

Computational and Machine Learning-Assisted Discovery and Experimental Validation of Conjugated Sulfonamide Cathodes for Lithium-Ion Batteries

Xuan Zhou Cheng Xu Xiaolong Guo Petru Apostol Alexandru Vlad René A. J. Janssen Süleyman Er*

Dr. Xuan Zhou

DIFFER – Dutch Institute for Fundamental Energy Research, De Zaale 20, 5612 AJ, Eindhoven, the Netherlands

Department of Applied Physics, Eindhoven University of Technology, 5600 MB, Eindhoven, the Netherlands

Cheng Xu

DIFFER – Dutch Institute for Fundamental Energy Research, De Zaale 20, 5612 AJ, Eindhoven, the Netherlands

Helmholtz Institute Ulm, Helmholtzstraße 11, D-89081, Ulm, Germany

Dr. Xiaolong Guo, Dr. Petru Apostol, Prof. Alexandru Vlad

Institute of Condensed Matter and Nanosciences, Molecular Chemistry, Materials and Catalysis, Université catholique de Louvain, B-1348, Louvain-la-Neuve, Belgium

Prof. René A. J. Janssen

DIFFER – Dutch Institute for Fundamental Energy Research, De Zaale 20, 5612 AJ, Eindhoven, the Netherlands

Molecular Materials and Nanosystems, Institute for Complex Molecular Systems, Eindhoven University of Technology, 5600 MB, Eindhoven, the Netherlands

Dr. Süleyman Er*

DIFFER – Dutch Institute for Fundamental Energy Research, De Zaale 20, 5612 AJ, Eindhoven, the Netherlands

Email Address: s.er@differ.nl

Keywords: *High-throughput Computational Screening, Conjugated Sulfonamides, Lithium-ion Batteries, Electrochemistry*

Conjugated sulfonamides (CSAs) stand out for their propitious electroactivity and notable stability in ambient conditions, making them suitable candidates for high-potential cathode materials in lithium-ion batteries (LIBs). This study employs a combination of machine learning, semi-empirical quantum mechanics, and density functional theory methods to evaluate a large library of 11,432 CSA molecules, focusing on material properties crucial for application in batteries, such as synthetic complexity, redox potential, gravimetric charge capacity, and energy density. After applying the thresholds for the synthetic complexity score at 2.62 and the redox potential at 3.25 V vs Li/Li⁺, we identify 50 CSA molecules that are easy-to-synthesize and suitable for the positive electrode in LIBs. By ranking on the basis of redox potential, 13 CSA molecules having potentials greater than 3.50 V vs Li/Li⁺ are identified. Through further investigations using molecular dynamics simulations on these reactant molecules and their lithiated products, we single out a molecule for synthesis and electrochemical evaluation. This molecule, lithium (2,5-dicyano-1,4-phenylene)bis((methylsulfonyl)amide)(Li2-DCN-PDSA), demonstrates a redox potential surpassing those previously reported within the class of CSA molecules. Moreover, the study explores the quantitative structure-property relations of CSAs, yielding insights for the development of CSA-based LIB cathode materials, informed by the comprehensive data assembled.

1 Introduction

Organic compounds with various redox functionalities have been increasingly developed and researched as electrode materials in rechargeable lithium-ion batteries (LIBs), owing to their distinct merits of natural abundance, environmental benignancy, structural diversity, and molecular-level controllability [1, 2, 3, 4].

According to the characteristics of uptaking and donating electrons in the neutral state of redox-active moieties, organic electroactive materials can be classified into *n*-type, *p*-type, and bipolar-type systems [5]. *N*-type organic species, when compared to their *p*-type counterparts, often offer less redox-inactive weight [6]. Furthermore, the lithiated *n*-type materials require lower amounts of organic electrolytes and Li salts, thus meeting the practical requirements for battery cathodes that need to be paired with Li-free anodes [7]. Moreover, *n*-type materials encompass a wide range of redox-active moieties, such as S-S, C=O, C=N, C=C, N=N, C≡N, and C≡C, allowing for the diverse design of cathode materials [6, 8].

To date, apart from the widely investigated organosulfur [9] and carbonyl-based compounds [10], new classes of organic materials, including imines [11], azo compounds [12], conjugated sulfonamides [13], oximates [14], triflimides [15], and cyanamides [15] have been successively applied as the electrode materials. These findings have enriched the library of organic electrode materials, and have also led to major advances in the electrochemical performance of organic materials not only for LIBs, but also for Na-ion and K-ion storage. Meanwhile, these studies also demonstrate the challenges in designing high-potential *n*-type organic cathodes, which are necessary for achieving high-energy-density LIBs. Among the *n*-type organic compounds, the recently disclosed conjugated sulfonamides (CSAs) [13] have achieved high redox potentials that can outcompete the carbonyl-based compounds [16] (2 – 3 V vs Li/Li⁺). The CSA derivatives also showed favorable cathode performance, such as excellent cycling stability, and good power density for both LIBs and other alkali (Na, K)-ion batteries [13, 17]. Some lithiated CSA phases have been also found with intrinsic oxidation and hydrolysis resistance under ambient atmosphere, enabling their direct handling and use as cathode materials in LIBs. Hence, they bring the development of organic cathode materials one step closer towards practical commercial use. This new class of organic materials, with their redox chemistry regulated by C=N-SO₂CH₃ molecular subunits, holds great promise as high-potential battery cathodes.

The reported CSA chemistries have been mainly studied experimentally, and the few compounds covered in previous studies [13, 17, 18] are based on intuition or chemical design. To the best of our knowledge, computational studies on the design of CSA cathode materials have not been reported yet. Thus, a vast chemical space awaits exploration to unearth completely new CSA cathode candidates for LIBs that could possess energy-related properties superior to the experimentally validated CSAs. Multi-scale computational simulations have been extensively applied to provide insights into the physical and chemical mechanisms of electrode materials across various battery applications [19, 20, 21, 22]. With the rapid advancement of artificial intelligence (AI) technologies and their emerging applications in material science, machine learning (ML) techniques are increasingly being integrated with these methods for material property prediction and inverse design [23, 24, 25, 26]. This integration has accelerated the discovery of battery materials. Additionally, incorporating combinatorial computational techniques into high-throughput computational screening (HTCS) has proven to be an effective strategy for navigating large chemical spaces and identifying potential candidates for functional battery materials [27, 28].

Here, we present a material discovery workflow that integrates multiple computational techniques, including an ML model, semi-empirical quantum chemistry and DFT methods, molecular dynamic simulations, and automated searches of chemical vendors, culminating in the synthesis of a top candidate compound. This approach has been applied to explore a sub-chemical space of CSAs, effectively screening for high-redox candidate molecules. First, a virtual library of 11,432 CSA molecules is created by exhaustively enumerating functionalized derivatives on specially-chosen CSA backbones with chemical R-groups. Then, we determine the most efficient and accurate approach for performing electronic structure simulations across the entire chemical molecular library, and apply it to predict the redox potential values of all the CSAs. Next, we use the simplified molecular-input line-entry system (SMILES) representations of CSAs in a machine-learning (ML) model to predict the synthetic complexity score (SCScore) of the molecules. In addition, we calculate important energy-related properties of the new compounds, such as the gravimetric specific capacity (Q) and gravimetric energy density (W). Subsequently, the SCScore and predicted redox potential are employed to down-select the easy-to-synthesize and high-potential CSA molecules. Moreover, molecular dynamics (MD) simulations are performed for a group of CSA molecules with a predicted redox potential greater than 3.50 V vs Li/Li⁺ to investigate their thermodynamic stability, and a new molecule is selected from this group for experimental validation. Lastly, the molecular design

principles are elucidated to guide the future engineering of CSA molecules for LIBs.

2 Results and Discussion

The virtual library of CSA molecules is constructed through the fusion of chemical building blocks within predefined design principles, followed by a systematic enumeration of functional derivatives using chemically functional R-groups. Three chemical building blocks, namely *para*-CSA (redox-active groups at the *para* position), *ortho*-CSA (redox-active groups at the *ortho* position), and benzene, are used to generate the core CSA structures. The use of *para*-CSA and *ortho*-CSA building blocks potentially facilitates $2\text{Li}^+ + 2e^-$ reactions for the generated molecules during the (de)lithiation process (Supporting Information Figure S1(a)). In this molecular fusion process, we apply boundary conditions for the total number of rings and redox-active groups allowed on molecules. For instance, the upper limit for the number of six-membered rings in the core CSA structures is three. Besides keeping the size of chemical library feasible for computational screening, this limitation is based on the observation that the single-molecule model is prone to have larger errors when predicting the properties of cathode materials that are formed by multiple six-membered rings, which is essentially a result of the increased intermolecular π -stacking interactions [29]. Additionally, each CSA core is designed to include two redox-active groups, aiming to be consistent with the molecules that have been researched in experimental works [13] on CSAs. In this context, the chemical building blocks *para*-CSA and *ortho*-CSA are exclusively combined with benzene, resulting in 14 unique core CSA structures. These structures consist of 2 one-ring, 3 two-ring, and 9 three-ring molecules, which are shown in Figure 1a.

Next, a comprehensive R-group enumeration is performed to systematically generate functionalized CSA derivatives. In this process, the aromatic H atoms of the core CSA structures are substituted with various specific chemical R-groups. Six different chemical R-groups are considered for this purpose, including $-\text{CN}$, $-\text{F}$, $-\text{Cl}$, $-\text{CH}_3$, $-\text{OCH}_3$, and $-\text{NH}_2$, as shown in Figure 1a. These groups, which either withdraw or donate electrons (highlighted in orange and blue in Figure 1a), are integrated into the core molecular structures to broaden their range of electrochemical properties. The complete enumeration process, as shown in Figure 1b, allows for substitutions from none to the maximum number of aromatic hydrogen atoms found in the core CSA structures, details of which are provided in Supporting Information Table S1. Using the Custom R-Group Enumeration package within SMSS, a virtual chemical library of 11,432 CSA molecules is created for computational screening. Further information on the generated CSA molecules and the number of functionalized derivatives obtained by using each unique R-group are provided in Supporting Information Tables S1 and S2.

The commercial vendor information for the enumerated CSA molecules is automatically retrieved from ChemSpace [30] and Molport [31] platforms using the ChemPrice [32] software tool and the ZINC database [33] via an *in-house* developed tool, respectively. Both tools utilize SMILES strings of CSA molecules as inputs and collect the available vendor information from the molecular databases. A comprehensive search of our molecular library across the ChemSpace, Molport, and ZINC databases revealed that none of the compounds were listed for available to purchase. As a result, it was not possible to directly acquire any new CSA compounds for experimental evaluation. This finding emphasizes the uniqueness of CSA molecules and underscores the vital role of computational research in investigating their characteristics. Given that no commercial vendors offer any compounds from our library, we assessed the synthesizability of these virtual molecules using an ML model previously applied to quinone-based battery materials [27] to quantify their synthesis complexity.

To determine a workable and yet accurate method for performing large-scale electronic structure simulations on our virtual molecular library, we evaluate the performance of an adapted chemical descriptor. This descriptor is calculated using different approaches, including density functional theory (DFT), and a hybrid approach of semi-empirical quantum mechanics (SEQM) and DFT, aimed at predicting the experimentally observed redox potentials. Selecting the lowest unoccupied molecular orbital (LUMO) energy of CSA molecules as the chemical descriptor, we use the coefficient of determination (R^2) and root-mean-square error (RMSE) to gauge prediction performance (Supporting Information Figures S1 (b) and (c)). The

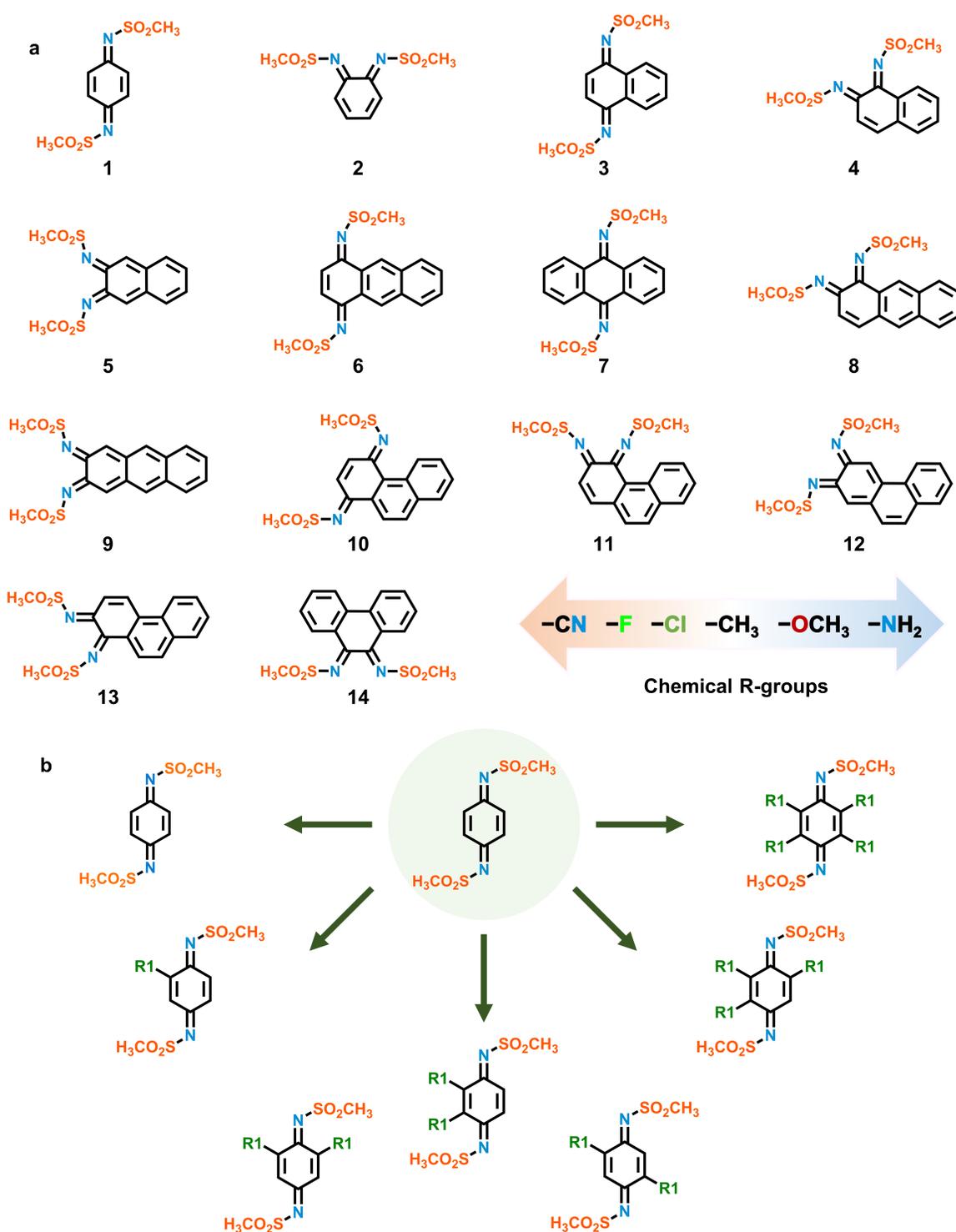


Figure 1: (a) Two-dimensional (2D) structural representations of the 14 core molecules, alongside the selected chemical R-groups. In the color bar, the orange and blue represent the electron-withdrawing and electron-donating R-groups, respectively. (b) R-group enumeration process for *para* conjugated sulfonamide.

LUMO energy of reactant CSA molecules, with $R^2 > 0.93$ and $RMSE < 0.10$ V for all the computational methods tested, is found to be a sufficiently good descriptor for predicting the redox potentials of CSA cathode materials via HTCS. The hybrid approach scheme involving PBE single-point-energy calculations with solvation effects, applied on PM7 optimized geometries ($R^2 = 0.96$, $RMSE = 0.057$ V), is chosen for predicting the redox potentials of 11,432 CSA molecules. This approach offers comparable accuracy to the full PBE approach ($R^2 = 0.97$, $RMSE = 0.047$ V), but at a computational cost approximately 13 times lower than the purely PBE approach. Further details on redox potential predictions are provided in the Supporting Information.

The relationships between the ML-predicted SCScore values and the calculated $E_{PM7/PBE}^o$ and Q for all CSA molecules within the virtual library are depicted using scatter plots as shown in Figures 2(a) and 2(b), respectively. Orange plus signs on these plots highlight the seven experimental CSAs, with their predicted values given in Supporting Information Table S3. The virtual molecular library has a broad distribution across each predicted property, as shown in Figures 2(a) and 2(b), allowing for the identification of CSA candidates with desirable features for cathode materials, such as high redox potential, low SCScore, and sizable gravimetric charge capacity. The SCScore values for the experimentally studied CSAs vary between 1.91 and 2.62, placing them in the lower regions of Figures 2(a) and 2(b). This highlights the significance of SCScores as a metric as well as the experimental CSA molecules' structural simplicity, usually characterized by a single six-membered ring and no more than two R-groups, in addition to their redox-active components. These molecules also achieve remarkably high theoretical Q values. Given the unavailability of CSA molecules for direct purchase in vendor databases, the SCScores of the experimentally studied CSAs serve as a threshold criterion in the first tier of HTCS. Consequently, CSA molecules with SCScores above this limit are not considered for further analysis. With an emphasis on experimental feasibility and synthesis simplicity, the highest SCScore from the experimental set, 2.62, is used as the threshold criterion, marked by a horizontal dotted line in Figures 2(a) and 2(b). As a result, a narrowed group of 231 CSA molecules is left for further study.

Following the virtual screening based on the SCScore, the selection process continues with a further refined analysis within the virtual library. As shown in Figure 2(a), the range of redox potentials among the 231 shortlisted CSA molecules spans from 1.85 V to 4.52 V vs Li/Li⁺. Previously, the experimentally studied CSA molecules have demonstrated redox potentials within the range of 2.38 V to 3.25 V vs Li/Li⁺. Among the 231 CSAs, 198 molecules have redox potentials above 2.38 V, and among these, 50 exceed the 3.25 V threshold. Given the need for high redox potentials in *n*-type cathodes for LIBs, a benchmark redox potential of 3.25 V is set as a minimum cutoff, marked by a vertical dotted line in Figure 2(a). Applying this criterion leads to the identification of 50 CSA molecules as promising candidates, featuring theoretical Q values ranging from 84 mAh/g to 187 mAh/g.

Further studies involving molecular dynamics (MD) simulations are conducted to estimate the stability of the candidate molecules under conditions of 300 K and 1 atm for a duration of 100 ns. This evaluation focuses on observing the root mean square deviation (RMSD) across all atoms. Among the 50 screened high-redox-potential CSA candidates, those with the highest potentials were prioritized for further investigation using MD methods. After ranking the redox potentials, 13 molecules with values greater than 3.50 V vs Li/Li⁺ are selected. These MD simulations are conducted to investigate both the reactant molecules and their corresponding lithiated products. The results, including the distribution of $E_{PM7/PBE}^o$ and average RMSD values, are shown in Figure 2(c). Additional information, such as 2D structural representations, and numerical values for $E_{PM7/PBE}^o$, theoretical Q , W , average RMSD, and the maximum root mean square fluctuation (RMSF), is available in Table S4 of the Supporting Information. This analysis indicates that both the pure and lithiated versions of the molecules exhibit considerable structural stability, with RMSD values not exceeding 1.55 Å. For experimental assessment, we chose lithium, we chose lithium (2,5-dicyano-1,4-phenylene)bis((methylsulfonyl)amide) (Li₂-DCN-PDSA) from the top list of 13 molecules due to its high redox potential, low SCScore value, and a synthetic protocol similar to that in the previous study [13], thereby enabling feasible experimental validation. The selected Li₂-DCN-PDSA, marked by a purple dot in Figure 2(c), shows promising stability over the 100-ns MD simulation, evidenced by an average RMSD 1.22 Å as shown in Figure 2(d).

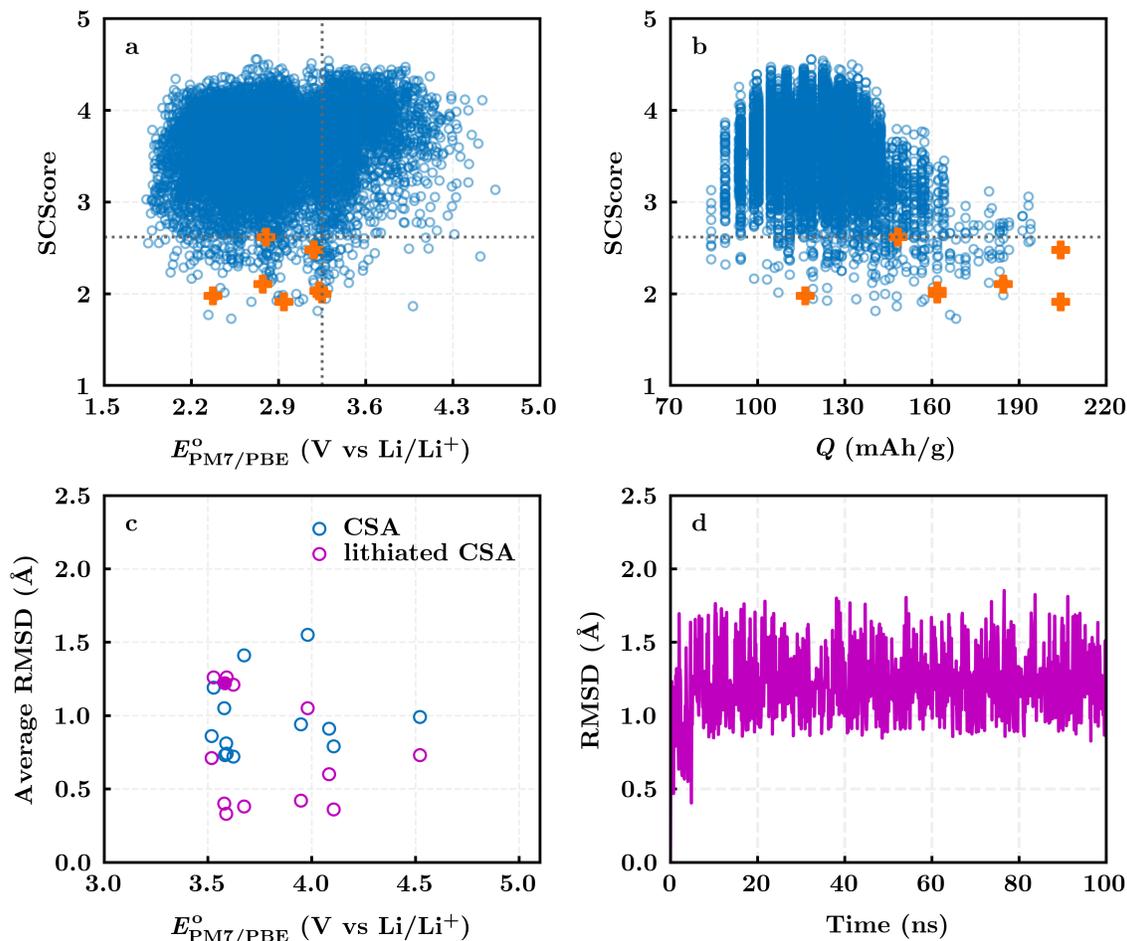


Figure 2: 2D scatter distribution of (a) SCScore versus $E_{PM7/PBE}^o$ and (b) SCScore versus Q . The orange pluses represent the experimentally tested CSAs. The dotted horizontal lines in both (a) and (b) represent the SCScore cutoff value of 2.62. The dotted vertical line in (a) indicates the redox potential cutoff value, $E_{PM7/PBE}^o = 3.25$ V. (c) The distribution of the average values of RMSD versus $E_{PM7/PBE}^o$ for 13 CSA molecules and their associated lithiated molecules. The blue and purple circles denote the reactant and product molecules, respectively. The purple dot represents the experimentally validated molecule. (d) The 100-ns MD trajectory of the experimentally validated molecule.

The Li₂-DCN-PDSA molecule is synthesized according to the reaction process outlined in Figure 3(a). Four sulfonyl groups are firstly grafted onto 2,5-diaminoterephthalonitrile (H₄-DCN-DAB) by reacting it with methanesulfonyl chloride in pyridine. Then, a solution of tetrabutylammonium fluoride (TBAF) in tetrahydrofuran (THF) is added to yield the protonated compound H₂-DCN-PDSA. The Li₂-DCN-PDSA compound is prepared through a lithiation reaction by adding lithium methoxide in anhydrous methanol. The synthesis details are available in the Supporting Information. The structure of the targeted molecule is confirmed by the proton nuclear magnetic resonance (¹H NMR) and Fourier transform infrared (FTIR) spectroscopy, as shown in Figures 3(b) and (c). In the FTIR spectrum (red) of Li₂-DCN-PDSA, shown in Figure 3(c), the peaks at 2937, 2814 cm⁻¹, and 1263, 1114 cm⁻¹ correspond to the vibrations of the C–H bond and the S=O bond in the redox-active C=N–SO₂CH₃ node, respectively, which are in good agreement with the previous study [13]. The peak at 2232 cm⁻¹ is attributed to the C≡N bond vibration, consistent with previously reported results [34]. Then, the cyclic voltammetry (CV) profile of Li₂-DCN-PDSA is carried out in 0.1 M LiTFSI in DMSO, shown in Figure 3(d), reveals two redox processes centered at 0.15 V and -0.19 V vs Fc/Fc⁺ (equivalent to 3.58 V and 3.24 V vs Li/Li⁺, represented by vertical dotted lines in Figure 3(d)), aligning closely with the predicted value of $E_{PM7/PBE}^o = 3.58$ V obtained using the PM7/PBE method. The Li₂-DCN-PDSA compound exhibits a superior redox potential compared to previously reported values [13] within the class of CSAs. These findings demonstrate that

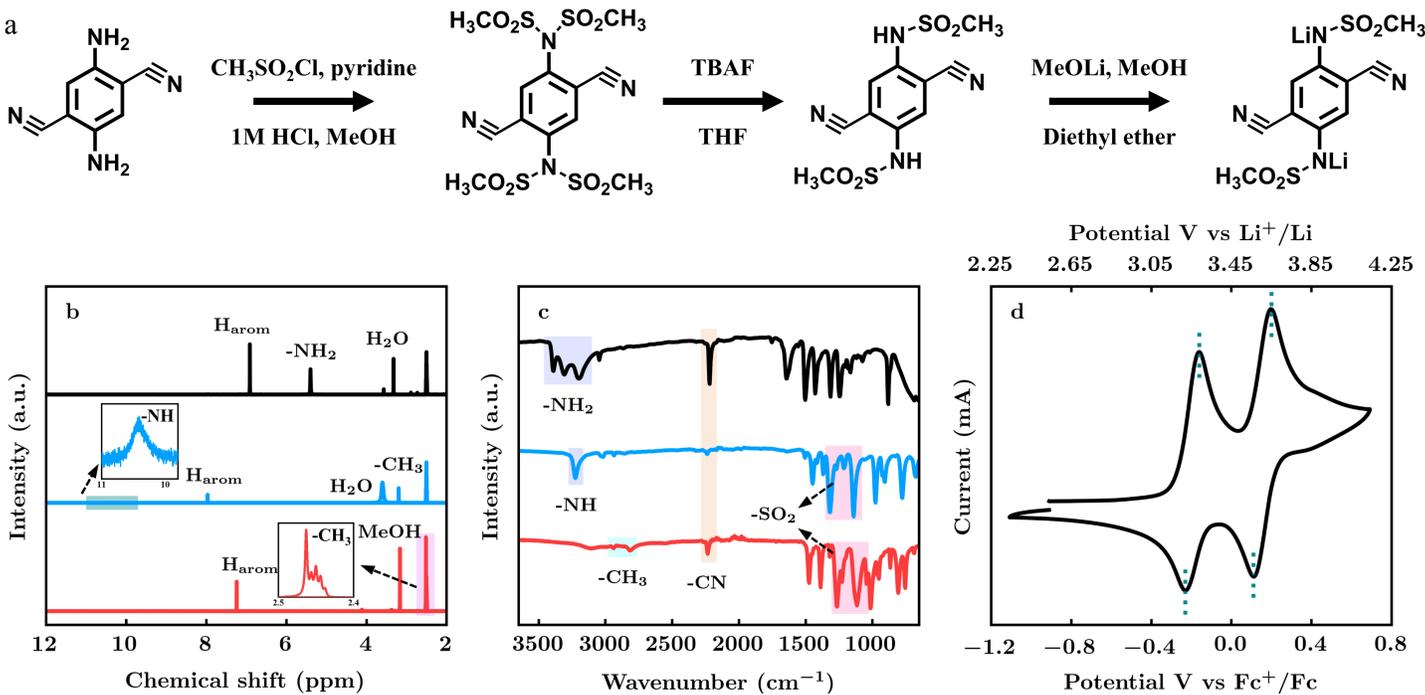


Figure 3: (a) Synthetic route for Li₂-DCN-PDSA compound. (b) ¹H NMR and (c) FTIR spectra of H₄-DCN-DAB (black), H₂-DCN-PDSA (blue), and Li₂-DCN-PDSA (red) compounds. (d) Cyclic voltammetry of Li₂-DCN-PDSA with the electrolyte of 0.1 M LiTFSI in DMSO.

our computational strategy is both a promising and reliable method for screening high-redox CSA cathode candidates for LIBs. Moreover, the designed material discovery workflow can be directly applicable to an expanded chemical space of CSA materials, involving a broader variety of core structures and R-groups, and integrating these in various ways in the enumeration process. Moreover, computational modeling across various spatial and temporal scales [22] can be further integrated into this workflow for the study of CSA materials, incorporating validated crystal and aggregated structure models in future studies. To advance the application of CSA materials in LIB technologies, it is essential to complement computational findings with extensive experimental exploration in future studies. This includes galvanostatic testing and full electrochemical characterization of top CSA compounds, as well as synthesizing and testing additional candidate molecules. CSAs and their lithiated molecules have shown excellent air stability and good compatibility with commonly used organic electrolytes, such as lithium salts (LiTFSI and LiPF₆) and organic solvents (tetraglyme, PC, DOL/DME, and EC/DMC) [13, 35, 36]. Moreover, optimizing electrode formulations and their associated organic electrolytes is also crucial for the practical use of CSA materials in LIB applications.

Next, the relationships between the structural features of molecules—specifically, the types of chemical functional groups, the count of aromatic rings, and the varieties of redox-active groups—and their calculated properties like redox potential and SCScore are explored. These insights are useful for future molecular engineering efforts, aiming to develop other unique CSA-derived compounds as candidate cathode materials for LIBs. The effects of different chemical functional groups on the predicted $E_{\text{PM7/PBE}}^{\circ}$ and SCScores are shown in Figures 4(a) and 4(d), with a detailed examination of ring size effects provided in Figure S2 of the Supporting Information. Figure 4(a) reveals how redox potentials progressively increase with the electron-withdrawing capabilities of functional groups, ranked as $-\text{CN} > -\text{F} > -\text{Cl} > -\text{CH}_3 > -\text{OCH}_3 > -\text{NH}_2$. Notably, CSA molecules with $-\text{CN}$ groups achieve significantly higher redox potentials compared to those with other electron-withdrawing groups, such as $-\text{F}$ and $-\text{Cl}$. Additionally, the analysis presented in Figure S2 suggests that CSA molecules with a single aromatic ring attain a narrower range of redox potentials relative to the molecules with two or three rings, especially for those functionalized with $-\text{F}$ and $-\text{Cl}$, hinting at the subtle impact of these groups on the redox potentials of single-ring CSA

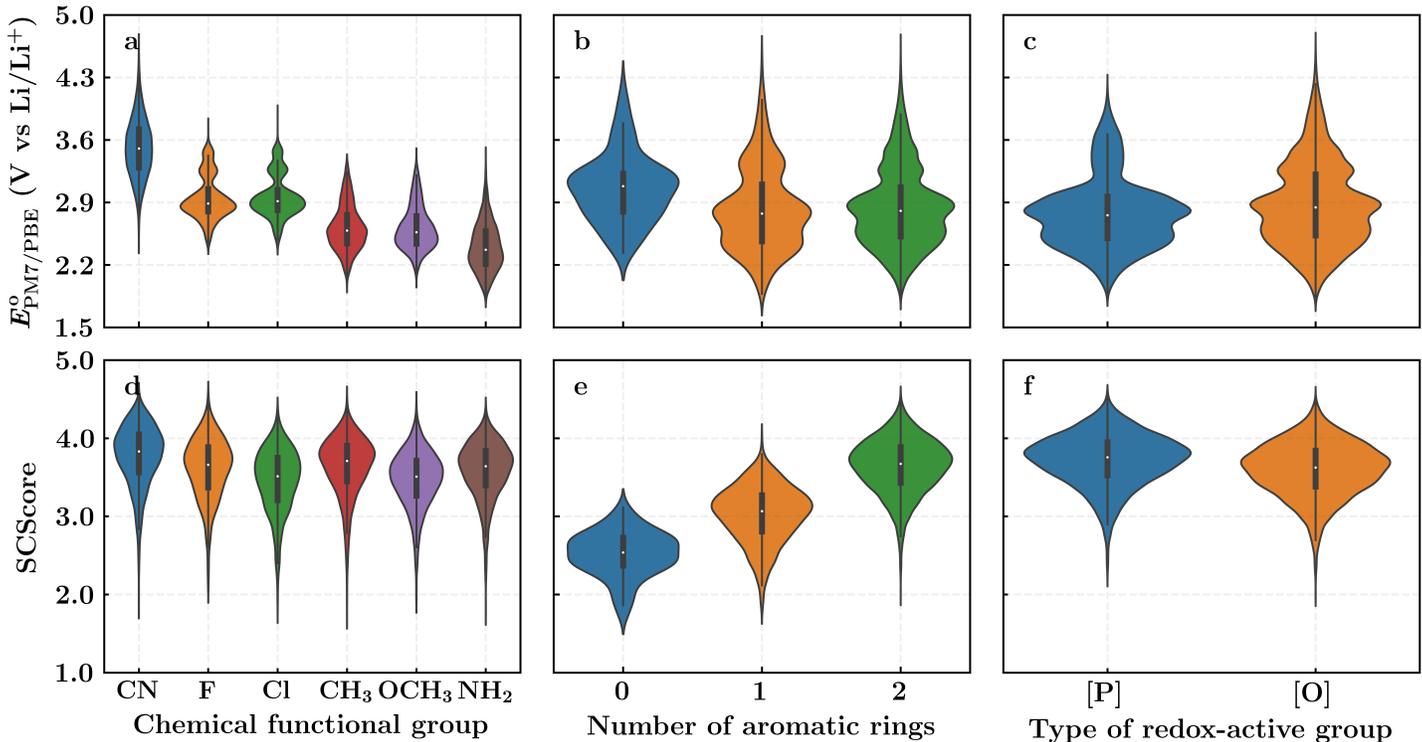


Figure 4: Violin plots show the distributions of predicted redox potential $E_{\text{PM7/PBE}}^{\circ}$ (a, b, c) and SCScore (d, e, f) for 11,432 CSA candidates within the virtual library, categorized by the type of chemical functional group (a, d), the number of aromatic rings (b, e), and type of redox-active group (c, f) found on the molecule. In each plot, the median values are indicated by the white dots inside the black bars.

variants.

No clear relationship has been established between the SCScore and the type of chemical functional group, as shown in Figure 4(d). Here, CSA molecules with different functional groups exhibit comparable SCScore profiles. Nevertheless, molecules that incorporate the $-\text{Cl}$ functional group generally show marginally lower SCScores, a pattern that mirrors findings in organic quinone studies [27].

Regarding the effect of the number of aromatic rings, shown in Figures 4(b) and 4(e), CSA molecules that contain aromatic rings demonstrate a wide range of predicted $E_{\text{PM/PBE}}^{\circ}$, owing to the increased opportunities for structural alterations through substitutions, which can adjust their redox properties. Molecules having fewer aromatic rings are typically associated with higher redox potentials and lower SCScores, which are both desirable traits for organic cathode materials. However, the challenge with smaller molecules lies in their solubility in organic electrolytes, which can compromise capacity retention in cathode materials within LIBs. Therefore, it's essential to aim for a well-rounded balance among key properties such as E° , Q , SCScore, and solubility, in the design of potential cathode materials for LIBs.

Within the virtual library, the positioning of redox-active groups in CSA molecules is categorized into either *para* ([P]) or *ortho* ([O]) configurations. As shown in Figure 4(c), *ortho*-CSA variants exhibit slightly higher predicted $E_{\text{PM7/PBE}}^{\circ}$ values in comparison to their *para*-CSA equivalents. Additionally, a significant portion of *para*-CSA compounds are predicted to have $E_{\text{PM7/PBE}}^{\circ}$ values below 3 V. The influence of redox-active group placement on the SCScore is indistinct, similar to the observed impact of functional group types, as shown in Figure 4(f). Furthermore, the SCScore profiles for both *para*-CSA and *ortho*-CSA molecules appear alike. Further analysis comparing the effects of redox-active group positioning on both the predicted E° and SCScore, considering various ring configurations, is detailed in Figure S3 of the Supporting Information. With the rise of generative AI in material discovery and design [37], accompanied by the shift of materials research from "Fitting-Generation" to "Pretraining-Prompting-Generation", it becomes feasible to perform the inverse design of CSA materials using structure-activity relationship data for future LIB technologies.

3 Conclusions

In this work, we conduct a computational investigation of an extensive collection of 11,432 conjugated sulfonamide (CSA) compounds to assess their viability as potential cathode materials for lithium-ion batteries (LIBs). Using a comprehensive computational approach that merges semi-empirical quantum mechanics (SEQM-PM7) with density functional theory (DFT-PBE), we achieve the accuracy required for effective high-throughput computational screening (HTCS). We evaluate four key properties of these molecules—redox potential, SCScore, gravimetric charge capacity, and gravimetric energy density—to rank them across the entire virtual library. Utilizing the SCScore and redox potential benchmarks established by previously analyzed CSA molecules, we identify 231 CSAs with promising synthetic accessibility. Subsequent refinement spots 50 candidates with redox potentials surpassing 3.25 V vs Li/Li⁺. From this group, a top-ranking subset of 13 molecules, each with a potential greater than >3.50 V vs Li/Li⁺, undergoes molecular dynamics (MD) simulations, revealing average RMSD values of up to 1.55 Å. Among these, one molecule is chosen for experimental validation, showing properties aligning well with computational predictions. Lastly, an exploration of structure-property correlations provides insights aimed at steering the future synthesis and optimization of CSA compounds.

4 Methods

4.1 Electronic Structure Calculations

In this study, Schrödinger Materials Science Suite (SMSS) [38] is used for all the calculations. The computational scheme for large-scale electronic structure calculations includes the force field (FF) conformational search on all the generated molecules, structural optimizations (OPT) on these FF-minimized geometries using SEQM methods and DFT-level single-point energy (SPE) calculations on optimized coordinates in an implicit solvent phase. The standard Poisson–Boltzmann Formalism (PBF) [39] was applied as an implicit solvation model, requiring the solvent’s dielectric constant, molecular weight, and density. These physical parameters were set according to tetraethylene glycol dimethyl ether, the electrolyte used in experiments on conjugated sulfonamides [13]. The FF conformational search, SEQM optimizations, and DFT SPE calculations are performed using the MacroModel [40] module, MOPAC2016 [41] package, and Jaguar [42] program, respectively, all as implemented in SMSS. The OPLS3e [43] FF, PM7 [44] method and PBE functional are employed for finding the lowest-energy conformers, SEQM OPT simulations, and DFT SPE calculations, respectively. Further information on the computational methodology is provided in the Supporting Information.

4.2 Synthetic Complexity Prediction

To prioritize the compounds that can be produced in the lab within our custom-made molecular library, we apply a ML model developed by Coley *et al.* [45]. This model evaluates the synthetic complexity of a molecule by analyzing its association with the number of steps involved in its synthesis history. Molecules deemed simpler to synthesize receive lower scores, while those requiring more complex synthesis processes are assigned higher scores. The ML algorithm, which underpins the SCScore, has been refined through training on a dataset comprising 22 million reactant-product pairs from Reaxys database [46], which contains high-quality reaction data on 12 million reaction instances. Molecular representation in a machine-readable format is essential for predicting molecular properties [47]. In this study, we used the SMILES representation of molecules as the ML model for synthetic complexity predictions operates with this format. Upon entering a molecule’s SMILES notation into the model, it generates an SCScore on a scale from 1 to 5. This score is then utilized as a criterion for filtering a selection of organic CSA compounds, distinguishing those with a higher likelihood of straightforward synthesis. An SCScore value close to 1 indicates low complexity in the synthesis of a compound, whereas a value near 5 indicates high complexity. In this study, we set an SCScore threshold of 2.6 and selected CSA compounds with SCScores ranging from 1.0 to 2.6 for further screening. It is important to consider additional factors such as cost,

productivity, recyclability, toxicity, and safety when synthesizing these compounds at scale to better assess their practicality for large-scale use.

4.3 Gravimetric Charge Capacity and Energy Density

In addition to the redox potential, we calculate and utilize the theoretical gravimetric charge capacity (Q) and energy density (W) for CSA cathode materials, with the relevant equations provided in the Supporting Information.

4.4 Molecular Dynamics Simulations

For all molecular dynamics (MD) simulations in this study, the Desmond package [48] within the SMSS is utilized. Each MD system is set up within a cubic simulation box, measuring 50 Å along each edge. The OPLS4 FF is utilized to model atomistic interactions. The simulations run for a duration of 100 nanoseconds, featuring a relaxation period of 1 picosecond, and trajectory frames are captured at intervals of 100 picoseconds. The thermodynamic stability of the molecules is assessed by monitoring the root mean square deviation (RMSD) of all atoms and the root mean square fluctuation (RMSF) of the heavier atoms, excluding hydrogen. The average RMSD value is derived from an analysis of the 1000 frames collected during the simulation. Further specifics on the simulation parameters can be found in a previous study [49].

4.5 Experimental Measurements

The BioLogic Science Instruments SP-300 is utilized for electrochemical measurements. A three-electrode system with screen-printed platinum as the working and counter electrodes, and screen-printed silver as the pseudo-reference electrode are employed. The supporting electrolyte is 0.1 M LiTFSI in DMSO. 5 mM of Li₂-DCN-PDSA are dissolved in the organic solution. The Cyclic Voltammetry (CV) is performed with a scan rate of 100 mV/s. Following the completion of the test, ferrocene is introduced into the solution to serve as an internal reference.

Supporting Information

Supporting Information is available from the Wiley Online Library.

Code & Data Availability

The computed properties of 11,432 CSA molecules have been compiled into a dataframe containing information on SMILES representations, molecular weight, Q , $E_{\text{PM7/PBE}}^{\text{LUMO}}$, W , SCScore values, and commercial vendor page retrieved from the ZINC database whenever available. This dataframe is available as a self-contained Supplementary File in CSV format.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

X.Z. and C.X. contributed equally to this work. We acknowledge Murat Cihan Sorkun for technical support with price search using the ChemPrice tool. X.Z. and C.X. acknowledge financial support from China Scholarship Council, respectively, grant no. 201806240322 and 202006220082. A.V. is indebted to the

European Research Council for support under the European Union's Horizon 2020 research and innovation program (grant agreement no. 770870, MOOIRE), as well as partial support from F.R.S.-FNRS through the F.4552.21-P-MIS-CSA-LION grant. DIFFER is part of the institutes organisation of NWO. This work was sponsored by NWO Exact and Natural Sciences for the use of supercomputer facilities.

References

- [1] J. Kim, Y. Kim, J. Yoo, G. Kwon, Y. Ko, K. Kang, *Nat. Rev. Mater.* **2023**, *8* 54.
- [2] X. D. Chen, X. J. Yin, J. Aslam, W. W. Sun, Y. Wang, *Electrochem. Energy Rev.* **2022**, *5* 12.
- [3] G. Kwon, Y. Ko, Y. Kim, K. Kim, K. Kang, *Acc. Chem. Res.* **2021**, *54* 4423.
- [4] T. Huang, M. Q. Long, J. Xiao, H. Liu, G. X. Wang, *Energy Mater.* **2021**, *1* 100009.
- [5] Z. P. Song, H. S. Zhou, *Energy & Environ. Sci.* **2013**, *6* 2280.
- [6] Y. Lu, J. Chen, *Nat. Rev. Chem.* **2020**, *4* 127.
- [7] Y. Lu, Q. Zhang, F. J. Li, J. Chen, *Angew. Chem. Int. Ed.* **2023**, *62* e202216047.
- [8] Y. Lu, Q. Zhang, L. Li, Z. Q. Niu, J. Chen, *Chem* **2018**, *4* 2786.
- [9] Z. Shadike, S. Tan, Q. C. Wang, R. Q. Lin, E. Y. Hu, D. Y. Qu, X. Q. Yang, *Mater. Horiz.* **2021**, *8* 471.
- [10] H. Q. Yang, J. Y. Lee, J. Y. Cheong, Y. F. Wang, G. G. Duan, H. Q. Hou, S. H. Jiang, I. D. Kim, *Energy & Environ. Sci.* **2021**, *14* 4228.
- [11] T. Matsunaga, T. Kubota, T. Sugimoto, M. Satoh, *Chem. Lett.* **2011**, *40* 750.
- [12] C. Luo, O. Borodin, X. Ji, S. Hou, K. J. Gaskell, X. L. Fan, J. Chen, T. Deng, R. X. Wang, J. J. Jiang, C. S. Wang, *Proc. Natl. Acad. Sci.* **2018**, *115* 2004.
- [13] J. D. Wang, A. E. Lakraychi, X. L. Liu, L. Sieuw, C. Morari, P. Poizot, A. Vlad, *Nat. Mater.* **2021**, *20* 665.
- [14] J. D. Wang, P. Apostol, D. Rambabu, X. L. Guo, X. L. Liu, K. Robeyns, M. Y. Du, Y. Zhang, S. Pal, R. Markowski, F. Lucaccioni, A. E. Lakraychi, C. Morari, J. F. Gohy, D. Gupta, A. Vlad, *Sci. Adv.* **2023**, *9*, 17 6079.
- [15] X. L. Guo, P. Apostol, X. Zhou, J. D. Wang, X. D. Lin, D. Rambabu, M. Y. Du, S. Er, A. Vlad, *Energy & Environ. Sci.* **2023**, *17* 173.
- [16] L. M. Zhu, G. C. Ding, L. L. Xie, X. Y. Cao, J. P. Liu, X. F. Lei, J. X. Ma, *Chem. Mater.* .
- [17] J. D. Wang, X. L. Liu, H. Jia, P. Apostol, X. L. Guo, F. Lucaccioni, X. Z. Zhang, Q. Zhu, C. Morari, J. F. Gohy, A. Vlad, *ACS Energy Lett.* **2022**, *7* 668.
- [18] J. Wei, P. B. Zhang, T. Y. Shen, Y. Z. Liu, T. F. Dai, Z. X. Tie, Z. Jin, *ACS Energy Lett.* **2023**, *8*, 1 762.
- [19] A. Massaro, F. Fasulo, A. Pecoraro, A. Langella, A. B. Muñoz-García, M. Pavone, *Phys. Chem. Chem. Phys.* **2023**, *25*, 28 18623.
- [20] S. Zhang, J. Ma, S. M. Dong, G. L. Cui, *Electrochem. Energy Rev.* **2023**, *6*, 1 4.
- [21] Q. He, B. Yu, Z. H. Li, Y. Zhao, *Energy & Environ. Mater.* **2019**, *2*, 4 264.

- [22] S. Q. Shi, J. Gao, Y. Liu, Y. Zhao, Q. Wu, W. W. Ju, C. Y. Ou Yang, R. Y. Xiao, *Chinese Physics B* **2015**, *25*, 1 018212.
- [23] H. Y. Guo, Q. Wang, A. Stuke, A. Urban, N. Artrith, *Front. Energy Res.* **2021**, *9* 695902.
- [24] Y. Liu, B. R. Guo, X. X. Zou, Y. J. Li, S. Q. Shi, *Energy Storage Mater.* **2020**, *31* 434.
- [25] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, A. Walsh, *Nature* **2018**, *559*, 7715 547.
- [26] Y. Liu, T. L. Zhao, W. W. Ju, S. Q. Shi, *Journal of Materiomics* **2017**, *3*, 3 159.
- [27] X. Zhou, A. Khetan, J. Zheng, M. Huijben, R. A. J. Janssen, S. Er, *Digit. Discov.* **2023**, *2* 1016.
- [28] R. Pollice, G. dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. Nigam, C. T. Ser, Z. P. Yao, A. Aspuru-Guzik, *Acc. Chem. Res.* **2021**, *54* 849.
- [29] S. Manzhos, *Curr. Opin. Green Sustain. Chem.* **2019**, *17* 8.
- [30] ChemSpace, URL <https://chem-space.com/>.
- [31] Molport, URL <https://www.molport.com/>.
- [32] M. C. Sorkun, S. Baptiste, S. Er, *ChemRxiv* **2024**.
- [33] J. J. Irwin, B. K. Shoichet, *J. Chem. Inf. Model.* **2005**, *45* 177.
- [34] B. Bouchet-Fabre, E. Marino, G. Lazar, K. Zellama, M. Clin, D. Ballutaud, F. Abel, C. Godet, *Thin Solid Films* **2005**, *482* 167.
- [35] J. D. Wang, X. L. Guo, P. Apostol, X. L. Liu, K. Robeyns, L. Gence, C. Morari, J.-F. Gohy, A. Vlad, *Energy & Environ. Sci.* **2022**, *15*, 9 3923.
- [36] J. T. Liu, E. Grignon, A. M. Battaglia, M. Imran, C. Copeman, H. A. Mills, A. J. Howarth, E. H. Sargent, D. S. Seferos, *Chem. Mater.* **2023**, *35*, 22 9692.
- [37] Y. Liu, Z. W. Yang, Z. Y. Yu, Z. T. Liu, D. H. Liu, H. L. Lin, M. Q. Li, S. C. Ma, M. Avdeev, S. Q. Shi, *J. Materiomics* **2023**, *9*, 4 798.
- [38] Use Manual. Schrödinger, New York NY, USA, 2019, URL <https://www.schrodinger.com>.
- [39] B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff, B. Honig, *J. Phys. Chem.* **1996**, *100* 11775.
- [40] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11* 440.
- [41] J. J. P. Stewart, *J. Comput.-Aided Mol. Des.* **1990**, *4* 1.
- [42] S. M. Bachrach, *J. Am. Chem. Soc.* **2004**, *126* 5018.
- [43] K. Roos, C. J. Wu, W. Damm, M. Reboul, J. M. Stevenson, C. Lu, M. K. Dahlgren, S. Mondal, W. Chen, L. L. Wang, R. Abel, R. A. Friesner, E. D. Harder, *J. Chem. Theory Comput.* **2019**, *15* 1863.
- [44] J. J. P. Stewart, *J. Mol. Model.* **2013**, *19* 1.
- [45] C. W. Coley, L. Rogers, W. H. Green, K. F. Jensen, *J. Chem. Inf. Model.* **2018**, *58* 252.
- [46] Reaxys – An expert-curated chemistry database, URL <https://www.elsevier.com/solutions/reaxys>.
- [47] S. C. Liu, W. L. Nie, C. P. Wang, J. R. Lu, Z. R. Qiao, L. Liu, J. Tang, C. W. Xiao, A. Anandkumar, *Nat. Mach. Intell.* **2023**, *5*, 12 1447.

- [48] K. J. Bowers, D. E. Chow, H. F. Xu, R. O. Dror, M. P. Eastwood, B. A. Gregersen, J. L. Klepeis, I. Kolossvary, M. A. Moraes, F. D. Sacerdoti, J. K. Salmon, Y. B. Shan, D. E. Shaw, *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing* **2006**, 84–es.
- [49] X. Zhou, R. A. Janssen, S. Er, *Energy adv.* **2023**, *2* 820.