proposed by Wang et al. and also experimentally confirmed by adding anhydrous LiOH to an EC/LiPF₆ solution.27 In the EC + LiOH → LEMC reaction, we identified two mechanisms through DFT (purple and black paths in Figure 4a): one is a single-step process where LiOH directly attacks the ethylene carbon in the EC; another one is a stepwise process where LiOH attacks the carbonyl carbon to form a tetrahedral intermediate, and a subsequent shoulder bond breaking and proton transfer creates LEMC. The former one-step process has a high barrier of 2.24 eV (Struct 3 → TS 5; 1.95 eV with an explicit EC molecule, as shown in Figure S6, Struct 1 → TS 3), whereas the stepwise mechanism exhibits an effective overall barrier of 1.33 eV (Struct 3 → TS 4; 0.95 eV with an explicit EC molecule, as shown in Figure S6, Struct 1 → TS 2). The reaction barrier can be significantly lowered if the hydroxide ion is free and not bound to lithium (0.25 eV, Struct 7 → TS 6; 0.27 eV with an explicit EC molecule, as shown in Figure S6, Struct 6 → TS 4). This is consistent with the fact that Li⁺ is a strong Lewis acid, and the nucleophilicity of bare hydroxide anions is expected to be greatly reduced in the presence of Li⁺.82,83 Similarly, the electrophilicity of EC would be enhanced by Li⁺, thus also increasing the reactivity between the two species. We consider OH⁻ + LiEC⁺ → LEMC a more realistic scenario for EC hydrolysis, as (i) LiOH has extremely high

Figure 2. Overview of the reaction networks generated in this work.

Figure 3. Direct paths to LEMC from EC, Li⁺, and H₂O identified by the reaction network.
low solubility in organic solvents; (ii) OH\(^{-}\) should be expected to be present because the electron-abundant anode represents a highly basic environment,\(^{84}\) and (iii) LiEC\(^{+}\) is considered to be abundant in battery electrolytes, and the Li\(^{+}\)-coordinated EC is reduced at a higher potential than the pure solvent.\(^{30}\)

Path D in Figure 3 as proposed from the reaction network represents another possibility of LEMC formation. In path D, a neutral LiEC is further reduced to LiEC\(^{-}\) and reacts with water to form LEMC and a hydride ion. (c—EC\(^{2-}\))Li\(^{+}\), where c—EC\(^{2-}\) represents the cyclic EC dianion, was proposed previously by Leung et al.,\(^{36}\) and the reduction rate of c—EC\(^{-}\) was estimated on the basis of Marcus theory to be significantly higher (at \(\sim 10^4\) order) than that of EC but lower than that of \(\sigma-EC\(^{-}\), which is the ring-opened EC anion. Leung et al. also suggested that the doubly reduced (c—EC\(^{2-}\))Li\(^{+}\) favorably ring opens at the shoulder bond and release CO. Regardless of the existence of c—EC\(^{2-}\) (or its lifetime), the reaction network suggests a reaction path according to EC + Li\(^{+}\) + 2e\(^{-}\) + H\(_2\)O \(\rightarrow\) LEMC + H\(^{-}\). This poses a possibility of a simultaneous (presumably asynchronous) formation of hydride.

Figure 4. LEMC paths from DFT.
and LEMC. We anticipate that the hydride will react rapidly with any available Li⁺ ion to form LiH, which is likely to—in the presence of water—further react to LiOH on its surface. Experimentally, LiH was first characterized by IR spectroscopy on Li metal with wet electrolyte solutions and later observed as a main dendritic product in the Li-anode-SEI by cryo-STEM mapping and XRD experiments.

Examining this hypothetical path more carefully, we obtain a mechanism that resembles the black path in Figure 4a. Note that the overall charge state for the structures in the energy profile in Figure 4a is neutral, whereas in Figure 4b, the overall charge is −1. We were unable to identify a path that proceeds through (c−EC^−)Li⁺, presumably because water is more readily reduced than c−EC (the hydrogen evolution potential occurs at 2.21 V relative to Li/Li⁺, whereas the reduction potential of c−EC^− is calculated to be 1.16 V relative to Li/Li⁺). Since the reaction network assigns a potentially overestimated endothermic penalty (depending on the distance to the fresh Li metal surface) to water reduction (see Sec 2.3), it proposes reduction of LiEC rather than H₂O. In the DFT refinement, we see a favorable water splitting in the anion state and identify a path corresponding to H−Li−OH^−+EC → LEMC + H^+ (Figure 4b). We include a water splitting step with LiEC^− at the beginning only for the sake of completing the energy diagram; however the LiEC in Struct 1 is Li⁺ EC in character, as revealed by the Mulliken atomic charge (−1.17 e) and spin (0.0 au), and the Li in the structure is considered a Li surface proxy, coordinating to EC. A similar argument can be made for the water splitting step in Figure 4a, where the Li is neutral in character (Struct 1). As mentioned in Sec 2.3, metallic surface effects are neglected in the reaction network, and to confirm that water splitting is indeed spontaneous within our level of theory, we used a cluster DFT model consisting of 32 Li atoms and found that the reaction is almost barrierless (see SI Sec III) and consistent with previous work (0.12 eV). We note that although water reduction reactions are usually considered to be equivalent to hydrogen evolution reaction (HER), this is a simplified picture as the detailed mechanisms involve adsorbed species H_ads and OH_ads. Hence, the as-formed hydride ion, most likely created through water splitting on the Li surface, is presumably stabilizing the TS structures along the path, and the overall barrier (0.72 eV as the rate-limiting barrier, Struct 4 → TS 3) is much lower than that for the black path in Figure 4a. Further investigation is necessary to determine the plausibility of this path and whether cooperative effects could further lower the barrier in this pathway.

Generally comparing our results with previous investigations of water reactivity with the Li metal surface, our results are consistent with Shang et al. who found the final products of H₂O reacting with a Li surface to include LiOH, a small amount of LiH, H₂ gas, and various water complexes.

3.2. Reaction Paths to LEMC from EC, Li⁺, and H₂O Involving Recombinant Molecules. By including recombinant molecules into the reaction network, we were able to identify some pathways to LEMC involving new intermediate molecules en route to the formation of LEMC, as shown in Figure 5. Note that the inclusion of LEMC and LEDC, along with other recombinants is a direct consequence of graph enumeration. In other words, those molecules are naturally considered irrespective of any prior knowledge of the SEI components previously studied in the literature.

The hydrolysis reaction of LEDC (path A in Figure 5) was both proposed by Wang et al. and Rinkel et al. This pathway was successfully identified by our reaction network, albeit with a low ranking (ranking 352 among the 434 unique paths), presumably owing to the endergonicity of the last LEDC + H₂O → LEMC + LiHCO₃ reaction step (0.07 eV). We also note that the LEDC and LEMC conformers used in the reaction network are linear configurations because of inadequate sampling of the conformational space. Applying a conformational search under implicit solvation identified a puckered LEDC structure (see Figure 6a), which is lower in energy than the linear conformer by 0.6 eV. While this puckered conformer obtained here may or may not be representative under explicit solvation (to our knowledge, there is no known crystal structure for LEDC), the significant energy difference highlights the impact of LEDC conformation to the reaction energetics.

Figure 5. Paths to LEMC involving recombinant molecules from EC, Li⁺, and H₂O identified by the reaction network.
Examining more closely path A and its kinetics, we found that the single-step LEDC hydrolysis step exhibits a high barrier (3.05 eV, Figure S2, Struct 1 → TS 1). However, hydrolysis reactions do not usually proceed in a single step, and therefore we attempt to obtain the barrier for a stepwise hydrolysis of LEDC passing through a tetrahedral intermediate, either with neutral water or OH\(^{-}\). The rate-limiting barrier of the water version of this path is 1.48 eV (Figure 6a, Struct 1 → Struct 4; 1.37 eV with an explicit EC molecule, as shown in Figure S7, Struct 1 → TS 3), which is lowered significantly compared with the single-step path. Under basic conditions (starting from OH\(^{-}\)), the barrier was calculated to be 1.76 eV from the entrance complex (Struct 1 → TS 1), 0.62 eV from the starting materials (LEDC + OH\(^{-}\) → TS 1) for the black path, and 1.09 eV (LEDC + OH\(^{-}\) → TS 4) for the green path (Figure 6b).

In polymolecular reactions, such as the hydrolysis of LEDC, there is a possibility that the entrance complex before entering the transition state region is lower in energy than the sum of the starting materials, in which case the rate-limiting step may be determined by the energy difference between the TS and the entrance complex. Hence, whenever there is a non-negligible discrepancy between using the sum of the free energies of the individual reactants and the free energy of the entrance complex as the baseline for the barrier height, we included the results for both scenarios.

We observe that the LEDC hydrolysis reactions are less kinetically favorable compared to the EC hydrolysis pathway and the hydride pathway based on our calculations; however, LEDC hydrolysis was found to occur both in DMSO-d\(_6\) solutions (with pure LEDC plus 1.5 equiv of water) and in the solid state on a time scale of minutes, as evidenced by \(^1\)H NMR and FTIR.\(^{27}\) Explicit solvation effects, especially in the case of water, are known to significantly impact kinetics,\(^{32,95,96}\) and further investigation is needed to elucidate the fast kinetics of this reaction. Indeed, if the surrounding water plays a pivotal, catalytic role in the hydrolysis of LEDC to LEMC, it may be less favorable in the battery environments, where water is assumed to be at impurity concentrations. The presence of the DMSO solvent could also be of relevance (pure LEDC can only be synthesized as LEDC·2DMSO\(^{27}\)) as...

Figure 6. LEMC paths from DFT.
LEDC is highly soluble in DMSO, and solvation effects are known to impact reaction thermodynamics as well as kinetics.

Besides the LEDC mediated pathways, two additional representative recombinant molecules were identified in the formation of LEMC: LODEC (lithium 2-(2-oxido-1,3-dioxolan-2-yl)ethyl carbonate) in path B and LODC (lithium 2-oxido-1,3-dioxolan-2-yl carbonate) in path C in Figure 5. These molecules have not been proposed, to our knowledge, in the SEI literature, although LODEC was observed in an AIMD simulation of EC-based electrolytes.97 The reaction network suggests both recombinants undergo hydrolysis and form LEMC. Examining the mechanistics of both reactions (Figure S3), we are unable to find a single-step hydrolysis reaction for both recombinants. Allowing for multistep reactions, in path B, LODEC and water transform to EC, LEC, and LiOH, crossing a 1.62 eV barrier, and we surmise that EC and LiOH subsequently react according to the path in Figure 4a to create LEMC. For path C, LODC initially decomposes to EC and Li$_2$CO$_3$ (which may exist as LiEC$^-$-LiCO$_3^{--}$), and LiCO$_3^{--}$ may facilitate water splitting and lead to the formation of a tetrahedral intermediate (Struct 4 in Figure S3b) with the subsequent formation of LEMC. However, this is equivalent to the EC + LiOH → LEMC path (Figure 4a), black path).

3.3. Reaction Paths to LEMC without Water. There are 8 reaction-network-identified paths starting from EC, Li$^+$ that do not involve water (see SI Sec V for the paths). However, these paths exhibit highly complex concerted reactions (involving 4 or 5 bond changes) connecting unstable open-shell species, with likely short lifetimes; those reactions were then verified to exhibit high barriers by DFT (see SI Sec V). We note that reactions involving up to 5 bond changes were included in the reaction network to accommodate hydrolysis reaction mechanisms. Hence, to explore nonwater containing reactions, another reaction network was constructed, limiting the concerted reactions to ≤2 bond changes and this reaction network consisted of 418 640 reactions. This procedure reveals some potentially feasible paths to LEMC starting with EC, Li$^+$ only (two representative paths are shown in Figure 7) and are discussed below. We note that these two paths are the two shortest paths from the 2-bond-change reaction network. The
The last step (LiHCO$_3$ \( \rightarrow \)) as the process (according to ref 98) is presumably LiEC and the associated reaction barriers. In fact, the C-H bond breaking in LEDC, in the absence or presence of radicals in the neighboring environment. The favorable thermodynamics of this path suggests that other radical species and/or the Li surface could facilitate a similar bond cleavage (a homolytic cleavage of the (CH$_2$O)O- moiety which would immediately abstract another H$^\cdot$ (from for example EC) to form LEMC. Soto et al. reported LEDC decomposition in AIMD simulations, including CO$_2$ and CO$_2$ detachment from LEDC, in the absence or presence of radicals in the solution. In fact, the C-O bond breaking of LEDC on the Li (100) surface, creating CO$_2$ bond to the surface, was calculated from DFT (PBE functional) as $\Delta E = 0.22$ eV and $\Delta E = -1.94$ eV. The resulting fragment may be further decomposed by breaking the -CH$_2$-O bond, however with a high associated barrier (1.8 eV). Hence, it is reasonable to assume the possibility of LiOC(==O)CH$_2$OLi undergoing H-abstraction reactions to create LEMC, although further investigations are recommended regarding the H$^\cdot$ or H$^+$ source and the associated reaction barriers.

For path B in Figure 7, we were unable to locate the TS for the last step (LiHCO$_3$ + o-LiEC \( \rightarrow \)) LEDC + LiCO$_2$ with DFT. An overall neutral TS was identified, corresponding to LiHCO$_3$ + o-LiEC \( \rightarrow \) LEMC + LiCO$_2$ (Figure S4); however, the step exhibits a high barrier of 2.75 eV, and hence, under the current modeling conditions, it is presumably not kinetically feasible.

Finally, we comment on what the presented work means in the likelihood of LEMC being present in the early stage Li metal or graphite SEI, instead of LEDC. All identified, feasible paths to LEMC (the EC hydrolysis path, the hydride path, the LEDC hydrolysis path, the path involving (CH$_2$O)O-OLi bond breakage in LEDC) either required water and/or LEDC. Indeed, the currently proposed reaction mechanisms for forming LEMC in the literature all require water participation. Hence, assuming trace water impurities in commercial electrolytes (~10 ppm of water in rigorously dried electrolyte) would not produce detectable amount of LEMC. However, it is possible that some water may originate from the cross-talk between the positive and negative electrodes (e.g., from the oxidation of EC at the cathode to produce water and CO$_2$). Our only nonwater containing path is effectively a conversion of LEDC into LEMC and as such requires LEDC to form first.

4. CONCLUSIONS

In this work, we developed a graph reaction network based on a database of systematically generated DFT-calculated molecular fragments and recombinants, specifically to explore reaction mechanisms for the formation of LEMC, relevant for the interface formation in lithium-ion batteries. The network allows for reactivity under a grand canonical formalism, treating all species as open. Our constructed reaction network is very dense, consisting of over 570 species and ~9 M reactions; this rich data allows us to embrace an unprecedented number of electrochemical reaction possibilities and predict reaction mechanisms without explicit knowledge of intermediates. A cost function based on the Gibbs free energy of the pathways was employed to rank the pathways, which were also manually inspected. Predominant pathways were further evaluated for mechanistic and kinetic feasibility using DFT. All relevant molecules, including our target molecule LEMC, were generated from fragmentation/recombination of the starting materials (EC, Li$^+$ and H$_2$O) with no imposed bias toward known SEI components. We note that our methodology can in principle be applied to any electrochemical, reactive system, although caution should be applied to avoid the rapid growth of combinatorial complexity. To prune the combinatorial complexity of possible recombinant molecules, machine learning was employed to predict reaction free energies to determine the feasibility of recombinant molecules. However, we emphasize that, in the present work, no other limitations (e.g., a priori identification of reactive sites) are imposed except for the elimination of recombinants with irregular valence.

Exploring the formation of LEMC with any amount of EC, H$_2$O, Li$^+$, and electrons at the Li metal chemical potential, we successfully recovered—automatically—the two major literature-proposed mechanisms, namely, EC and LEDC hydrolysis. Based on detailed DFT calculations of the reaction barriers, the direct EC hydrolysis mechanism was identified as the most kinetically feasible under the current model conditions, with the further suggestion of being significantly accelerated under basic conditions. In addition to these known paths, we also identify a hydride-generating mechanism and a radical mechanism through the (CH$_2$O)-O-OLi bond cleavage in LEDC, although those two mechanisms are more likely to be relevant on/near the Li metal surface.

Importantly, we find that nearly all feasible routes to LEMC, identified by the reaction network, require the participation of water, which is consistent with previously proposed mechanisms. Since the formation of LEDC does not require water, our results suggest that decreasing or increasing the
Despite these caveats, this work represents the first exploration of reaction pathways to LEMC formation assuming model conditions, such as the absence of an explicit surface as well as explicit solvation effects, both of which can contribute significantly to the mechanistics of the reactions. Despite these caveats, this work represents the first exploration of reaction pathways to LEMC based on a general, graph network of first-principles thermodynamics. The predicted pathways reproduce both existing, literature-suggested pathways as well as novel ones. Such predictions are critical for understanding the early stage formation of the SEI and may be leveraged to develop next-generation electrolytes/additives for Li-ion batteries. Our work illustrates that our methodology can be used for discovering novel reaction mechanisms and demonstrate the potential of the chemically consistent reaction network formalism for automated probing of complex, out-of-equilibrium electrochemical reaction mechanisms.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c05807.

Details on the MILP algorithm; details on the LEMC paths from the reaction network; details on the water splitting reaction on the Li surface; supplementary free energy diagrams; details on kinetically infeasible LEMC paths without water participation; effects of explicit solvation on the reaction paths; effects of using different dielectric constants in the implicit solvation model; analysis of the BondNet prediction error (PDF) xyz files for all the stationary point structures (ZIP) Information on molecules used to construct the reaction network and all unique LEMC paths from the reaction network (ZIP) (more description given in SI Sec II) Visualization of the unique LEMC paths from the S-bond-change reaction network (PDF) Visualization of the unique LEMC paths from the 2-bond-change reaction network (PDF)

## AUTHOR INFORMATION

### Corresponding Author

**Kristin A. Persson** — Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States; Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; orcid.org/0000-0003-2495-5509; Email: kapersson@lbl.gov

### Authors

**Xiaowei Xie** — Department of Chemistry, University of California, Berkeley, California 94720, United States; Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; orcid.org/0000-0001-5618-8768

**Evan Walter Clark Spotte-Smith** — Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States; orcid.org/0000-0003-1554-197X

**Mingjian Wen** — Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

**Hetal D. Patel** — Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Materials Science and Engineering, University of California, Berkeley, California 94720, United States

**Samuel M. Blau** — Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; orcid.org/0000-0003-3132-3032

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/jacs.1c05807

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

Early software infrastructure development and applications to Li metal reactions was supported by the Battery Materials Research (BMR) program directed by Tien Duong under the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the U.S. Department of Energy, Contract DEAC02-05CH11231. Data production and network algorithmic development was supported by the Joint Center of Energy Storage Research (JCESR), an Energy Innovation Hub funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences. The machine learning algorithm development was supported by the Silicon Electrolyte Interface Stabilization (SEISta) Consortium directed by Brian Cunningham under the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the U.S. Department of Energy, Contract No. DE-AC02-05CH11231. Calculations were performed at National Energy Research Scientific Computing Center (NERSC), Lawrence computational cluster at Lawrence Berkeley National Laboratory (LBNL), the Eagle HPC system at the National Renewable Energy Laboratory (NREL), as well as the UC Berkeley Molecular Graphics and Computation Facility (MGCF) supported by grant NIH S10OD02332. We acknowledge Schrodinger Inc. for providing AutoTS software licenses and technical assistance. X.X. thanks Dr. David W. Small for assistance with DFT calculations.

## REFERENCES

1. Winter, M. The solid electrolyte interphase—the most important and the least understood solid electrolyte in rechargeable Li batteries. Z. Phys. Chem. 2009, 223, 1395–1406.


(20) Yang, C.; Wang, Y.; Wan, C. Composition analysis of the passive film on the carbon electrode of a lithium-ion battery with an EC-based electrolyte. J. Power Sources 1998, 72, 66–70.


(37) Leung, K.; Budzien, J. L. Ab initio molecular dynamics simulations of the initial stages of solid–electrolyte interphase formation on lithium ion battery graphite anodes. Phys. Chem. Phys. 2010, 12, 6583–6586.


