Multi-objective goal-directed optimization of de novo stable organic radicals for aqueous redox flow batteries

Shree Sowndarya S. V.[1], Jeffrey Law[2], Charles Tripp[3], Dmitry Duplyakin[3], Erotokritos Skordilis[3†], David Biagioni[3*], Robert S. Paton[1*], Peter C. St. John[2*]

[1] Department of Chemistry, Colorado State University, Fort Collins CO 80523, USA

[2] Biosciences Center, National Renewable Energy Laboratory, Golden CO 80401, USA

[3] Computational Sciences Center, National Renewable Energy Laboratory, Golden CO 80401, USA

ABSTRACT: Advances in the field of goal-directed molecular optimization offer the promise to find feasible candidates for even the most challenging molecular design applications. However, several obstacles remain in applying these tools to practical problems, including lengthy computational or experimental evaluation, synthesizability considerations, and a vast potential search space. As an example of a fundamental design challenge with industrial relevance, we search for novel stable radical scaffolds for an aqueous redox flow battery that simultaneously satisfy redox requirements at the anode and cathode. To meet this challenge, we develop a new open-source molecular optimization framework based on *AlphaZero* coupled with a fast, machine learning-derived surrogate objective trained with nearly 100,000 quantum chemistry simulations. The objective function comprises two graph neural networks: one that predicts adiabatic oxidation and reduction potentials and a second that predicts electron density and local 3D environment, previously shown to be correlated with radical persistence and stability. With no hand-coded knowledge of organic chemistry, the reinforcement learning agent finds molecule candidates that satisfy a precise combination of redox, stability, and synthesizability requirements defined at the quantum chemistry level, many of which have reasonable predicted retrosynthetic pathways. The optimized molecules show that alternative stable radical scaffolds may offer a unique profile of stability and redox potentials to enable low-cost symmetric aqueous redox flow batteries.

INTRODUCTION

The development of materials with precisely tuned electrochemical and physical properties is critical in enabling next-generation energy technologies. One example appears in redox flow batteries (RFBs), which offer the potential to deliver a low-cost and reliable energy storage solution at the grid-scale.[1] While vanadium-based RFB chemistries are currently the most well-studied, battery formulations using organic molecules as the active species are a promising alternative as they are domestically manufacturable, decoupled from markets for transition metals, and have a lesser ecological footprint.[2–4] The stability window of the solvent determines the desired electrochemical potential for both anode and cathode half-reactions: for water-based batteries, this dictates a maximum thermodynamically stable voltage of 1.23 V at 25°C, although in practice voltages of approximately 2.0 V are possible due to the slow kinetics of hydrogen evolution on carbon electrodes.[5]

A wide range of organic redox couples exist and have been explored as charge carriers in flow battery applications.[6] Among these, persistent organic radicals are a promising class of active species with highly reversible redox processes.[7] These molecules have an unpaired valence electron that can either be donated or paired with an accepted electron to form a closed-shell species. However, due partly to their unique and complex chemistry, relatively few stable radical-containing materials have been explored.[7,8] As a result, most studies have focused on chemical modifications of a handful of well-known stable radical scaffolds,[9,10] primarily via mechanism-based approaches that identify optimal side-chains to improve performance such as increasing solubility or limiting possible decomposition reactions.[11–18] The scarcity of radical scaffolds complicates the tuning of their physical and electrochemical properties to meet the strict demands of high-performance, low-cost RFBs.[2,3] For example, TEMPO (2,2,6,6-tetramethylpiperidine-N-oxyl) is currently one of the best organic catholyte candidates due to its persistence and ability to undergo reversible one-election oxidation (Fig. 1). However, water-soluble TEMPO derivatives remain uneconomical, with a low oxidation potential of +0.8 V vs. Standard Hydrogen Electrode (SHE) compared to the thermodynamic limit of +1.23 V in water imposed by the oxygen evolution half-reaction.[19,20] Viologen derivatives have similarly been explored as anolyte materials due to their highly reversible +1/+2 redox couple with a standard reduction potential of -0.45 V vs. SHE.[16,21] However, with a molecular weight of 257 g/mol (vs. 156 g/mol for TEMPO), batteries based on methyl viologen are unlikely to satisfy capital cost requirements for commercial viability.[3] Additionally, the use of separate electrolytes for the anode and cathode results in capacity fade with chemical cross-over driven by concentration gradients.[22] The discovery of new stable organic radical scaffolds may therefore unlock performance and cost targets unachievable with current materials.

Recent work has demonstrated that the stability of organic radicals, viewed in terms of thermodynamic stabilization and kinetic persistence, can be estimated using density functional theory (DFT).[23] In addition to stability, the electron exchange half-reactions of RFB candidates must have a precise redox potential for optimal performance. For organic radicals, their single-electron half-reaction potentials are determined by their adiabatic electron affinity and ionization energy, which can be reliably estimated via DFT.[10] Computational screening of many requirements for new redox-active moieties is therefore feasible, enabling a high-throughput search for potential candidates.

The field of goal-directed molecular optimization has evolved rapidly in recent years, boosted in part by improved machine learning (ML) tools and generative algorithms.[24] Computational lead generation has been predominantly studied in pharmaceutical research, often through generating serialized molecular structures as SMILES strings that resemble a given training database of compounds.[25-27] While algorithmic approaches like Bayesian optimization can help to efficiently navigate predefined chemical spaces in combination with surrogate models, the number of pre-enumerated chemical species is constrained. In contrast, performing on-the-fly molecular generation during exploration enables searching even larger chemical spaces.[24] Techniques from reinforcement learning (RL) have shown an excellent ability to generate valid molecules with desired properties without relying on an existing database of molecular structures to learn valid structural motifs.[28,29] In particular, methods based on a direct tree search of molecular structures using techniques such as Monte Carlo Tree Search (MCTS) offer the ability to precisely control the search space of candidate molecules.[30-34] Most molecular optimization work has been benchmarked on fairly simple and fast-to-compute functions, such as Quantitative Estimate of Druglikeness (QED)[35] or penalized octanol-water partition coefficient (Penalized Log-P)[36] that do not provide a realistic picture of the challenges of molecular design,[37] or optimize just a single objective.[38] Additionally, generative models explore millions of potential candidates during a typical search - precluding the incorporation of more detailed and computationally intensive molecular evaluation criteria using DFT or other high-fidelity simulations.

Machine learning (ML) surrogate models, given sufficient training data, have been shown to reproduce quantum chemical calculations at a fraction of the computational cost.[39-42] In this study, we develop a complex and multi-factored objective function for organic radical charge carriers that includes radical stability, redox potential, and synthesizability considerations backed by $O(10^5)$ DFT calculations. We next implement a scalable RL approach based on single-player AlphaZero[43,44] that guarantees validity and low synthetic accessibility scores[35] for optimized molecules.

As an even more ambitious goal than simply improving on the single redox couple performance of either TEMPO or methyl viologen, we sought to find stable radical candidates that simultaneously satisfy both the oxidation and reduction potential requirements for a symmetric aqueous redox flow battery. Compared to the optimization of an asymmetric battery candidate, this requirement imposes a more complex multi-objective optimization challenge, as the quantum chemical energies of two one-electron redox processes must be balanced within a single small radical scaffold. Further, we performed a "synthetically aware" exploration of chemical space by evaluating synthetic accessibility on the fly and pruning the search tree where it enters synthetically or topologically impractical regions.

The generative model yielded a large distribution of molecules predicted to meet the desired stability criteria while simultaneously having suitable oxidation and reduction potentials. The accuracy of these ML surrogate predictions was then validated against DFT calculations, with many radical candidates passing all criteria at the DFT level. Furthermore, we performed a post hoc analysis of the predicted retrosynthetic routes for the optimized molecules, finding many molecules with reasonable synthetic pathways.[45] This study demonstrates that goal-directed molecular optimization, coupled with a highly detailed ML surrogate model, can produce realistic candidates for demanding applications. Additionally, this study suggests that stable radical scaffolds for RFBs are likely more abundant than the limited but well-known set of experimentally characterized motifs.

RESULTS AND DISCUSSION

Required features of organic active species

We began by defining the features required for organic stable radical active species to be viable candidates for redox flow batteries (Fig. 1). For commercial viability, RFBs need to achieve a high charge density and high reversibility (i.e., longevity) at low cost. A benefit of organic active species is their ability to be incorporated as pendant groups into large macromolecules.[21,46] Polymer-based RFBs allow cheaper size-exclusion membranes to be used and provide an additional lever for tuning the solubility of active materials. Pendant groups must therefore have a precisely tuned redox potential to take full advantage of the solvent's electrochemical stability window, and a highly stable, long-lived radical center to avoid reactions that might reduce the battery's capacity over time.[47] Finding a single redox-active species that can perform both the oxidation and reduction reactions simplifies the battery design and reduces capacity fade through membrane cross-over.[48]

Radical groups must also be synthetically accessible. In this study, we estimate standard redox potentials with adiabatic (i.e., geometry-optimized) ionization potentials and electron affinities obtained from implicitly solvated DFT thermochemistry including vibrational zero-point energy. Further, we estimate radical stability using a recently developed metric that incorporates both thermodynamic and kinetic stabilization of the radical center using 3D structural features and electron spin density obtained via DFT.[23] Highly-delocalized and sterically protected radicals are prioritized by this approach. Synthesizability is considered by
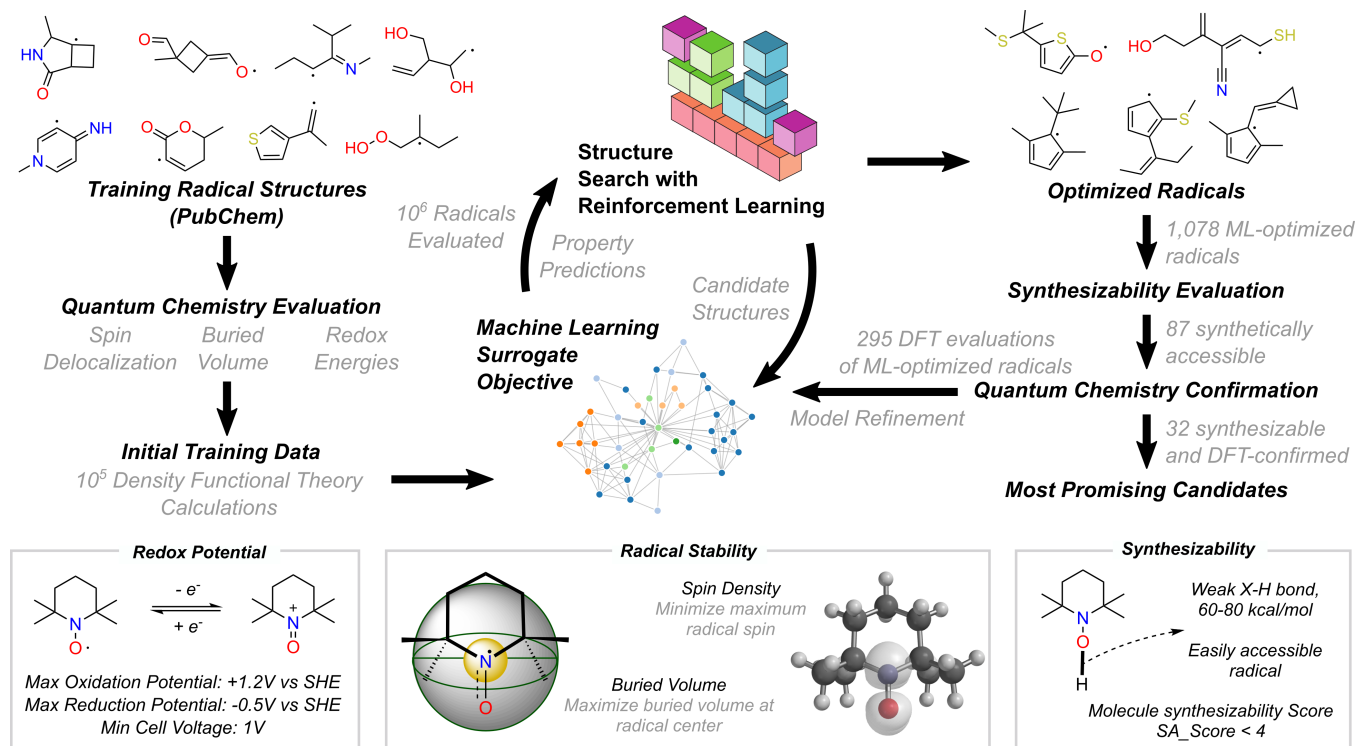
Figure 1. Overview of molecular design criteria used for goal-directed optimization of redox-active stable radical moieties for use in redox flow batteries.

constraining the synthetic accessibility score (SAscore) of the closed-shell R-H molecule to be less than 4.0[35,49] and by ensuring that the R-H bond is relatively weak, with a homolytic bond dissociation enthalpy (BDE) of 60-80 kcal/mol.[50,51] While many thermal and photochemical synthetic protocols exist to form radicals from a closed-shell parent organic compound (e.g., deoxygenation, dehalogenation, etc.), this BDE constraint limits our candidates to those that could be generated by a facile and selective late-stage H-atom abstraction.

Development of a fast surrogate multi-objective function

With the criteria for feasible candidates defined, we next determined a suitable computational approach for scoring generated radicals. To ensure the accuracy of aqueous redox calculations, we first benchmarked a wide number of different density functional, basis set, and solvation model combinations on an experimental dataset of 174 redox potentials (Fig. 2A, Fig. S1).[52] The lowest mean absolute error (MAE) was achieved using M06-2X/def2-TZVP[53] and the SMD solvation model.[54] Full DFT estimation of the adiabatic (i.e., including the effects of geometric relaxation and vibrational zero point energy) ionization energy (IE) and electron affinity (EA) for a given radical takes hours per candidate, and requires three separate geometry optimizations to obtain standard-state Gibbs energies of the neutral radical and both anionic and cationic closed-shell species. We further obtain oxidation potential (OP) and reduction potential (RP) values in Volts by referencing the standard-state Gibbs energy changes to the absolute potential of the SHE.[55]

To enable goal-directed molecular optimization, we constructed a database of 50,547 OP and 81,854 RP calculations by re-optimizing radical and charged structures from

an existing database of organic radicals in an implicit water solvent.[56] We impose several quality checks to ensure convergence of the DFT optimization and validity of the resulting energy calculations, including checking for normal termination of the DFT method, ensuring that bonds were not broken or formed during optimization, and that the optimized open-shell molecules have minimal spin contamination (see Methods).

We next trained a graph neural network (GNN) model to predict both OP and RP directly from a radical's chemical connectivity, i.e., only based on atoms and bonds without considering a specific 3D conformation (Fig. 2B).[39] A test set of 2000 radicals was withheld for validation, consisting of 1773 and 1052 converged RP and OP calculations, respectively. Learning curves plot the models' prediction error as a function of database size (Fig 2C) and demonstrate the models continue to benefit from additional data even at the full database limit. Distributions of prediction errors (in volts) for test-set compounds using the entire training dataset are shown in Fig. 2D, with an MAE of 47.4 and 37.4 mV (1.1 and 0.9 kcal/mol) for OP and RP, respectively, close to the 'chemical accuracy' target of 1 kcal/mol.

Using the same chemical connectivity inputs, we trained a second surrogate GNN model on a recently published database of radical stability scores.[23,57] In this dataset, radical stability is correlated with two quantum chemical descriptors: the delocalization of the radical electron's spin, and the buried volume at the location of maximum spin (calculated as the fractional occupancy of a 3.5 Å radius sphere surrounding a target atom).[58] This GNN is trained to predict local aspects of the optimized 3D-geometry along with the quantum mechanical electron density (more precisely, the density difference between α– and β–spin electrons) at each atomic position. Buried volume and spin density are fractional quantities bounded between 0 and 100 percent.
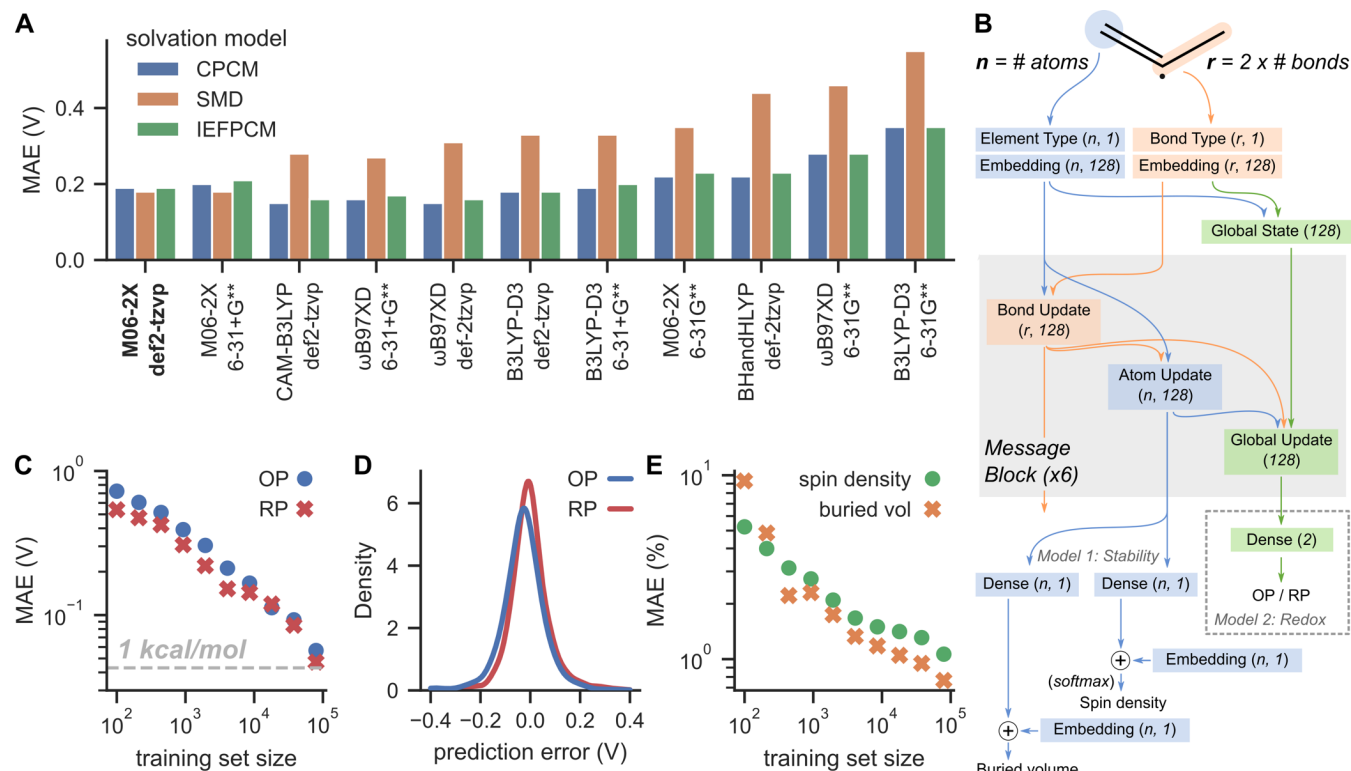
3

Figure 2. Development of a fast surrogate objective function. (A) Prediction accuracy as a function of density functional, basis set, and solvation model on an experimental database of 174 redox potentials.[52] (B) Graph neural network topology for predicting stability and redox potential. Two separate models are trained. The first predicts spin density and buried volume for an input molecule at each atom. The second predicts the oxidation (OP) and reduction (RP) potentials for an entire molecule. (C) Learning curve for redox potentials showing prediction accuracy vs. number of training molecules. (D) Distribution of redox prediction errors for the final trained model. (E) Learning curve of the prediction of parameters governing radical stability.

The model achieves an MAE of 1% in buried volume prediction and 0.7% in predicting QM spin densities on each heavy (i.e., non-hydrogen) atom on 5000 radicals withheld for validation.

Stabilized radicals tend to have highly delocalized electronic structures, where substituents can potentially have a long-range influence on stability and redox potential. As demonstrated by the learning curves in Figs. 2C and 2E, the trained GNN models continue to benefit from additional training data even with nearly 100,000 training examples. The GNNs employed in this study use six message-passing layers and are therefore able to exchange localized chemical information within a radius of 6 bonds. The ability of GNN models to learn long-distance and nonlinear interactions between functional moieties given sufficient training data is an advantage over traditional fingerprint or descriptor-based methods.

These two trained ML models, one for redox potential and one for radical stability, quickly and accurately predict many of the relevant parameters for organic radical viability in RFB applications, thus fulfilling the role of a viable surrogate for DFT calculations. Since RL frameworks typically operate with scalar reward functions, we converted the outputs of these two models into a single reward value as follows. First, we computed radical stability scores by combining the maximum predicted spin and the buried volume at the location of maximum spin. Stability scores range from near zero for highly unstable radicals (i.e., the methyl radical) to 75 or higher for radicals known to be stable experimentally.[23] Second, for the redox potential score, a

maximum of 100 extra points were awarded for meeting each of four separate criteria (25 points each): (i) a reduction potential between -0.5 V and +0.2 V, (ii) an oxidation potential between +0.5 V and +1.2 V, (iii) a total voltage difference of at least 1 V, and (iv) an R-H BDE between 60-80 kcal/mol. BDEs were predicted for the hydrogen-terminated radical using a recently developed ML model.[39] We added these two scores together to obtain a single reward value. Further details on the exact structure of the reward function are provided in the Methods section.

After constructing an efficient surrogate objective function, we next sought to find radicals that maximize this function. Molecule optimization was posed as a search over a directed acyclic graph (DAG), beginning the search at an initial state of a lone carbon atom. In a similar fashion to previous studies, we next considered possible actions to transition between states.[28,29] In this study, each action adds a new bond to the molecule, either between two atoms with free valence in the original molecule (forming a ring) or between an atom in the original molecule and one of a set of possible atom additions. We considered only C, N, O, or S atoms as common elements found organic electronic materials (Fig. 3A). To ensure the molecules we generated were realistic, we refined the set of possible successor states from a given starting structure by (i) enumerating possible stereoisomers, (ii) canonicalizing molecules to tautomer forms,[59] and (iii) removing of molecules with high SAscore values or highly constrained ring systems. Additionally, we removed molecules containing moieties that differed substantially from the redox and stability training database.

Hydrogen atoms were handled implicitly and filled free valence positions in each final molecule. To generate radical structures, additional terminal successor states were created from intermediate molecules where one atom has a hydrogen atom replaced with an unpaired electron. A more complete description of the action space, including a comparison against previous approaches, is given in the Methods section.

Candidate optimization through Reinforcement Learning

In this study, we limited constructed molecules to a maximum of 12 heavy atoms (approximately the size of TEMPO), as lower molecular weight redox-active moieties are preferred for a higher charge-to-mass ratio. Including selecting of the location of the radical electron, this yields a search space of approximately 10^9 possible valid radicals, estimated via extrapolating from smaller maximum sizes and matching previous results.[60] The computational cost of enumerating this grows exponentially with the maximum molecule size, motivating a more efficient strategy for finding top-performing molecules (Fig. S2).

A framework for MCTS optimization over the defined DAGs was implemented that allows for transpositions, where the same molecule can be reached through multiple paths.[61] Following the approach of AlphaZero[43], this framework was augmented with a policy model that replaces the simulation phase (using a random policy) of MCTS with a value score predicted from a GNN, which also initializes the prior scores for successor states from the given molecule. This policy model is trained in a concurrent process by maintaining a buffer of recent MCTS rollouts, sampling in-progress molecules, and minimizing a multi-objective loss function. The loss function contains both the difference between the predicted value score and the final rollout reward and the difference between predicted prior probabilities and the actual search probabilities for each of the molecule's successor nodes (see Methods). As MCTS and the AlphaZero framework were originally designed for competitive games, the ranked reward strategy was used to enable *tabula rasa* self-play for the single-player combinatorial optimization problem.[44] In this strategy, the final reward of a rollout is rescaled to {0, 1} depending on whether the reward is greater than the 75th percentile of the last 250 results. An overview of the connectivity between the rollouts, the data buffer, and the policy model is shown in Fig. 3B. In this fashion, the policy-guided rollouts evolve from an initial random walk over molecular space to a highly targeted exploration of regions likely to contain high-reward molecules.

To search for potential candidate radicals, 200 rollout workers were split across 50 compute nodes for four hours, with a single node equipped with dual Tesla V100 GPUs handling the continual training of the policy model. This approach resulted in a total of 34,626 rollouts and over 3.8 million terminal state radicals evaluated with the surrogate objective function. Fig. 3C plots the final reward from each molecule rollout as a function of time, along with the loss values for training the policy model to predict the final value and prior probabilities for intermediate molecule states. Using ranked rewards to rescale the final reward as a function of recent rollouts means the policy model is forced to continually adapt to predict which intermediate states are the most likely to lead to higher-performing radicals.

By maintaining a cache of all surrogate reward calculations performed during the search, we can easily query the data buffer for the top radical candidates found during the optimization. Of the 3.8 million radicals evaluated, 1,078 had a total surrogate reward greater than 195, corresponding to a minimum stability score of 95. The radical stability metric rewards molecules with highly delocalized electrons and bulky groups offering steric protection of the radical center. As such, the stability metric tends to have a higher maximum value for larger molecules. Known stable radicals in this size range include TEMPO (with a stability score of 93.9) and the phenoxy radical (77.2). The reward function includes a maximum of 100 points for meeting all redox and
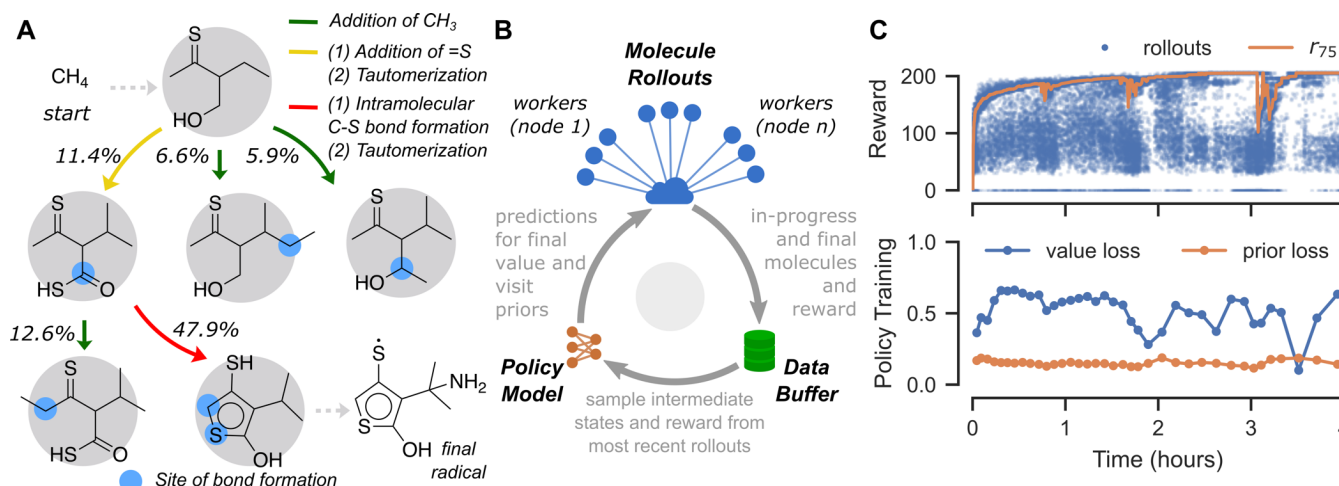

Figure 3. Overview of the Reinforcement Learning (RL) structure optimization strategy. (A) An example of how molecules are constructed through the iterative addition of bonds and atoms. Percentages indicate transition probabilities between states near the end of the RL optimization (not all possible states are shown). (B) Schematic of the architecture of the computational search. Molecule rollouts and policy training are performed asynchronously and coordinated by a data buffer. (C) Evolution of the rewards for individual molecule rollouts (top) and losses for the policy model (bottom) versus time as the optimization proceeds. The 75th percentile of the final reward from the most recent 200 rollouts is denoted as r_{75} (top), and is used to reshape the reward through the ranked rewards strategy.
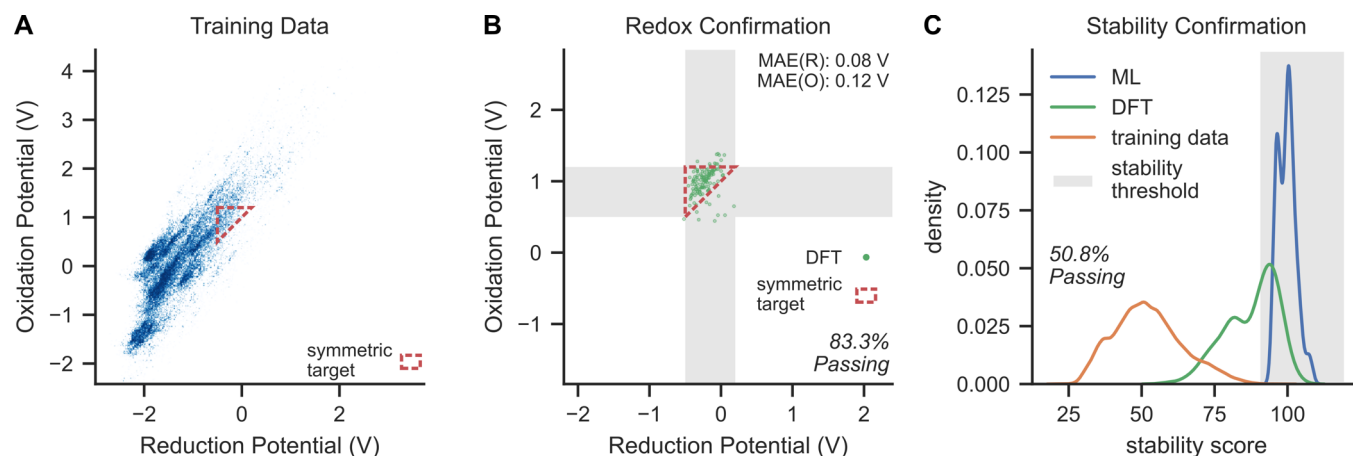
5

Figure 4. Confirmation of top-scoring RL-generated radicals with DFT calculations. (A) Distribution of redox potentials for radicals in the training database, highlighting the target zone for an aqueous, symmetric redox battery (B) DFT computed redox properties for RL optimized radicals that were predicted to fall within the target zone. (C) A comparison of ML (blue) and DFT-derived (green) stability estimates for optimized radicals, compared with the distribution of stability scores for radicals in the training database (orange). A stability threshold of 90 (gray) was used as a lower bound for determining whether a radical could be classified as stable.

bond strength criteria in addition to the radical stability score. From the radical training database, no radicals were found that had a stability score greater than 90 while satisfying the redox criteria.

Confirmation of RL-optimized candidates with DFT

From the 1078 molecules predicted to have the desired properties, a subset of 208 molecules was chosen for subsequent analysis with DFT to verify the accuracy of the ML models. Most top-performing candidates had close to the maximum molecule size. As smaller molecules satisfying the constraints were desired, the top 100 molecules were selected from each size. The subset consisted of 100 radicals containing twelve heavy atoms, 100 containing eleven heavy atoms, six radicals with ten heavy atoms, and two radicals with nine heavy atoms.

In Fig. 4A, we plot the ML-predicted redox voltages for the chosen subset, which all lie within the target triangle to permit a single radical to function as both the electron donor and acceptor in an aqueous redox flow battery with a total voltage of at least 1V. For radicals for which the DFT calculations converged, 83.3% fell within the desired target region (Fig. 4B). The stability scores of the radicals predicted via machine learning were then checked against those obtained via DFT. Fig. 4C shows the distribution of stability scores for both approaches and that stability scores obtained via DFT tended to be lower than those predicted with the surrogate objective function. Using a cutoff score of 85, well within the stability scores observed for experimentally known stable species, 63.3% of radicals were still classified as stable. As shown in Fig. S3, while buried volume predictions for optimized radicals were highly consistent with those obtained from DFT, accurate prediction of spin density was more difficult for these highly delocalized radicals. Additional training data in this region of molecular space may improve accuracy in subsequent experiments, as the generated radicals tended to be much more stable than those found in the training data (Fig. 4C). Evaluation of the synthesizability of generated molecules

The synthesizability of molecules proposed by generative algorithms has been identified as an area of concern, as theoretically optimized molecules that cannot be experimentally tested are of limited practical value.[49] To address this concern, the ASKCOS retrosynthesis prediction web service was applied post hoc to evaluate the 1,078 top-ranked candidates in addition to on-the-fly constraints imposed during optimization to prioritize the search of synthetically tractable space (see Methods).[45,62] Of these, 87 returned putative synthetic routes with a median of 5 synthetic routes per candidate and an average depth of 7.9 steps.

We performed DFT confirmation for all high-scoring molecules found to be synthetically accessible. A total of 32 molecules were confirmed to satisfy the redox requirements while having high stability (>90). Chemical structures for a representative subset of these molecules are depicted in Fig. 5A. The RL-optimized molecules show structural variability through the varied inclusion of N, S, and O heteroatoms and extended delocalized structures, frequently with unsaturated carbo- and heterocyclic (e.g., cyclopentadienyl, pyrrole, furan, thiophene) cores. As required by the objective function, all radicals demonstrate high spin delocalization and high steric protection of the site of highest spin density. In Fig. 5B, we visually compare spin delocalization and buried volumes for both experimentally known and RL-optimized radicals. As expected, a high predicted stability is achieved by delocalizing the radical electron density across multiple atoms and centering the location of highest spin density on an atom with a high buried volume. We note that in Fig. 5A-B, the surrogate model correctly predicts that the spin is predominantly focused at the location of highest buried volume, matching DFT results, even though the radical center is formally specified at a different atom in the SMILES string.

We next investigated predicted retrosynthetic pathways by which the wide variety of top-performing radicals might be experimentally prepared. In Fig. 5C, we show a putative pathway from ASKCOS for the hydrogenated form of a thiophene-based radical. Thiophenes are well-known fragments in organic electronics, where their semi-conducting
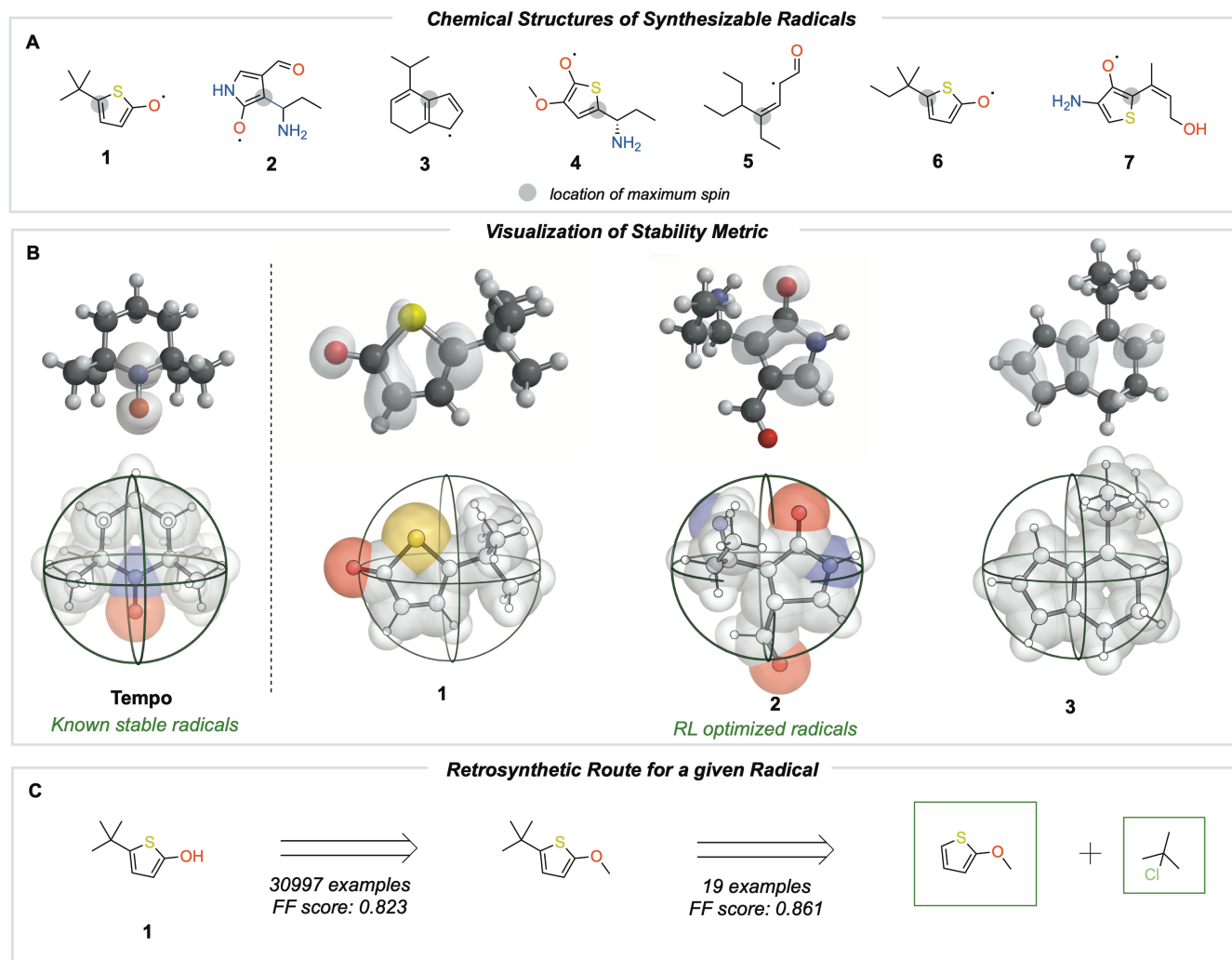
6

Figure 5. De novo structures generated by RL. (A) Radical structures passing the design criteria subsequently validated with DFT. (B) Visualization of the radical stability metric, consisting of both spin delocalization (top row) and buried volume at the center of maximum spin (bottom row) for known stable radicals (left) and those generated via the RL algorithm (right). (C) A possible retrosynthetic pathway for the hydrogenated form of radical **1** generated by ASKCOS. FF score is the estimated plausibility value for each reaction.

properties are exploited for high efficiency.[63] The retrosynthetic route consists of a minimum of two well-established transformations involving a Friedel-Crafts alkylation and an acidic methyl ether cleavage, starting from commercially available 2-methoxythiophene and *tert*-Butyl chloride.[64–66]

Error analysis of the surrogate objective function

The surrogate objective function successfully guided molecule optimization towards regions meeting the desired criteria at the DFT level. However, approximately half of the radicals predicted to meet the desired criteria ultimately fell short upon DFT confirmation. Understanding the primary modes by which the surrogate objective fails will help better understand the limitations of machine-learning guided molecular design and further improve predictive methods for subsequent rounds of optimization.

In Fig. 6A, we show optimized radicals with a substantially lower DFT-calculated stability than that predicted with the surrogate model. A significant reason for such failure is the incorrect prediction of the maximum spin location, with a higher fraction of spin residing on an atom that is not highly shielded by bulky substituents. This failure partly represents a weakness of the chosen stability metric, as minor differences in predicted resonance form can often lead to large swings in the combined score. However, erroneously predicted loci of maximum spin were chemically reasonable, generally corresponding to the location of the second highest DFT spin density. Extended conjugated thiocarbonyl-based radicals are encountered frequently in these outliers. With DFT relaxation, the maximum spin typically locates on the terminal S atom, while the surrogate objective model predicts greater spin at $\alpha-$ or $\gamma-$ positions, in accordance with the general principle of vinylogy.[67] Retraining the surrogate objective with additional examples of these systems may improve predictions in subsequent generation rounds.

Errors in redox predictions also tended to occur for functional groups absent from the training data. In Fig. 6B we show the structure of one such outlier. Using the embeddings assigned by the surrogate model's penultimate prediction layer, we can explore which training set molecules are closest in structure to the target prediction. A nearest-

neighbors search on this latent space reveals several cyclopentadienyl radicals with calculated redox potentials close
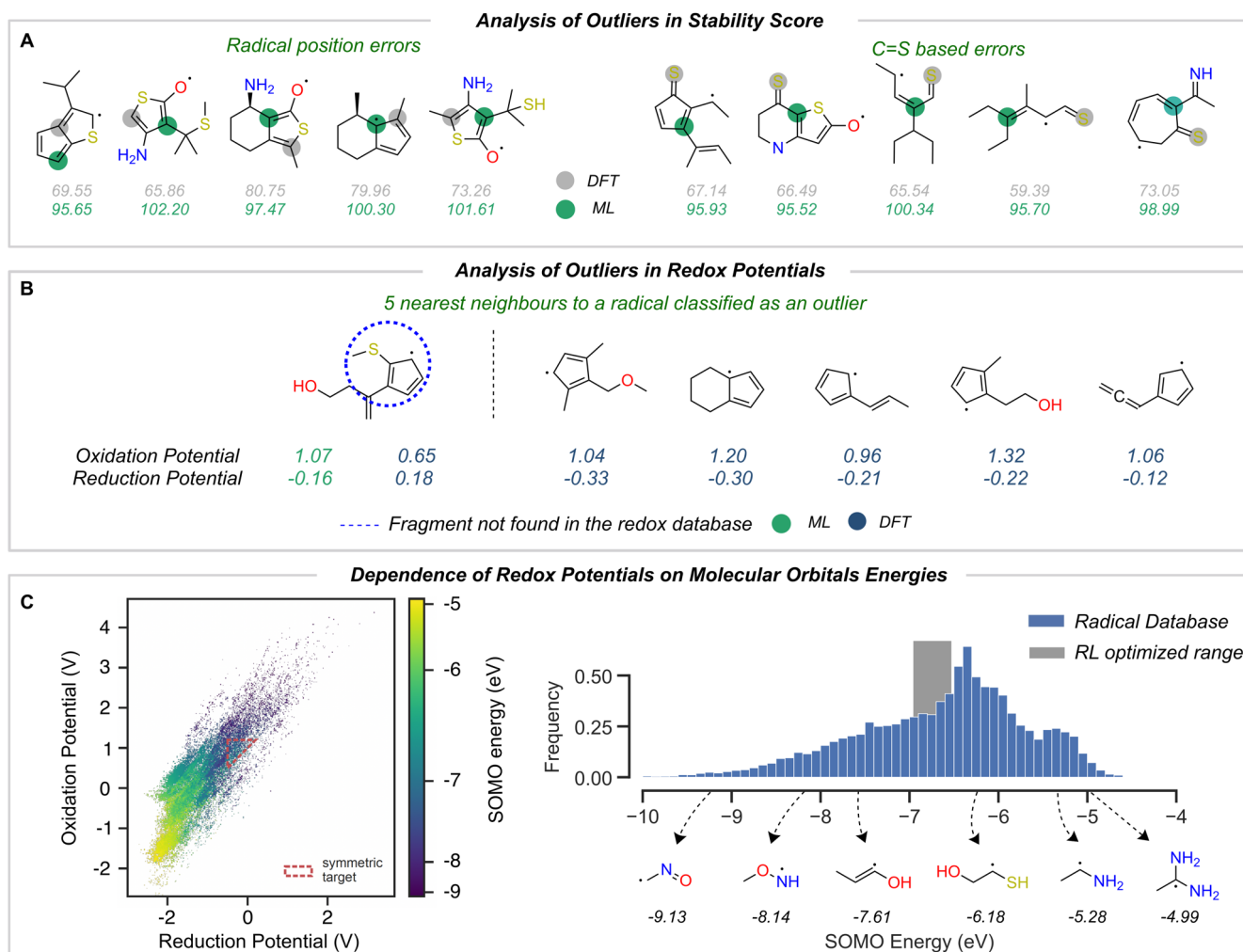

Figure 6. Sources of error in the ML surrogate model and strategies for tuning redox potential. (A) Outliers in ML-predicted stability scores relative to DFT values. The predicted location of maximum spin is highlighted for both methods, and the resulting stability score is shown in gray for DFT and green for ML. (B) An example prediction error for redox potential due to lack of similar molecules present in the training database. The input radical is shown on the left, with the five closest training set radicals shown on the right. (C) On the left, the distribution in redox potentials for all training set molecules is shaded by each radical's SOMO energy. On the right, the distribution of SOMO energies of radicals in the training database is shown, with the gray region extending from the 5th to 95th percentiles of the range observed in RL-optimized candidates. Example structures from the radical database are shown for various SOMO energies.

to the erroneous prediction. The thioether substituent on a cyclopentadienyl, which is not found in any of the molecules in the redox potentials training database, strongly influences redox behavior in a way not captured by the surrogate model. The sulfur atom provides additional stabilization of the oxidized form through resonance and of the reduced form through inductive effects (i.e., stabilization of the α-anion resonance structure). These types of prediction outliers could be remedied by augmenting the redox potential database with additional structural diversity. Learned strategies for precisely tuning redox potential

Searching for a symmetric electrolyte candidate places a challenging constraint on the electronic properties of optimized molecules, as both the oxidation and reduction potentials must be precisely and independently tuned. To explore the strategies used by the RL algorithm, we plot the relationship between oxidation potential (derived from the

ionization energy) and the reduction potential (derived from the electron affinity) in Fig. 6C. Electron-rich, more readily oxidized (e.g., planar aminal) radicals are found in the lower-left corner, while electron-deficient, more readily reduced (e.g., heterocyclic sp²) radicals are found in the top right corner of this plot (Fig. S4).

For open-shell molecules studied with spin-unrestricted Kohn–Sham DFT, the analog to Koopmans' theorem relates the energy of the highest singly-occupied molecular orbital (SOMO) with the vertical ionization energy.[68] Further, the lowest unoccupied molecular orbital (LUMO) is linked closely to vertical electron affinity, mainly when using long-range corrected density functionals.[69] From our computations we observe a correlation between a radical's SOMO energy and both redox potentials (Fig. 6C). This interdependence of the SOMO energy and both redox potentials illustrates the challenge of independently tuning the anode

and cathode half-reactions. Qualitatively, electron-poor (low SOMO energy, electrophilic) radicals tend to be easily reduced and difficult to oxidize, while electron-rich (high SOMO energy, nucleophilic) radicals are easily oxidized and hard to reduce.

However, the RL algorithm still manages to find candidates that meet both redox criteria. Radicals with the required redox properties for aqueous batteries have a SOMO energy in the range of -6.5 to -7 eV (gray region in Fig. 6C). To independently optimize oxidation and reduction potentials at a fixed SOMO energy, AlphaZero learns to harness captodative stabilization of the radical center.[70] Captodative stabilization involves the incorporation of conjugated electron-donating and electron-withdrawing groups, and provides enhanced stability to all three important redox states: the radical, oxidized, and reduced states. Interestingly, this strategy mirrors the use of bipolar redox-active molecules (BRMs), an emerging strategy in the development of non-aqueous RFBs such as 2-phenyl-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide (PTIO).[71] The algorithm thus rediscovers a fundamental concept in radical chemistry that has shown promise in the development of symmetric RFBs. However, unlike existing BRMs with relatively bulky functional groups, those discovered by RL more efficiently blend all required functionality into a much lower molecular weight moiety.

CONCLUSION

In this study, we have performed a search for molecular structures that simultaneously satisfy several complex quantum chemical phenomena important in advanced energy applications. We have demonstrated that combining high-fidelity quantum chemistry simulations, machine learning predictive models, and state-of-the-art reinforcement learning strategies is an effective tool in efficiently exploring molecular space. Without being explicitly programmed on how to construct resonantly stabilized radicals with appropriate orbital energies, the RL algorithm learns a range of strategies that lead to high-performance candidates. As the optimization criteria in molecular design challenges grow more complex, the need for efficient search algorithms also grows.

The construction and optimization of a surrogate model for an otherwise costly simulation or experiment is widely used in many fields, including molecular design. While in this study, candidates were found with reasonable efficiency (50% of optimized radicals), iterative refinement of the surrogate model with respect to the ground-truth calculations would improve the model's accuracy. Additionally, while molecules with putative synthetic routes were found from among the top-performing candidates, more accurate and faster methods of searching synthetically accessible space are required. Additional refinement of the top-performing candidates is also required before they are likely to be applicable in aqueous organic redox flow batteries. Optimizing solubility with predictive models[72,73] and including charged moieties in both the training data and action space will be particularly important in achieving high charge density. Addressing these limitations to achieve holistic prediction of improved bipolar redox actives candidates remains a future goal. All software, data, and models generated in this study have been made available as open-source resources (see Methods).

METHODS

Calculation and validity analysis of redox potentials

Gaussian 16[74] was used for all DFT calculations with a default ultra-fine grid for all numerical integration. The primary database of redox potentials was built using the M06-2x/def2-TZVP level of theory by separately optimizing the neutral, oxidized, and reduced radical species. The calculations were performed using the SMD solvation model with a water solvent at 298K.[54] The same initial structures were used for all three calculations and were taken from previous calculations performed in the gas phase.[56] Iodine-based molecules in the experimental redox benchmark were optimized with the LAN2DZ basis set in combination with 6-31G(d,p) and 6-31g+(d,p).

An automated workflow was developed to check optimizations for convergence by ensuring the absence of imaginary vibrational frequencies and that all bond lengths remained within 0.4 Å of the sum of their covalent radii. Additionally, molecules were inspected to see whether new bonds were formed during optimization, as this often led to difficult-to-predict redox potentials. Atom adjacency matrices were used to determine if any two atoms were closer than 1.3 times the sum of their covalent radii, and these molecules were removed from the training dataset. This primarily occurred during oxidation, as 8,566 oxidized molecules, 602 reduced molecules, and 177 neutral radicals were removed from the database in this fashion.

$$\text{Reduction Potential (V)} = \frac{G(R\cdot) - G(R^-)}{F} - E^{\ominus}\left(\frac{H^+}{H_2}\right)_{abs}$$

$$\text{Oxidation Potential (V)} = \frac{G(R^+) - G(R\cdot)}{F} - E^{\ominus}\left(\frac{H^+}{H_2}\right)_{abs}$$

Spin contamination was checked by looking at the expectation value of the total spin, $\langle S^2 \rangle$. Radicals were expected to have an $\langle S^2 \rangle = 0.75$, and a handful of optimizations were discarded where spin contamination resulted in an $\langle S^2 \rangle > 0.8$. Anions and cations were assumed to adopt a closed-shell singlet state, with an $\langle S^2 \rangle \sim 0$. To improve the consistency of the dataset, open-shell anions and cations with $\langle S^2 \rangle > 0.25$ were removed.

Training the surrogate objective models

Two separate machine learning models were developed to predict quantum mechanical properties as a function of a candidate radical's SMILES[75] notation. The first model predicts spin delocalization and buried volume on each heavy atom in the molecule. The second model predicts the radical's oxidation and reduction potential (in V relative to SHE). SMILES strings were first converted to a graph representation using the nfp[76] and RDKit[77] python libraries. Atoms and bonds were classified depending on features determined via RDKit. For atoms, this included their atomic type, chirality, presence in a ring, number of heavy atom neighbors (degree), aromaticity, number of neighboring hydrogens, and presence of a formal radical center. For bonds, this included the atom types of the joined atoms, the bond type (single, double, aromatic), and presence in a ring, and

Z/E stereochemistry (if present). The GNN edges are directional, and therefore two graph edges are created for each bond in the molecule, one pointing from atom A to atom B and another pointing from atom B to atom A. Each model consisted of a GNN with a similar core structure depicted in Fig. 2 and further detailed in Fig. S5. The GNN generates representative embeddings at the atom, bond, and global level by passing the initial features through a series of message blocks.[78] In the stability prediction GNN, the final atom feature vector is reduced to two output predictions for each atom's buried volume and spin density. In the redox GNN, the final global feature vector is reduced to two output predictions for the reduction and oxidation potentials. Both models are trained with a batch size of 128 molecules for 500 epochs over the training data, using the AdamW optimizer with an initial learning rate of 1E-4, decayed by 1E-5 each update step. The weight decay was set to an initial value of 1E-5 (1E-6 for the redox model) and was decayed by 1E-5 each update step.

Data and trained models for both radical stability and redox potential are available at github.com/pstjohn/redox-models. The initial training data for redox potential is provided as SMILES strings with associated potentials (vs SHE) in Volts. For the final set of 32 radicals, optimized 3D coordinates from DFT for the oxidized, radical, and reduced states are provided, as well as the calculated stability and redox potentials.

Details of the reward function

To find radicals that meet all the desired criteria, predictions and desired ranges for multiple properties were synthesized into a scalar reward function. A continuous piecewise linear function was used to convert predictions to a score between 0 and 1, where a 1 was assigned if the predicted was inside the desired range, 0 outside, and a linear transition between the two scores if the prediction was near the boundary (with width equal to one sixth the width of the desired region).

The overall reward function was then composed by summing over individual scores from different properties

$$reward = 50\,(1 - \max s_i) + 100\,BV_j$$
$$+ 25\,\text{window}(R_{pred}, [-.5\,V, 0.2V])$$
$$+ 25\,\text{window}(O_{pred}, [0.5\,V, 1.2V])$$
$$+ 25\,\text{window}(R_{pred} - O_{pred}, [1\,V, 0.2V])$$
$$+ 25\,\text{window}(BDE_{pred}, [60, 80])$$

where s_i represents the predicted fractional spin on atom i, and BV is a vector of predicted buried volumes. The reward function was constructed to place approximately equal weight between the stability score (including spin and buried volume contributions, typically near 100 for highly stable radicals) and the remaining BDE (in kcal/mol) and redox requirements.

Description of the molecular action space

Beginning with the initial state of a single carbon atom (i.e., methane after adding implicit hydrogens), possible actions were enumerated following a series of expansion and filtering steps. First, all possible tautomers of the given starting molecule were considered as possible starting states.[59] From each starting state, a new bond was added between an atom in the molecule and a second atom, either already in the molecule (forming a ring) for an unbonded C, N, O, or S atom. New molecules were generated for every possible atom pair and bond type (single, double, or triple) for which valency rules were satisfied. From this set of all possible next actions, molecules were filtered according to several ring, saturation, and synthetic accessibility criteria,[79] including restricting molecules to a maximum SAScore of 4.0. The action space was then further expanded by enumerating all possible stereochemical configurations of the starting molecule, followed by a reduction to canonical tautomer forms. Next actions were then de-duplicated by SMILES string.

Our action space differs from previously described molecular 'environments' in several ways. Unlike the environment proposed in MolDQN,[29] our approach results in a directed acyclic graph (DAG) over possible molecules by eliminating the possibility to remove atoms and bonds from molecules under construction. This DAG property prevents the search from searching cyclically and guarantees forward progress when building a radical. It also makes learning the value function easier by eliminating conflation and cross-contamination from cyclical paths in the search graph. Our approach is similar in that respect to the generation environment proposed by You *et al.*,[28] where atom and bond additions are guided by a policy network. Unlike in You *et al.*, we do not decompose the action into a node selection and link selection step, and instead only evaluate policy predictions once both the atom and bond type have been chosen. This allows us to easily filter the action space based on valency rules and dramatically reduces the number of invalid molecules constructed by the algorithm. It additionally allows us to easily consider additional modifications of the action space, including stereochemical enumeration (e.g., at tetrahedral carbon atoms and of double bonds), tautomerization, and synthesizability considerations.

Details of the RL algorithm

The RL optimization was performed using the rlmolecule library[81] (github.com/nrel/rlmolecule), which implements the AlphaZero approach for molecule and material design. In this study, the RL agent learned to select from a parametric action space, where the molecular structures resulting from possible next actions were passed through a trainable policy GNN. The policy GNN had a similar structure to that used for redox prediction (Fig. 2B), using only 3 message passing layers and a feature dimension of 64. The policy model was trained with the ADAM optimizer with a learning rate of 1E-3 and a batch size of 32 positions. Within each batch, the policy model is presented with a molecule state and a list of potential next actions from a recently played MCTS rollout. The policy model is trained to simultaneously predict the actual visitation frequency from MCTS, as well as the outcome of the resulting molecule rollout (0 or 1 as scored via ranked rewards). Starting at the root methane state, molecule rollouts consisted of conducting 250 MCTS samples (or for a maximum of 30 seconds) and selecting the subsequent molecule state with probability proportional to the softmax of the visit counts. This procedure is repeated until a terminal state is selected.

The wall time limit was imposed to mitigate the effects of problematic regions of chemical space where the number of possible next actions per molecule, and therefore the time required to enumerate them, vastly outnumbered typical molecules. This was typically encountered with molecules with many possible tautomers and resulted in rollouts being added to the replay buffer that used an outdated version of the policy model.

Communication between the policy network training script and the MCTS rollout workers is handled through a shared filesystem and a PostgreSQL server. Policy checkpoints are previously written to a shared filesystem location, which is checked at the beginning of each rollout by the workers. Final statistics and molecule reward calculations are then written to the shared SQL database. The policy training script in turn selects the 256 most recent rollouts each training epoch, with each epoch consisting of 100 training steps.

Synthesizability prediction

Retrosynthetic routes are predicted using the ASKCOS web interface tool (https://askcos.mit.edu) using the tree builder module. Settings were chosen to match those used in a previous study evaluating the synthesizability of generative models.[49] Specifically, the maximum tree depth was limited to 9 steps, the maximum branching ratio is set to 25, a maximum wall time of each expansion is limited to 60 s, a maximum reagent cost of $100/g, 1000 max templates, and a maximum target probability of 0.999. Employing these settings along with no defined banned chemicals and reactions for the radical 1(Fig. 5A), we obtained a total of 43 routes, containing 90 chemicals and 970 reactions.

ASSOCIATED CONTENT

Supporting Information

Supplemental Figures (PDF)

AUTHOR INFORMATION

Corresponding Author

* To whom correspondence should be addressed. Peter St. John: peter.stjohn@nrel.gov. Robert Paton: Robert.Paton@colostate.edu, David Biagioni: Dave.Biagioni@nrel.gov

Present Addresses

† Dept. of Business Technology, University of Miami, Miami Herbert Business School. 5250 University Dr., Coral Gables FL 33146, USA

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

ACKNOWLEDGMENT

REFERENCES

(1) Ding, Y.; Zhang, C.; Zhang, L.; Zhou, Y.; Yu, G. Molecular Engineering of Organic Electroactive Materials for Redox Flow Batteries. *Chem. Soc. Rev.* **2018**, *47* (1), 69–103. https://doi.org/10.1039/C7CS00569E.

(2) Ha, S.; Gallagher, K. G. Estimating the System Price of Redox Flow Batteries for Grid Storage. *J. Power Sources* **2015**, *296* (c), 122–132.

(3) Darling, R. M.; Gallagher, K. G.; Kowalski, J. A.; Ha, S.; Brushett, F. R. Pathways to Low-Cost Electrochemical Energy Storage: A Comparison of Aqueous and Nonaqueous Flow Batteries. *Energy Environ. Sci.* **2014**, *7* (11), 3459–3477.

(4) Hu, B.; Debruler, C.; Rhodes, Z.; Liu, T. L. Long-Cycling Aqueous Organic Redox Flow Battery (AORFB) toward Sustainable and Safe Energy Storage. *J. Am. Chem. Soc.* **2017**, *139* (3), 1207–1214. https://doi.org/10.1021/jacs.6b10984.

(5) Kühnel, R.-S.; Reber, D.; Battaglia, C. Perspective—Electrochemical Stability of Water-in-Salt Electrolytes. *J. Electrochem. Soc.* **2020**, *167* (7), 070544. https://doi.org/10.1149/1945-7111/ab7c6f.

(6) Kwabi, D. G.; Ji, Y.; Aziz, M. J. Electrolyte Lifetime in Aqueous Organic Redox Flow Batteries: A Critical Review. *Chem. Rev.* **2020**, *120* (14), 6467–6489. https://doi.org/10.1021/acs.chemrev.9b00599.

(7) Wilcox, D. A.; Agarkar, V.; Mukherjee, S.; Boudouris, B. W. Stable Radical Materials for Energy Applications. *Annu. Rev. Chem. Biomol. Eng.* **2018**, *9* (1), 83–103.

(8) Muench, S.; Wild, A.; Friebe, C.; Häupler, B.; Janoschka, T.; Schubert, U. S. Polymer-Based Organic Batteries. *Chem. Rev.* **2016**, *116* (16), 9438–9484. https://doi.org/10.1021/acs.chemrev.6b00070.

(9) Liu, B.; Tang, C. W.; Jiang, H.; Jia, G.; Zhao, T. Carboxyl-Functionalized TEMPO Catholyte Enabling High-Cycling-Stability and High-Energy-Density Aqueous Organic Redox Flow Batteries. *ACS Sustain. Chem. Eng.* **2021**, *9* (18), 6258–6265. https://doi.org/10.1021/acssuschemeng.0c08946.

(10) Lin, K.; Gómez-Bombarelli, R.; Beh, E. S.; Tong, L.; Chen, Q.; Valle, A.; Aspuru-Guzik, A. A.; Aziz, M. J.; Gordon, R. G. A Redox-Flow Battery with an Alloxazine-Based Organic Electrolyte. *Nat. Energy* **2016**, *1* (9), 1–8. https://doi.org/10.1038/nenergy.2016.102.

(11) Wei, X.; Pan, W.; Duan, W.; Hollas, A.; Yang, Z.; Li, B.; Nie, Z.; Liu, J.; Reed, D.; Wang, W.; Sprenkle, V. Materials and Systems for Organic Redox Flow Batteries: Status and Challenges. *ACS Energy Lett.* **2017**, *2* (9), 2187–2204. https://doi.org/10.1021/acsenergylett.7b00650.

(12) Dai, G.; Liu, Y.; Niu, Z.; He, P.; Zhao, Y.; Zhang, X.; Zhou, H. The Design of Quaternary Nitrogen Redox Center for High-Performance Organic Battery Materials. *Matter* **2019**, *1* (4), 945–958. https://doi.org/10.1016/j.matt.2019.05.009.

(13) Zhang, C.; Niu, Z.; Peng, S.; Ding, Y.; Zhang, L.; Guo, X.; Zhao, Y.; Yu, G. Phenothiazine-Based Organic Catholyte for High-Capacity and Long-Life Aqueous Redox Flow Batteries. *Adv. Mater.* **2019**, *31* (24), 1–8. https://doi.org/10.1002/adma.201901052.

(14) Yan, Y.; Robinson, S. G.; Vaid, T. P.; Sigman, M. S.; Sanford, M. S. Simultaneously Enhancing the Redox Potential and Stability of Multi-Redox Organic Catholytes by Incorporating Cyclopropenium Substituents. *J. Am. Chem. Soc.* **2021**, *143* (33), 13450–13459. https://doi.org/10.1021/jacs.1c07237.

(15) Hu, B.; Tang, Y.; Luo, J.; Grove, G.; Guo, Y.; Liu, T. L. Improved Radical Stability of Viologen Anolytes in Aqueous Organic Redox Flow Batteries. *Chem. Commun.* **2018**, *54* (50), 6871–6874. https://doi.org/10.1039/c8cc02336k.

(16) Liu, T.; Wei, X.; Nie, Z.; Sprenkle, V.; Wang, W. A Total Organic Aqueous Redox Flow Battery Employing a Low Cost and Sustainable

Methyl Viologen Anolyte and 4-HO-TEMPO Catholyte. *Adv. Energy Mater.* **2016**, *6* (3). https://doi.org/10.1002/aenm.201501449.

(17) Wang, W.; Luo, Q.; Li, B.; Wei, X.; Li, L.; Yang, Z. Recent Progress in Redox Flow Battery Research and Development. *Adv. Funct. Mater.* **2013**, *23* (8), 970–986. https://doi.org/10.1002/adfm.201200694.

(18) Shrestha, A.; Hendriks, K. H.; Sigman, M. S.; Minteer, S. D.; Sanford, M. S. Realization of an Asymmetric Non-Aqueous Redox Flow Battery through Molecular Design to Minimize Active Species Crossover and Decomposition. *Chem. - A Eur. J.* **2020**, *26* (24), 5369–5373. https://doi.org/10.1002/chem.202000749.

(19) Tian, Y.; Wu, K. H.; Cao, L.; Saputera, W. H.; Amal, R.; Wang, D. W. Unlocking High-Potential Non-Persistent Radical Chemistry for Semi-Aqueous Redox Batteries. *Chem. Commun.* **2019**, *55* (15), 2154–2157. https://doi.org/10.1039/c8cc09304k.

(20) Suo, L.; Borodin, O.; Gao, T.; Olguin, M.; Ho, J.; Fan, X.; Luo, C.; Wang, C.; Xu, K. "Water-in-Salt" Electrolyte Enables High-Voltage Aqueous Lithium-Ion Chemistries. *Science (80-.).* **2015**, *350* (6263), 938–943.

(21) Janoschka, T.; Martin, N.; Martin, U.; Friebe, C.; Morgenstern, S.; Hiller, H.; Hager, M. D.; Schubert, U. S. An Aqueous, Polymer-Based Redox-Flow Battery Using Non-Corrosive, Safe, and Low-Cost Materials. *Nature* **2015**, *527* (7576), 78–81. https://doi.org/10.1038/nature18909.

(22) Potash, R. A.; McKone, J. R.; Conte, S.; Abruña, H. D. On the Benefits of a Symmetric Redox Flow Battery. *J. Electrochem. Soc.* **2016**, *163* (3), A338–A344. https://doi.org/10.1149/2.0971602jes.

(23) Sowndarya, S. S. V.; St. John, P. C.; Paton, R. S. A Quantitative Metric for Organic Radical Stability and Persistence Using Thermodynamic and Kinetic Features. *Chem. Sci.* **2021**. https://doi.org/10.1039/d1sc02770k.

(24) Coley, C. W. Defining and Exploring Chemical Spaces. *Trends Chem.* **2021**, *3* (2), 133–145. https://doi.org/10.1016/j.trechm.2020.11.004.

(25) Segler, M. H. S. S.; Kogej, T.; Tyrchan, C.; Waller, M. P. Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. *ACS Cent. Sci.* **2018**, *4* (1), 120–131. https://doi.org/10.1021/acscentsci.7b00512.

(26) Yang, Y.; Zhang, R.; Li, Z.; Mei, L.; Wan, S.; Ding, H.; Chen, Z.; Xing, J.; Feng, H.; Han, J.; Jiang, H.; Zheng, M.; Luo, C.; Zhou, B. Discovery of Highly Potent, Selective, and Orally Efficacious P300/CBP Histone Acetyltransferases Inhibitors. *J. Med. Chem.* **2020**, *63* (3), 1337–1360. https://doi.org/10.1021/acs.jmedchem.9b01721.

(27) Moret, M.; Helmstädter, M.; Grisoni, F.; Schneider, G.; Merk, D. Beam Search for Automated Design and Scoring of Novel ROR Ligands with Machine Intelligence. *Angew. Chemie Int. Ed.* **2021**, 2–8. https://doi.org/10.1002/anie.202104405.

(28) You, J.; Liu, B.; Ying, R.; Pande, V.; Leskovec, J. Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation. In *NeurIPS*; Cornell University Library, 2018.

(29) Zhou, Z.; Kearnes, S.; Li, L.; Zare, R. N.; Riley, P. Optimization of Molecules via Deep Reinforcement Learning. *Sci. Rep.* **2019**, *9* (1), 10752. https://doi.org/10.1038/s41598-019-47148-x.

(30) Yang, X.; Aasawat, T. K.; Yoshizoe, K. Practical Massively Parallel Monte-Carlo Tree Search Applied to Molecular Design. **2020**, 1–19.

(31) Jensen, J. H. A Graph-Based Genetic Algorithm and Generative Model/Monte Carlo Tree Search for the Exploration of Chemical Space. *Chem. Sci.* **2019**, *10* (12), 3567–3572. https://doi.org/10.1039/c8sc05372c.

(32) Yang, X.; Zhang, J.; Yoshizoe, K.; Terayama, K.; Tsuda, K. ChemTS: An Efficient Python Library for de Novo Molecular Generation. *Sci. Technol. Adv. Mater.* **2017**, *18* (1), 972–976. https://doi.org/10.1080/14686996.2017.1401424.

(33) Rajasekar, A. A.; Raman, K.; Ravindran, B. Goal Directed Molecule Generation Using Monte Carlo Tree Search. **2020**, 1–6.

(34) Kajita, S.; Kinjo, T.; Nishi, T. Autonomous Molecular Design by Monte-Carlo Tree Search and Rapid Evaluations Using Molecular Dynamics Simulations. *Commun. Phys.* **2020**, *3* (1), 1–11. https://doi.org/10.1038/s42005-020-0338-y.

(35) Ertl, P.; Schuffenhauer, A. Estimation of Synthetic Accessibility Score of Drug-like Molecules Based on Molecular Complexity and Fragment Contributions. *J. Cheminform.* **2009**, *1* (1), 8. https://doi.org/10.1186/1758-2946-1-8.

(36) Kusner, M. J.; Paige, B.; Hernández-Lobato, J. M. Grammar Variational Autoencoder. *34th Int. Conf. Mach. Learn. ICML 2017* **2017**, *4*, 3072–3084.

(37) Brown, N.; Fiscato, M.; Segler, M. H. S.; Vaucher, A. C. GuacaMol: Benchmarking Models for de Novo Molecular Design. *J. Chem. Inf. Model.* **2019**, *59* (3), 1096–1108. https://doi.org/10.1021/acs.jcim.8b00839.

(38) Sumita, M.; Yang, X.; Ishihara, S.; Tamura, R.; Tsuda, K. Hunting for Organic Molecules with Artificial Intelligence: Molecules Optimized for Desired Excitation Energies. *ACS Cent. Sci.* **2018**, *4* (9), 1126–1133. https://doi.org/10.1021/acscentsci.8b00213.

(39) St. John, P. C.; Guan, Y.; Kim, Y.; Kim, S.; Paton, R. S. Prediction of Organic Homolytic Bond Dissociation Enthalpies at near Chemical Accuracy with Sub-Second Computational Cost. *Nat. Commun.* **2020**, *11* (1), 1–12. https://doi.org/10.1038/s41467-020-16201-z.

(40) Guan, Y.; Shree Sowndarya, S. V.; Gallegos, L. C.; St. John, P. C.; Paton, R. S. Real-Time Prediction of 1 H and 13 C Chemical Shifts with DFT Accuracy Using a 3D Graph Neural Network. *Chem. Sci.* **2021**, *12* (36), 12012–12026. https://doi.org/10.1039/D1SC03343C.

(41) Tabor, D. P.; Gómez-Bombarelli, R.; Tong, L.; Gordon, R. G.; Aziz, M. J.; Aspuru-Guzik, A. Mapping the Frontiers of Quinone Stability in Aqueous Media: Implications for Organic Aqueous Redox Flow Batteries. *J. Mater. Chem. A* **2019**, *7* (20), 12833–12841. https://doi.org/10.1039/c9ta03219c.

(42) Jinich, A.; Sanchez-Lengeling, B.; Ren, H.; Harman, R.; Aspuru-Guzik, A. A Mixed Quantum Chemistry/Machine Learning Approach for the Fast and Accurate Prediction of Biochemical Redox Potentials and Its Large-Scale Application to 315 »000 Redox Reactions. *ACS Cent. Sci.* **2019**, *5* (7), 1199–1210. https://doi.org/10.1021/acscentsci.9b00297.

(43) Silver, D.; Hubert, T.; Schrittwieser, J.; Antonoglou, I.; Lai, M.; Guez, A.; Lanctot, M.; Sifre, L.; Kumaran, D.; Graepel, T.; Lillicrap, T.; Simonyan, K.; Hassabis, D. A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go through Self-Play. *Science (80-.).* **2018**, *362* (6419), 1140–1144. https://doi.org/10.1126/science.aar6404.

(44) Laterre, A.; Fu, Y.; Jabri, M. K.; Cohen, A.-S.; Kas, D.; Hajjar, K.; Dahl, T. S.; Kerkeni, A.; Beguir, K. Ranked Reward: Enabling Self-Play Reinforcement Learning for Combinatorial Optimization. **2018**.

(45) Coley, C. W.; Rogers, L.; Green, W. H.; Jensen, K. F. Computer-Assisted Retrosynthesis Based on Molecular Similarity. *ACS Cent. Sci.* **2017**, *3* (12), 1237–1245. https://doi.org/10.1021/acscentsci.7b00355.

(46) Nguyen, T. P.; Easley, A. D.; Kang, N.; Khan, S.; Lim, S.-M.; Rezenom, Y. H.; Wang, S.; Tran, D. K.; Fan, J.; Letteri, R. A.; He, X.; Su, L.; Yu, C.-H.; Lutkenhaus, J. L.; Wooley, K. L. Polypeptide Organic Radical Batteries. *Nature* **2021**, *593* (7857), 61–66. https://doi.org/10.1038/s41586-021-03399-1.

(47) Sevov, C. S.; Hickey, D. P.; Cook, M. E.; Robinson, S. G.; Barnett, S.; Minteer, S. D.; Sigman, M. S.; Sanford, M. S. Physical Organic Approach to Persistent, Cyclable, Low-Potential Electrolytes for Flow Battery Applications. *J. Am. Chem. Soc.* **2017**, *139* (8), 2924–2927. https://doi.org/10.1021/jacs.7b00147.

(48) Tong, L.; Jing, Y.; Gordon, R. G.; Aziz, M. J. Symmetric All-Quinone Aqueous Battery. *ACS Appl. Energy Mater.* **2019**, *2* (6), 4016–4021. https://doi.org/10.1021/acsaem.9b00691.

(49) Gao, W.; Coley, C. W. The Synthesizability of Molecules Proposed by Generative Models. *J. Chem. Inf. Model.* **2020**, *60* (12). https://doi.org/10.1021/acs.jcim.0c00174.

(50) Henry, D. J.; Parkinson, C. J.; Mayer, P. M.; Radom, L. Bond Dissociation Energies and Radical Stabilization Energies Associated with Substituted Methyl Radicals. *J. Phys. Chem. A* **2001**, *105* (27), 6750–6756. https://doi.org/10.1021/jp010442c.

(51) Galli, C. Nitroxyl Radicals. In *The Chemistry of Hydroxylamines, Oximes and Hydroxamic Acids*; John Wiley & Sons, Ltd, 2008; pp 705–750. https://doi.org/https://doi.org/10.1002/9780470741962.ch15.

(52) Roth, H. G.; Romero, N. A.; Nicewicz, D. A. Experimental and Calculated Electrochemical Potentials of Common Organic Molecules for Applications to Single-Electron Redox Chemistry. *Synlett* **2016**, *27* (05), 714–723.

(53) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2007**, *120* (1–3), 215–241.

(54) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378–6396. https://doi.org/10.1021/jp810292n.

(55) Hammerich, O.; Speiser, B. Techniques For Studies Of Electrochemical Reactions In Solution. In *Organic Electrochemistry*; CRC Press, 2015; pp 117–188. https://doi.org/10.1201/b19122-8.

(56) St. John, P. C.; Guan, Y.; Kim, Y.; Etz, B. D.; Kim, S.; Paton, R. S. Quantum Chemical Calculations for over 200,000 Organic Radical Species and 40,000 Associated Closed-Shell Molecules. *Sci. Data* **2020**, 7 (1), 244. https://doi.org/10.1038/s41597-020-00588-x.

(57) Sowndarya S. V., S.; St. John, P.; Paton, R. Radical Stability Calculations for 89,320 Organic Radicals in the Water Phase. figshare May 2021. https://doi.org/10.6084/m9.figshare.14597556.v2.

(58) Cavallo, L.; Correa, A.; Costabile, C.; Jacobsen, H. Steric and Electronic Effects in the Bonding of N-Heterocyclic Ligands to Transition Metals. *J. Organomet. Chem.* **2005**, *690* (24–25), 5407–5413. https://doi.org/10.1016/j.jorganchem.2005.07.012.

(59) Sitzmann, M.; Ihlenfeldt, W. D.; Nicklaus, M. C. Tautomerism in Large Databases. *J. Comput. Aided. Mol. Des.* **2010**, *24* (6–7), 521–551. https://doi.org/10.1007/s10822-010-9346-4.

(60) Blum, L. C.; Reymond, J. L. 970 Million Druglike Small Molecules for Virtual Screening in the Chemical Universe Database GDB-13. *J. Am. Chem. Soc.* **2009**, *131* (25), 8732–8733. https://doi.org/10.1021/ja902302h.

(61) Saffidine, A.; Cazenave, T.; Méhat, J. UCD: Upper Confidence Bound for Rooted Directed Acyclic Graphs. *Knowledge-Based Syst.* **2012**, *34*, 26–33. https://doi.org/10.1016/j.knosys.2011.11.014.

(62) Coley, C. W.; Barzilay, R.; Jaakkola, T. S.; Green, W. H.; Jensen, K. F. Prediction of Organic Reaction Outcomes Using Machine Learning. *ACS Cent. Sci.* **2017**, *3* (5), 434–443. https://doi.org/10.1021/acscentsci.7b00064.

(63) Dou, L.; Liu, Y.; Hong, Z.; Li, G.; Yang, Y. Low-Bandgap Near-IR Conjugated Polymers/Molecules for Organic Electronics. *Chem. Rev.* **2015**, *115* (23), 12633–12665. https://doi.org/10.1021/acs.chemrev.5b00165.

(64) Belen'Kii, L. I.; Yakubov, A. P. Stable Heteroareniumions - VIII Some Transformations of Alkylthiophenium Ions and New Synthesis of 2-t-Butylthiophene. *Tetrahedron* **1984**, *40* (13), 2471–2477. https://doi.org/10.1016/S0040-4020(01)83499-2.

(65) 2-Methoxythiophene | Sigma-Aldrich https://www.sigmaaldrich.com/US/en/product/aldrich/331597 (accessed Mar 12, 2021).

(66) 2-Chloro-2-methylpropane | Sigma-Aldrich.

(67) Curti, C.; Battistini, L.; Sartori, A.; Zanardi, F. New Developments of the Principle of Vinylogy as Applied to π-Extended Enolate-Type Donor Systems. *Chem. Rev.* **2020**, *120* (5), 2448–2612. https://doi.org/10.1021/acs.chemrev.9b00481.

(68) Gritsenko, O. V.; Baerends, E. J. The Spin-Unrestricted Molecular Kohn-Sham Solution and the Analogue of Koopmans's Theorem for Open-Shell Molecules. *J. Chem. Phys.* **2004**, *120* (18), 8364–8372. https://doi.org/10.1063/1.1698561.

(69) Tsuneda, T.; Song, J. W.; Suzuki, S.; Hirao, K. On Koopmans' Theorem in Density Functional Theory. *J. Chem. Phys.* **2010**, *133* (17). https://doi.org/10.1063/1.3491272.

(70) Bordwell, F. G.; Lynch, T. Y. Radical Stabilization Energies and Synergistic (Captodative) Effects. *J. Am. Chem. Soc.* **1989**, *111* (19), 7558–7562. https://doi.org/10.1021/ja00201a043.

(71) Li, M.; Case, J.; Minteer, S. D. Bipolar Redox-Active Molecules in Non-Aqueous Organic Redox Flow Batteries: Status and Challenges. *ChemElectroChem* **2021**, *8* (7), 1215–1232. https://doi.org/10.1002/celc.202001584.

(72) Vermeire, F. H.; Green, W. H. Transfer Learning for Solvation Free Energies: From Quantum Chemistry to Experiments. *Chem. Eng. J.* **2021**, *418* (March), 129307. https://doi.org/10.1016/j.cej.2021.129307.

(73) Alibakhshi, A.; Hartke, B. Improved Prediction of Solvation Free Energies by Machine-Learning Polarizable Continuum Solvation Model. *Nat. Commun.* **2021**, *12* (1), 1–7. https://doi.org/10.1038/s41467-021-23724-6.

(74) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Rev. C.01. *Gaussian 16* **2016**.

(75) Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Model.* **1988**, *28* (1), 31–36.

(76) St. John, P. Nfp: Keras Layers for End-to-End Learning on Molecular Structure. 2019. https://doi.org/10.11578/dc.20190731.1.

(77) Landrum, G. A. RDKit: Open-source cheminformatics http://www.rdkit.org.

(78) Gallegos, L. C.; Luchini, G.; St. John, P. C.; Kim, S.; Paton, R. S. Importance of Engineered and Learned Molecular Representations in Predicting Organic Reactivity, Selectivity, and Chemical Properties. *Acc. Chem. Res.* **2021**. https://doi.org/10.1021/acs.accounts.0c00745.

(79) Ruddigkeit, L.; Van Deursen, R.; Blum, L. C.; Reymond, J. L. Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *J. Chem. Inf. Model.* **2012**, *52* (11), 2864–2875. https://doi.org/10.1021/ci300415d.

(80) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, *55* (12), 2562–2574.

(81) Biagioni, D.; Skordilis, E.; Tripp, C.; Duplyakin, D.; St. John, P. Rlmolecule: A Library for General-Purpose Material and Molecular Optimization Using AlphaZero-Style Reinforcement Learning. 2020. https://doi.org/10.11578/dc.20201221.3.