

Cite this: *Chem. Sci.*, 2023, 14, 13734 All publication charges for this article have been paid for by the Royal Society of ChemistryReceived 5th August 2023  
Accepted 1st November 2023

DOI: 10.1039/d3sc04084d

rsc.li/chemical-science

# Data science enabled discovery of a highly soluble 2,2'-bipyrimidine anolyte for application in a flow battery†

Adam R. Pancoast,<sup>ac</sup> Sara L. McCormack,<sup>ac</sup> Shelby Galinat,<sup>a</sup> Ryan Walser-Kuntz,<sup>bc</sup> Brianna M. Jett,<sup>bc</sup> Melanie S. Sanford<sup>bc</sup> and Matthew S. Sigman<sup>\*,ab</sup>

Development of non-aqueous redox flow batteries as a viable energy storage solution relies upon the identification of soluble charge carriers capable of storing large amounts of energy over extended time periods. A combination of metrics including number of electrons stored per molecule, redox potential, stability, and solubility of the charge carrier impact performance. In this context, we recently reported a 2,2'-bipyrimidine charge carrier that stores two electrons per molecule with reduction near  $-2.0$  V vs. Fc/Fc<sup>+</sup> and high stability. However, these first-generation derivatives showed a modest solubility of 0.17 M (0.34 M e<sup>-</sup>). Seeking to improve solubility without sacrificing stability, we harnessed the synthetic modularity of this scaffold to design a library of sixteen candidates. Using computed molecular descriptors and a single node decision tree, we found that minimization of the solvent accessible surface area (SASA) can be used to predict derivatives with enhanced solubility. This parameter was used in combination with a heatmap describing stability to de-risk a virtual screen that ultimately identified a 2,2'-bipyrimidine with significantly increased solubility and good stability metrics in the reduced states. This molecule was paired with a cyclopropenium catholyte in a prototype all-organic redox flow battery, achieving a cell potential up to 3 V.

## Introduction

Renewable energy sources, such as solar or wind power, have seen a steep decrease in cost over the last 10 years and consequently are envisioned to play an important role in global energy generation. Indeed, the International Renewable Energy Agency (IRENA) has projected that 90% of electricity will be derived from renewable sources by 2050.<sup>1</sup> Critical to realizing this goal is further development of effective energy storage technologies, such as redox flow batteries.<sup>2</sup> Redox flow batteries, which store energy in the form of flowable solutions of oxidized or reduced charge carriers (catholytes and anolytes, respectively), benefit from the separation of the electrolyte reservoirs from the cell stack.<sup>3</sup> This allows energy to be scaled independently from power, creating a high level of flexibility for various applications.<sup>4</sup>

The current state of the art in redox flow battery technology utilizes aqueous solutions of vanadium-based charge carriers. Despite success on a commercial level, vanadium redox flow batteries face several hurdles to broader utilization as energy storage devices, including modest energy densities due to the relatively narrow electrochemical window of water.<sup>4,5</sup> Conversely, non-aqueous solvents like acetonitrile present an attractive route to produce high energy density flow batteries due to the enhanced thermodynamic electrochemical window relative to that of water (6.1 *versus* 1.23 V, respectively).<sup>6</sup> As shown in eqn (1) (where  $C_{ap}$  is the battery limiting-capacity,  $V$  is the cell voltage, and  $\mu$  is the volume factor), increasing cell voltage should lead to improved energy densities, with all other factors being equal.<sup>3</sup>

$$\text{Theoretical energy density (W hr L}^{-1}\text{)} = C_{ap}V/\mu_V \quad (1)$$

The key to harnessing the electrochemical window of non-aqueous solvents is increasing the cell voltage through the development of charge carriers with electron transfer events that occur at extreme potentials.<sup>3</sup> This is difficult due to the propensity of highly energetic oxidized or reduced species to decompose by one or more pathways.<sup>7</sup> This challenge notwithstanding, there has been significant progress towards utilizing the expanded solvent window with highly reducing (negative)

<sup>a</sup>Department of Chemistry, University of Utah, 315 South 1400 East, Salt Lake City, Utah 84112, USA. E-mail: sigman@chem.utah.edu

<sup>b</sup>Department of Chemistry, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, USA

<sup>c</sup>Joint Center for Energy Storage Research, 9700 S. Cass Avenue, Argonne, Illinois 60439, USA

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d3sc04084d>



potentials for anolytes<sup>8</sup> and strongly oxidizing (positive) potentials for catholytes.<sup>9,10</sup>

Beyond stability considerations, another challenge lies in achieving high solubility.<sup>11</sup> Since conventional redox flow batteries store energy in the form of dissolved redox active molecules, energy density is directly correlated to the charge carrier concentration.<sup>3</sup> The relationship between capacity and solubility has prompted efforts to understand the molecular features dictating solubility for flow battery materials,<sup>12,13</sup> as well as the introduction of alternative configurations based on an insoluble charge storing species.<sup>14</sup> There are practical limitations to solubility, as the tradeoff between energy density, solution viscosity, and conductivity can become unfavorable at higher concentrations of active electrolyte in non-aqueous systems.<sup>15</sup> Increasing solution viscosity can decrease mass transfer of both active and supporting electrolytes, ultimately causing decreased conductivity, increased overpotentials, and impeded battery performance.<sup>16</sup>

$$\text{Theoretical capacity (A hr L}^{-1}\text{)} = nCF/3600 \quad (2)$$

Considering the tradeoff of solubility and viscosity, a general target of  $\sim 0.5$  M for charge carriers was proposed by Zhang and coworkers in 2018 as an upper limit for solubility due to negative impacts of mass transfer above this concentration.<sup>15</sup> Shown by eqn (2), where  $n$  is the number of electrons transferred,  $C$  is the concentration of the anolyte/catholyte,  $F$  is Faraday's constant, and 3600 is a conversion factor between coulombs and amp-hours, battery capacity is not solely dependent upon solubility.<sup>3</sup> As capacity is a measure of a battery's ability to store charge, anolytes/catholytes capable of storing multiple electrons per molecule will also increase battery capacity through an increased effective electron concentration while limiting mass-transfer issues present in systems with high molecular concentration.

Within this framework, we recently disclosed a promising class of anolytes based on the 2,2'-bipyrimidine scaffold.<sup>17</sup> These molecules exhibit two reduction events near  $-2.0$  V vs.  $\text{Fc}/\text{Fc}^+$  and are accessed through a modular synthetic route that enables the analysis of structure–function relationships. The original report utilized a combination of statistical modeling and mechanistic studies to identify a derivative with promising stability (Fig. 1). Despite this, early studies on anolyte solubility revealed that the neutral form exhibited a modest solubility ( $<0.2$  M) in acetonitrile. This prompted us to investigate the structure–solubility relationship for 2,2'-bipyrimidines.<sup>12,17,18</sup> Herein, we report the use of a single node decision tree and a heatmap to facilitate the design of an analogue with improved solubility while maintaining good stability. The optimized molecule demonstrates stable cycling in a prototype organic redox flow battery with a cell potential up to 3 V.

## Results and discussion

Our previous studies focused on understanding the structural features imparting stability to bipyrimidine anolytes.<sup>17</sup> This was achieved by incorporating sterically small, minimally polarizing

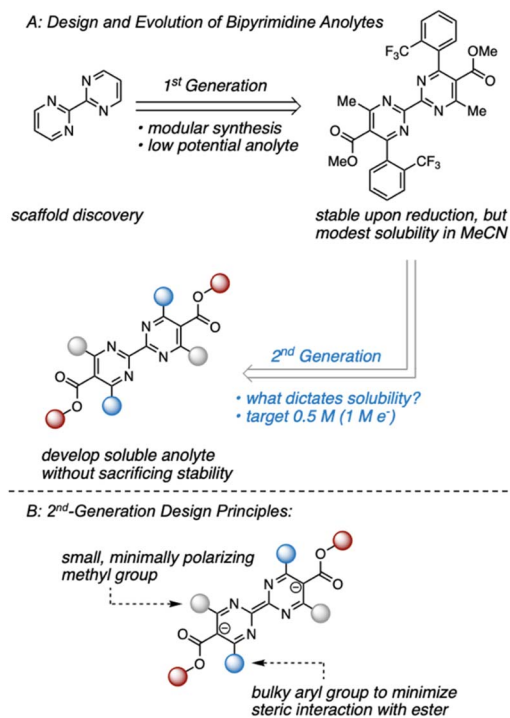


Fig. 1 (A) Evolution of bipyrimidine anolytes. (B) Design principles of 2nd-generation candidates aimed at maintaining a stable reduced state.

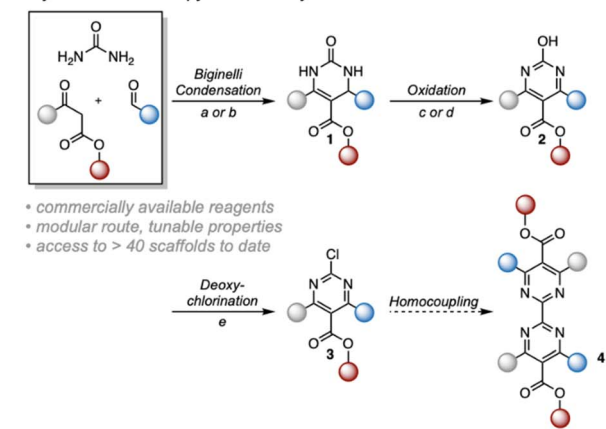
methyl groups at the 4- and 4' positions (grey in Fig. 1 and 2). In concert, *ortho* and *meta* substituted aryl groups were placed at the 6- and 6'-positions (blue in Fig. 1 and 2). These substituted arenes can rotate out of plane and thus minimize steric interactions with the ester. Based on this 1st generation scaffold, we focused on modifying two sites focused to enhance solubility: (1) the aryl groups (blue) and (2) the substituents on the esters (red).

Our previous synthesis commenced with a Biginelli condensation of commercial building blocks, followed by oxidative dehydrogenation and deoxychlorination to form a 2-chloropyrimidine (Fig. 2).<sup>17</sup> These first three steps can be readily performed on  $\geq 100$  mmol scale. However, the next step, a homocoupling of these 2-chloropyrimidine building blocks, was found to give poor yields that are highly dependent on the nature of the substrate. Therefore, to enable the synthesis of a wider range of new compounds for structure–solubility relationship studies, we investigated alternative synthetic approaches to form the bipyrimidine core. Through several model studies, we identified a nucleophilic aromatic substitution strategy based upon recent reports by the McNally group using P(v)-ligand coupling.<sup>19,20</sup> This process operates *via* sequential nucleophilic aromatic substitution and apicophilic ligand coupling,<sup>20</sup> necessitating the use of inherently electrophilic substrates like 2-chloropyrimidines. Gratifyingly, a slightly modified protocol proved to be a general method for the synthesis of the desired bipyrimidine products.

Once library synthesis was completed, each 2,2'-bipyrimidine was characterized with respect to reduction potential,

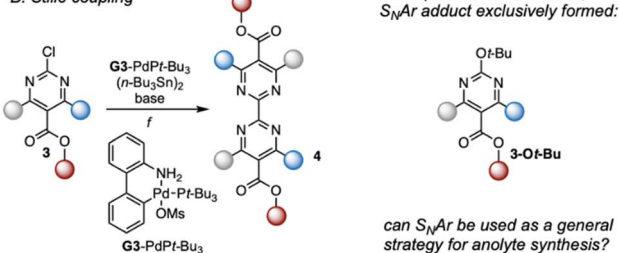


## A: Synthetic Route to Bipyrimidine Anolytes



despite robust nature of first three steps, final homocoupling found to be substrate dependent – necessitated evaluation of alternative coupling routes

## B: Stille coupling



## C: P(V)-ligand coupling

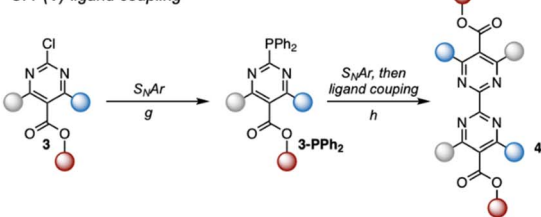


Fig. 2 Synthetic route to 2,2'-bipyrimidines. (A) Synthetic strategy used to access 2,2'-bipyrimidines. For (a) and (b), solvent and ester identity were matched to avoid transesterification. (a) 1 mol% CuCl<sub>2</sub>, 20 mol% H<sub>2</sub>SO<sub>4</sub>, ROH. (b) 0.6 eq. FeCl<sub>3</sub>·6H<sub>2</sub>O, cat. HCl, ROH. (c) 1.2 eq. K<sub>2</sub>S<sub>2</sub>O<sub>8</sub>, 3 : 2 MeCN : H<sub>2</sub>O, Δ. (d) 1 mol% CuCl<sub>2</sub>, 20 mol% K<sub>2</sub>CO<sub>3</sub>, 2.5 eq. *t*-BuOOH, DCM, Δ. (e) 0.5 eq. *N,N*-dimethylaniline, 8.5 eq. POCl<sub>3</sub>, 70 ° C. (f) 1.2 mol% Buchwald G3-PdPt-Bu<sub>3</sub>, 0.5 eq. (*n*-Bu<sub>3</sub>Sn)<sub>2</sub>, 1.4 eq. base, PhMe, Δ. When using K<sub>3</sub>PO<sub>4</sub> as base, no reaction was observed. (g) 1.2 eq. HPPPh<sub>2</sub>, 2,2,2-trifluoroethanol, 75 ° C. (h) 1.2 eqn (2)-chloropyrimidine, 1.2 eq. HOTf, 2.2 eq. NaOTf, 10 eq. H<sub>2</sub>O, 2,2,2-trifluoroethanol, 75 ° C. (B) Evaluation of a Stille coupling strategy. (C) P(V)-Ligand coupling strategy used to access 2nd-generation anolytes.

stability (measured by the temporal rate of fade during H-cell cycling; see ESI† for details), and solubility. Solubility measurements were performed according to a previously reported procedure.<sup>9,12</sup> The values reported in Fig. 3 represent the mean of ≥3 trials conducted on each anolyte. Anolyte **4a** was the most stable analog identified in the first generation but exhibited only modest solubility (0.17 M in neat acetonitrile) in the neutral form. However, in the presence of tetra-*n*-butylammonium (TBA) counterions, the doubly reduced state was found to be a viscous oil that was miscible with acetonitrile (see ESI†). These results indicate that the neutral bipyrimidine is the solubility-limiting form; as such, all solubility measurements

going forward were performed on the neutral structures. Cyclic voltammetry (CV, see ESI† for details) showed that each novel candidate (except for **4k** and **4l** due to their poor solubility in supporting electrolyte solution) undergoes reversible electron transfer in a diffusion-controlled manner.

Variation of the ester led to minimal change in the electrochemistry of **4a–4e** (Fig. 3), while variation of the aryl group had a more pronounced impact on both reduction potential and stability as observed in the first-generation library.<sup>17</sup> The range of anolyte stability (0.66% fade per hour to 5.3% fade per hour) was relatively narrow, consistent with the 2nd generation library design principles. For the candidates with the highest rates of fade (**4j**: 4.0% fade per hour and **4o**: 5.3% fade per hour, respectively), an irreversible peak corresponding to the oxidation of the protonated bipyrimidine (Fig. 4) was observed. This is consistent with the previously reported decomposition pathway, which is promoted by distortion of the reduced anolyte, either through ring puckering or rotation of the ester from the bipyrimidine plane.<sup>17</sup> Comparing **4o** and **4j**, **4o** has a greater distortion relative to **4j** by ~4°, leading to a greater amount of the protonated decomposition product (Fig. 4D). Accordingly, we hypothesized that the stability profiles of the 1st and 2nd generation anolyte libraries could be analyzed simultaneously.

Anolyte solubility was markedly impacted by ester identity, with an approximately six-fold decrease in solubility as the ester is changed from methyl (**4a** in Figure 3) to 2-methoxyethyl (**4e** in Fig. 3). This stands in stark contrast to the reported ability of alkoxyether substituents to enhance solubility in other systems.<sup>21</sup> The impact of the aryl substituent was less clear: changing from an *ortho*-methyl substituted phenyl (**4f**) ring to a 2,6-dimethyl substituted phenyl ring (**4g**) led to a six-fold increase in solubility (Fig. 5). However, no clear trend was evident, and the most soluble candidate, **4g** at 0.3 M (0.6 M e<sup>-</sup>), fell short of the desired target of 0.5 M (1 M e<sup>-</sup>). Notably, only 3 out of 16 2nd generation candidates had solubilities of ≥0.15 M. This provided the impetus to employ statistical modeling techniques capable of deconvoluting the impact of structure on solubility (*vide infra*).

Previously we found that computational analysis of these systems was expedited by using a truncated monoanionic surrogate of the bipyrimidine scaffold, as this strategy greatly reduced the number of accessible conformers.<sup>17</sup> Similarly, we opted to compute a truncated version of the neutral (for solubility modeling) and reduced (monoanionic form for stability modeling) bipyrimidines for statistical modeling efforts (see ESI† for details). Conformational searches were performed with a 5 kcal mol<sup>-1</sup> energetic window and the OPLS4 force field.<sup>22</sup> The ensuing conformational ensembles were then subjected to geometry optimization with DFT using the M06-2x/6-31+G(d,p) level of theory and the implicit conductor-like polarizable continuum model (CPCM) in acetonitrile.<sup>23</sup> Following single point energy calculations with M06-2x/def-TZVP, molecular descriptors such as Sterimol values, partial charges (assessed by natural bond orbital (NBO) calculations), and surface area measurements, were collected for each conformer. In addition to extracting the descriptor values for the lowest energy conformer, the Boltzmann-weighted average and minimum/



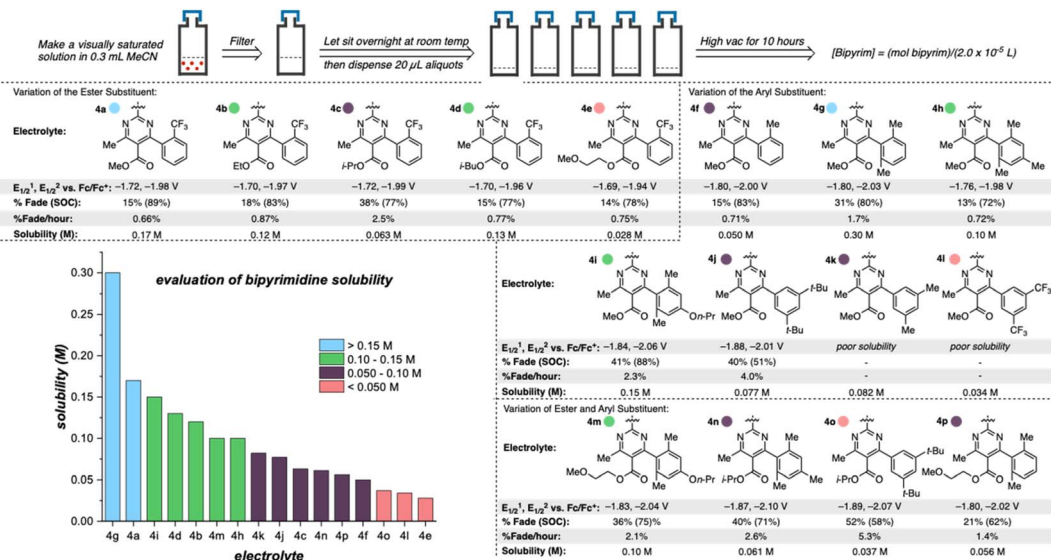


Fig. 3 Library of 2,2'-bipyrimidines. Solubilities were measured according to the method outlined at the top of the figure and reduction potentials internally referenced vs.  $Fc/Fc^+$ . A reported H-cell cycling assay was used to measure the percent capacity fade through the course of 50 cycles, state of charge (SOC; reported value is the highest measured), and percent fade per hour (total capacity fade per total experiment time).

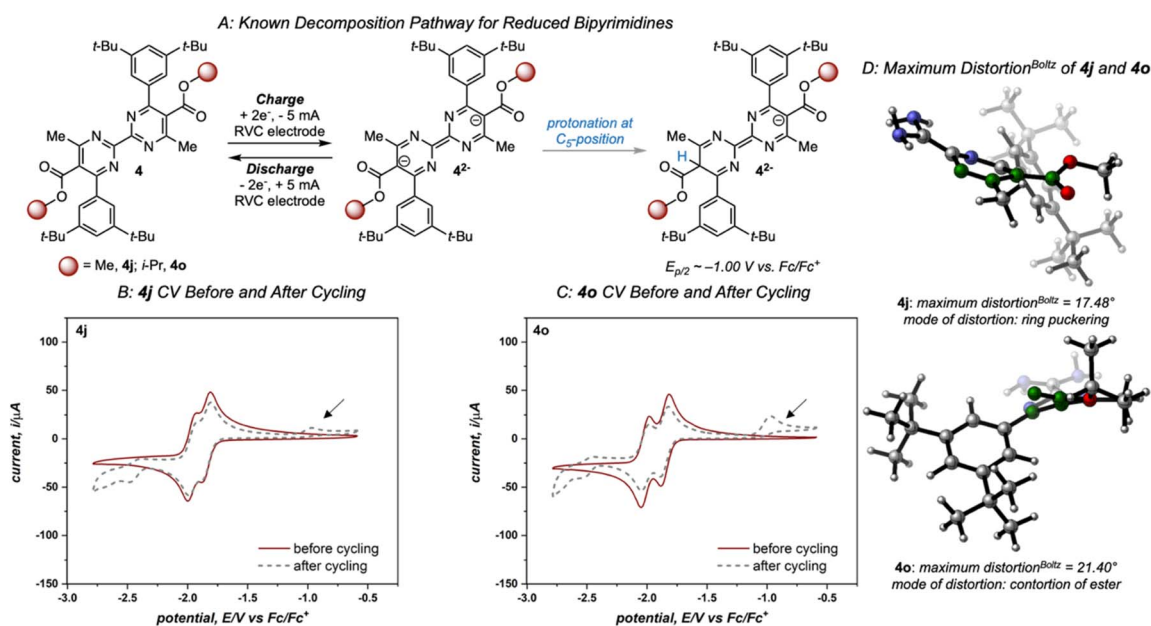


Fig. 4 (A) Known decomposition pathway of reduced bipyrimidine analytes. (B) and (C) Cyclic voltammograms of 2nd-generation candidates 4j and 4o, respectively, recorded before and after cycling at 100  $mV s^{-1}$  and an initial negative scan direction. The arrow indicates the diagnostic oxidation peak of the protonated bipyrimidine. (D) DFT computed structures of 4j and 4o with the maximum mode of distortion for each molecule (relevant atoms highlighted in green). While each mode serves to alleviate steric strain present in the reduced state, 4j exhibited a maximum distortion via ring puckering while 4o demonstrated torsion of the ester out of the plane of the pyrimidine core.

maximum values of each descriptor were collected to account for analyte flexibility. In addition to these DFT descriptors, descriptors from Schrödinger's QikProp library were also collected.

Following descriptor calculation, we next evaluated the data distribution to determine which statistical modeling technique

was best suited for solubility data analysis. Traditionally, linear regression has been used to correlate structure to function, but these methods failed in this study, likely due to poor data distribution.<sup>24</sup> This could be due to a poor representation of analytes with solubility >0.15 M. To expand the distribution of data, multiple 1st-generation bipyrimidines were included in



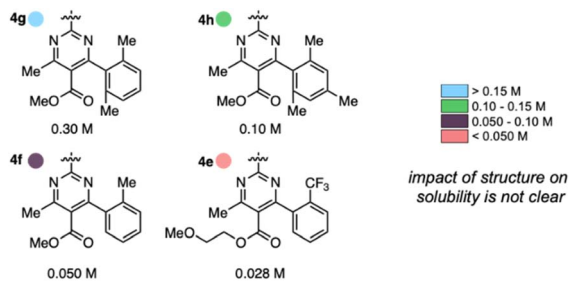


Fig. 5 Impact of anolyte structure on solubility with representative structures of each solubility class.

solubility statistical modeling efforts, even though their stability profile may have been undesirable. Additionally, this type of data distribution prompted the use of a classification algorithm whereby “good” versus “poor” measured solubilities can be binned using an algorithm-identified molecular descriptor and a user-defined threshold. This process, initially reported within the context of catalyst speciation in cross-coupling reactions,<sup>25</sup> has been integrated as a standard part of our statistical modeling workflow.<sup>26</sup>

This algorithm identified a solvent accessible surface area (SASA) descriptor to classify bipyrimidine solubility. Applying a user defined cutoff of 0.15 M, the single node decision tree (Fig. 6) revealed that anolytes with average solvent accessible surface area SASA, derived by moving a 1.4 Å probe over the enclosed molecular surface<sup>27</sup> values of  $<579 \text{ \AA}^2$  exhibit reasonable solubilities ( $\geq 0.15 \text{ M}$ ; blue squares in upper left quadrant), while anolytes with a SASA  $>579 \text{ \AA}^2$  have poor solubilities ( $\leq 0.15 \text{ M}$ , red squares in bottom right quadrant).

Within this classification, there are four false positives that represent molecules that the model identifies as soluble but have measured solubilities below the 0.15 M cut-off. Although the exact molecular feature(s) that contribute to the decreased solubilities of these four candidates is not clear, the model directs us to a specific property region to target for improved solubility.

Since anolyte stability is also a crucial metric, we sought to combine both the solubility and stability profiles of the 1st generation and 2nd generation libraries in a virtual screening workflow. Previously, linear regression was used to correlate the

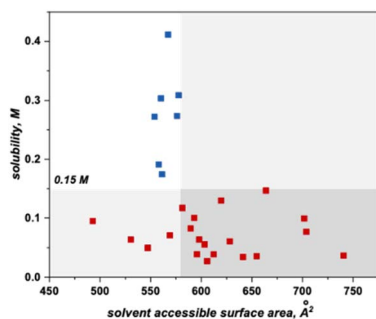


Fig. 6 Threshold of  $579 \text{ \AA}^2$  SASA based on a user-defined cutoff of 0.15 M.

experimental rate of fade to two molecular descriptors: the Boltzmann averaged maximum distortion (either ring puckering or contortion of the ester out of the bipyrimidine plane) and the Boltzmann-averaged NBO charge at both  $C_4/C_6$  (corresponding to the polarization of the  $C=N$  double bonds in the bipyrimidine). A heatmap that plots these two descriptors for all 1st and 2nd-generation anolytes was constructed, with points colored according to their rates of fade. Similar to the solubility threshold, the stability heatmap identifies a region of chemical space to target for design of stable candidates: bipyrimidines with maximum distortion values less than approximately  $15^\circ$  or NBO values less than 0.30 are envisioned to possess favorable stability.

As the next step, we utilized the solubility threshold and stability heatmap for a data-guided design of bipyrimidine candidates that meet or surpass the target solubility and de-risk future anolyte synthesis. Anolyte candidates were virtually screened according to a two-step workflow utilizing the SASA threshold for solubility and the heatmap describing anolyte stability as a function of the previously reported molecular descriptors (Fig. 7). Advantageously, the SASA term dictating solubility and the maximum distortion term describing stability are complementary to one another, in that smaller anolytes could potentially be soluble while also minimizing steric

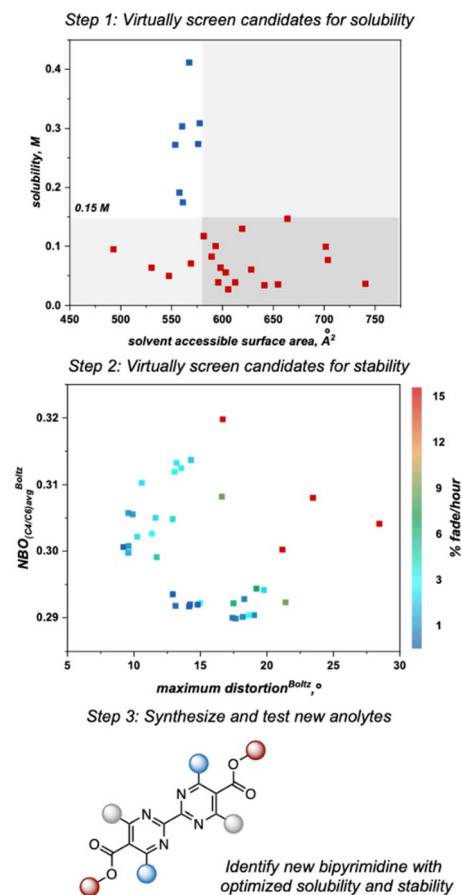


Fig. 7 Workflow utilized to virtually screen new anolytes with respect to solubility and stability.



interactions that induce decomposition. However, since structural variations designed to minimize the maximum distortion descriptor could also impact the average NBO charge at C<sub>4</sub>/C<sub>6</sub>, it can be difficult to design candidates that optimize both features simultaneously. Thus, candidates could be selected to optimize one of these features or to balance the two descriptors (Fig. 8).

A set of possible bipyrimidines was designed based on the computed SASA values (for solubility) and their positions in the heatmap describing anolyte stability. Notably, three candidates within the virtual screening process contain a tetramethyl substitution pattern. Despite the fact that this could lead to increased NBO charges at the C<sub>4</sub>/C<sub>6</sub> position (and potentially a slightly less stable anolyte),<sup>17</sup> this substitution pattern was pursued based on initial investigation of the SASA of **4aC**, along with two other analogs bearing tetramethyl substitution, **4aB** and **4aD** (Fig. 8A). Candidates with desirable SASA values along with minimized distortion (**4aB**, **4aD**, **4aE**, **4aF**), minimized NBO charge (**4aE**, **4aF**), or minimization of both stability descriptors (**4aE**, **4aF**) were selected for synthesis (Fig. 8B).

As before, experimental results demonstrated that ester identity had a substantial impact on solubility, represented by an order of magnitude increase in solubility when the methyl ester in **4aB** is changed to an ethyl ester in **4aC**. Furthermore, **4aD** was isolated as a viscous, miscible liquid that slowly solidified over a matter of weeks and was later found to have a solubility of ~1.3 M (Fig. 8C). This constitutes an ~50-fold increase in solubility of relative to **4e**. Accompanying this increase in solubility was a slight decrease in stability (1.6% fade per hour relative to the most stable anolyte **4a** with 0.66% fade per hour). This was not unexpected due the increased

average NBO charge induced by the tetra-methyl substitution pattern. Interestingly, **4aF** was highly unstable (13% fade per hour) despite the minimal distortion and NBO charge. This could be indicative of a yet unidentified mechanism of decomposition for bipyrimidines possessing a fluorinated aryl substituent. With the identification of **4aD** as a highly soluble candidate with good stability, we next sought to evaluate its performance in a flow battery.

To fully utilize the two highly energetic reduction events of **4aD** it was necessary to identify a compatible catholyte that undergoes multiple oxidation events at high potentials. Preliminary compatibility tests of catholytes (see ESI† for details) identified diaminocyclopropenium-phenothiazine **4-DMPP** as a suitable species, with two high-energy oxidation events (0.64 V and 1.00 V vs. Fc/Fc<sup>+</sup>)<sup>10e</sup> and reasonable compatibility (Fig. S3†).

Electrolytes **4aD** and **4-DMPP** were cycled in an asymmetric redox flow battery using a Fumasep FAP-375-PP separator and 25 mM solutions of **4aD** and **4-DMPP** in 0.5 M potassium hexafluorophosphate/acetonitrile (Fig. 9A). The solutions were flowed through the cell at 10 mL min<sup>-1</sup> and were electrochemically cycled at 15 mA cm<sup>-2</sup> with a cell potential up to 3 V, one of the highest cell potentials demonstrated in a non-aqueous two-electron redox flow battery.<sup>28</sup> This system achieved 49% material utilization on the first cycle and showed >95% capacity retention over 75 cycles with an average power density of 0.74 W h L<sup>-1</sup> power density (1.55 W h L<sup>-1</sup> theoretical power density).<sup>29</sup> However, catastrophic failure of the system was observed after these initial 75 cycles (see ESI†). Cyclic voltammetry analysis of the anolyte and catholyte solutions after

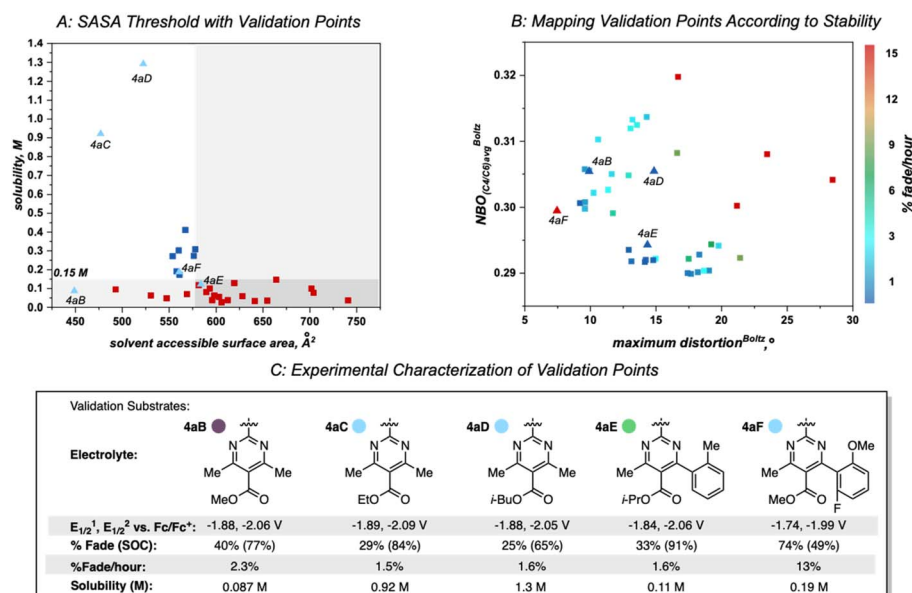


Fig. 8 (A). Threshold at 579 Å<sup>2</sup> SASA with validation points (labeled and shown by light blue triangles) plotted by their SASA value. (B) Heat map of the molecular descriptors dictating stability. Blue points indicate highly stable anolytes that have low rates of fade while red points indicate unstable molecules that have high rates of fade. Note that **4aC** was characterized with respect to stability in the first paper and not re-evaluated in the heatmap describing anolyte stability. (C) Characterization of validation substrates. Reduction potentials internally referenced vs. Fc/Fc<sup>+</sup>. A reported H-cell cycling assay was used to measure the percent capacity fade through the course of 50 cycles, state of charge (SOC; reported value is the highest measured), and percent fade per hour (total capacity fade per total experiment time).



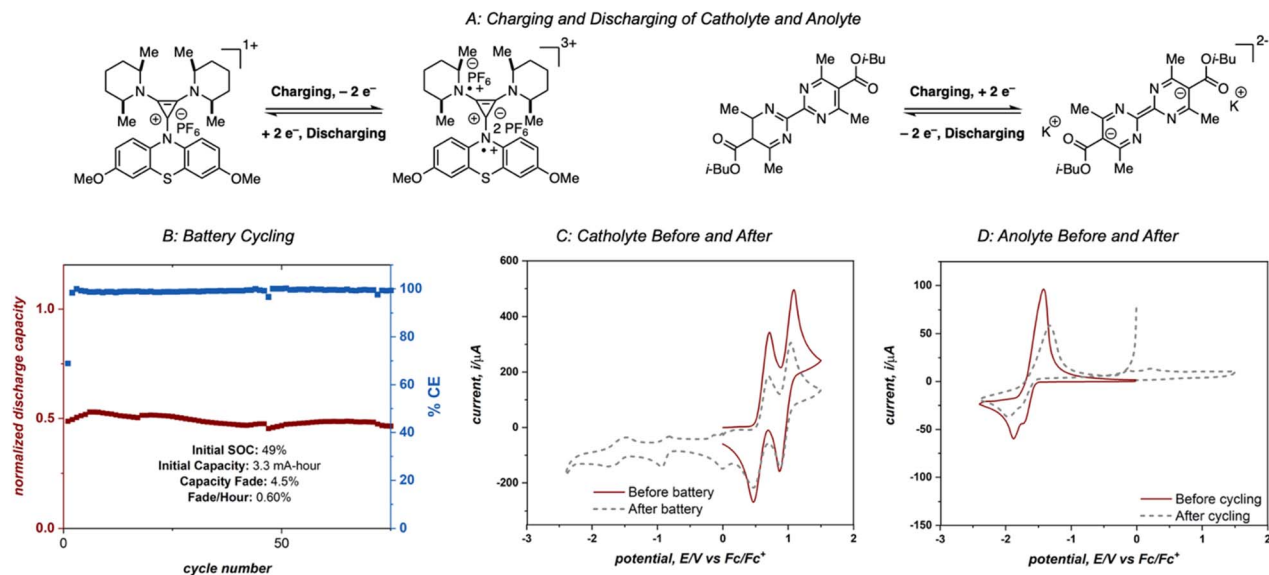


Fig. 9 (A) Electrolyte charging and discharging processes. (B) Flow cell cycling of 25 mM 4aD and 4-DMPP. (C) Cyclic voltammograms of catholyte recorded before and after cycling at  $100 \text{ mV s}^{-1}$  and an initial positive scanning direction. (D) Cyclic voltammograms of anolyte recorded before and after cycling at  $100 \text{ mV s}^{-1}$  and an initial negative scanning direction.

180 cycles revealed that crossover of the anolyte as well as molecular decomposition of both compounds contributed to the loss of capacity and system failure (Fig. 9C and D), ultimately disqualifying this system from operating in a high-concentration battery. However, this experiment demonstrates the cycling capability of 4aD in a stable, high energy flow battery.

## Conclusions

In summary, we report herein the design of a library of 16 new 2,2'-bipyrimidine anolytes with the intent of analyzing the impact of structure on solubility. Additionally, the design of this 2nd-generation library was carried out with respect for the structural features responsible for stable bipyrimidines. Aiming for a target concentration of 0.5 M (1 M  $e^-$  since the bipyrimidine can store two-electrons per molecule) to avoid mass transfer limitations, a narrow range of solubilities was observed within the initial library. This prompted the inclusion of multiple 1st generation bipyrimidines (independent of the corresponding stability) to improve dataset diversity. A non-linear threshold relationship between the solvent accessible surface area and solubility was identified, and when paired with a heat map describing anolyte stability with previously identified molecular descriptors, five new anolyte candidates were designed in a virtual screening campaign and experimentally tested. One candidate was identified to have a 46-fold increase in solubility (relative to the least soluble anolyte) of 1.3 M while minimally sacrificing stability. Pairing this compound with a diaminocyclopropenium-phenothiazine derivative accessed a proof-of-principle organic redox flow battery with a cell potential up to 3 V. Although relatively stable cycling was achieved for the first 75 cycles, drastic battery failure was observed

beyond 75 cycles. Cyclic voltammograms after cycling suggest that electrolyte decomposition, potentially due to anolyte crossover, was a primary mechanism of failure. Going forward, evolution of the bipyrimidine scaffold will focus on maintaining the gains made in stability and solubility while also optimizing other battery relevant properties, such as crossover.

## Data availability

The ESI† is freely available at <https://doi.org/10.1039/d3sc04084d> and contains the following for this article: experimental procedures for all reactions and electrochemical measurements; spectroscopic characterization of new compounds; detailed computational methods and model evaluation; copies of  $^1\text{H}$ ,  $^{13}\text{C}$ , and  $^{19}\text{F}$  NMR spectra.

## Author contributions

A. R. P. and M. S. S. conceptualized the project. A. R. P. and S. L. M. performed experimental investigations. A. R. P., S. L. M., and S. L. G. carried out modeling studies. R. W. K. and B. J. prepared and operated the flow battery. All authors contributed to writing and editing the manuscript. M. S. S. and M. S. S. acquired funding for the project.

## Conflicts of interest

There are no conflicts to declare.

## Acknowledgements

We acknowledge Jeremy Griffin for material. This research was supported by the Joint Center for Energy Storage Research (JCESR), a U.S. Department of Energy, Energy Innovation Hub.



Computational resources were provided from the Center for High Performance Computing (CHPC) at the University of Utah. NMR results included in this report were recorded at the David M. Grant NMR Center, a University of Utah Core Facility. Funds for construction of the Center and the helium recovery system were obtained from the University of Utah and the National Institutes of Health awards 1C06RR017539-01A1 and 3R01GM063540-17W1, respectively. NMR instruments were purchased with support of the University of Utah and the National Institutes of Health award 1S10OD25241-01. We acknowledge Prof. Shelley Minter for helpful discussions.

## Notes and references

- IRENA, *World Energy Transitions Outlook 2022: 1.5 °C Pathway*, International Renewable Energy Agency, Abu Dhabi, 2022.
- (a) Z. Yang, J. Zhang, M. C. W. Kintner-Meyer, X. Lu, D. Choi, J. P. Lemmon and J. Liu, *Chem. Rev.*, 2011, **111**, 3577–3613; (b) G. L. Soloveichik, *Chem. Rev.*, 2015, **115**, 11533–11558; (c) P. Denholm and M. Hand, *Energy Policy*, 2011, **39**, 1817–1830.
- J. Luo, B. Hu, M. Hu, Y. Zhao and T. L. Liu, *ACS Energy Lett.*, 2019, **4**, 2220–2240.
- D. G. Kwabi, Y. Ji and M. J. Aziz, *Chem. Rev.*, 2020, **120**, 6467–6489.
- (a) J. Winsberg, T. Hagemann, T. Janoschka, M. D. Hager and U. S. Schubert, *Angew Chem. Int. Ed. Engl.*, 2017, **56**, 686–711; (b) K. Lourenszen, J. Williams, F. Ahmadpour, R. Clemmer and S. Tasnim, *J. Energy Storage*, 2019, **25**, 100844; (c) For reports on other metal-based batteries, the reader is directed to the following: Z. Zhu, T. Jiang, M. Ali, Y. Meng, Y. Jing, Y. Cui and W. Chen, *Chem. Rev.*, 2022, **122**, 16610–16751; (d) Y. Yao, H. Xu, Z. Tian, J. Zhang, F. Zhan, M. Yan and C. Jia, *ACS Appl. Energy Mater.*, 2021, **4**, 8052–8058; (e) Y. Long, Z. Xu, G. Wang, H. Xu, M. Yang, M. Ding, D. Yuan, C. Yan, Q. Sun, M. Liu and C. Jia, *iScience*, 2021, **24**, 103157–103169; (f) F. Zhong, M. Yang, M. Ding and C. Jia, *Front. Chem.*, 2020, **8**, 451–464; (g) G. Wang, H. Zou, Z. Xu, A. Tang, F. Zhong, X. Zhu, C. Qin, M. Ding, W. You and C. Jia, *Mater. Today Energy*, 2022, **28**, 101061–101068; (h) S. Yan, S. Huang, H. Xu, L. Li, H. Zou, M. Ding, C. Jia and Q. Wang, *ChemSusChem*, 2023, e202300710–e202300716.
- K. Gong, Q. Fang, S. Gu, S. F. Y. Li and Y. Yan, *Energy Environ. Sci.*, 2015, **8**, 3515–3530.
- A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, John Wiley & Sons, Inc., Hoboken, New Jersey, United States of America, 2nd edn, 2001.
- (a) K. H. Hendriks, C. S. Sevov, M. E. Cook and M. S. Sanford, *ACS Energy Lett.*, 2017, **2**, 2430–2435; (b) C. S. Sevov, K. H. Hendriks and M. S. Sanford, *J. Phys. Chem. C*, 2017, **121**, 24376–24380; (c) B. Hu and T. L. Liu, *J. Energy Chem.*, 2018, **27**, 1326–1332; (d) X. Wang, J. Chai, A. Lashgari and J. J. Jiang, *ChemElectroChem*, 2020, **8**, 83–89; (e) N. Daub, R. A. Janssen and K. H. Hendriks, *ACS Appl. Energy Mater.*, 2021, **4**, 9248–9257; (f) Y. Yan, L. Zhang, R. Walser-Kuntz, D. B. Vogt, M. S. Sigman, G. Yu and M. S. Sanford, *Chem. Mater.*, 2022, **34**, 10594–10605.
- C. S. Sevov, S. K. Samaroo and M. S. Sanford, *Adv. Energy Mater.*, 2017, **7**, 1602027.
- (a) J. Zhang, Z. Yang, I. A. Shkrob, R. S. Assary, S. O. Tung, B. Silcox, W. Duan, J. Zhang, C. C. Su, B. Hu, B. Pan, C. Liao, Z. Zhang, W. Wang, L. A. Curtiss, L. T. Thompson, X. Wei and L. Zhang, *Adv. Energy Mater.*, 2017, **7**, 1701272; (b) K. H. Hendriks, S. G. Robinson, M. N. Braten, C. S. Sevov, B. A. Helms, M. S. Sigman, S. D. Minter and M. S. Sanford, *ACS Cent. Sci.*, 2018, **4**, 189–196; (c) Y. Yan, T. P. Vaid and M. S. Sanford, *J. Am. Chem. Soc.*, 2020, **142**, 17564–17571; (d) Y. Yan, S. G. Robinson, T. P. Vaid, M. S. Sigman and M. S. Sanford, *J. Am. Chem. Soc.*, 2021, **143**, 13450–13459; (e) Y. Yan, D. B. Vogt, T. P. Vaid, M. S. Sigman and M. S. Sanford, *Angew Chem. Int. Ed. Engl.*, 2021, **60**, 27039–27045.
- X. Wang, R. K. Gautam and J. J. Jiang, *Batteries Supercaps*, 2022, **5**, e202200298.
- S. G. Robinson, Y. Yan, K. H. Hendriks, M. S. Sanford and M. S. Sigman, *J. Am. Chem. Soc.*, 2019, **141**, 10171–10176.
- (a) F. H. Vermeire, Y. Chung and W. H. Green, *J. Am. Chem. Soc.*, 2022, **144**, 10785–10797; (b) M. R. Tuttle, E. M. Brackman, F. Sorourifar, J. Paulson and S. Zhang, *J. Phys. Chem. Lett.*, 2023, **14**, 1318–1325.
- (a) C. M. Wong and C. S. Sevov, *ACS Energy Lett.*, 2021, 1271–1279; (b) D. Kim, M. S. Sanford, T. P. Vaid and A. J. McNeil, *Chemistry*, 2022, **28**, e202200149; (c) R. Yan and Q. Wang, *Adv. Mater.*, 2018, **30**, e1802406; (d) G. Lee, C. M. Wong and C. S. Sevov, *ACS Energy Lett.*, 2022, **7**, 3337–3344.
- J. Zhang, R. E. Corman, J. K. Schuh, R. H. Ewoldt, I. A. Shkrob and L. Zhang, *J. Phys. Chem. C*, 2018, **122**, 8159–8172.
- L. Zhang, R. Feng, W. Wang and G. Yu, *Nat. Rev. Chem*, 2022, **6**, 524–543.
- J. D. Griffin, A. R. Pancoast and M. S. Sigman, *J. Am. Chem. Soc.*, 2021, **143**, 992–1004.
- C. S. Sevov, D. P. Hickey, M. E. Cook, S. G. Robinson, S. Barnett, S. D. Minter, M. S. Sigman and M. S. Sanford, *J. Am. Chem. Soc.*, 2017, **139**, 2924–2927.
- B. T. Boyle, M. C. Hilton and A. McNally, *J. Am. Chem. Soc.*, 2019, **141**, 15441–15449.
- M. C. Hilton, X. Zhang, B. T. Boyle, J. V. Alegre-Requena, R. S. Paton and A. McNally, *Science*, 2018, **362**, 799–804.
- J. D. Milshtein, A. P. Kaur, M. D. Casselman, J. A. Kowalski, S. Modekrutti, P. L. Zhang, N. H. Attanayake, C. F. Elliott, S. R. Parkin, C. Risko, F. R. Brushett and S. A. Odom, *Energy Environ. Sci.*, 2016, **9**, 3531–3543.
- C. Lu, C. Wu, D. Ghoreishi, W. Chen, L. Wang, W. Damm, G. A. Ross, M. K. Dahlgren, E. Russell, C. D. Von Bargen, R. Abel, R. A. Friesner and E. D. Harder, *J. Chem. Theory Comput.*, 2021, **17**, 4291–4300.
- (a) M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681; (b) Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241; (c) V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.





- 24 (a) D. M. Lustosa and A. Milo, *ACS Catal.*, 2022, **12**, 7886–7906; (b) C. B. Santiago, J. Y. Guo and M. S. Sigman, *Chem. Sci.*, 2018, **9**, 2398–2412.
- 25 S. H. Newman-Stonebraker, S. R. Smith, J. E. Borowski, E. Peters, T. Gensch, H. C. Johnson, M. S. Sigman and A. G. Doyle, *Science*, 2021, **374**, 301–308.
- 26 (a) J. P. Liles, C. Rouget-Virbel, J. L. H. Wahlman, R. Rahimoff, J. M. Crawford, A. Medlin, V. S. O'Connor, J. Li, V. A. Roytman, F. D. Toste and M. S. Sigman, *Chem*, 2023, **9**, 1518–1537; (b) J. Rein, S. D. Rozema, O. C. Langner, S. B. Zacate, M. A. Hardy, J. C. Siu, B. Q. Mercado, M. S. Sigman, S. J. Miller and S. Lin, *Science*, 2023, **380**, 706–712.
- 27 W. L. Jorgensen and E. M. Duffy, *Adv. Drug Delivery Rev.*, 2002, **54**, 355–366.
- 28 (a) Y. Yan, S. G. Robinson, M. S. Sigman and M. S. Sanford, *J. Am. Chem. Soc.*, 2019, **141**, 15301–15306; (b) Although we are uncertain at this time as to the exact reason behind the increase in capacity after the first 20 cycles, cell and electrode conditioning processes may play a role.
- 29 For a range of contemporary power power densities in flow batteries, the reader is directed to the following references: (a) S. K. Pahari, T. C. Gokoglan, B. R. B. Visayas, J. Woehl, J. A. Golen, R. Howland, M. L. Mayes, E. Agar and P. J. Cappillino, *RSC Adv.*, 2021, **11**, 5432–5443; (b) W. Zhang, R. Walser-Kuntz, J. S. Tracy, T. K. Schramm, J. Shee, M. Head-Gordon, G. Chen, B. A. Helms, M. S. Sanford and F. Dean Toste, *J. Am. Chem. Soc.*, 2023, **145**, 18877–18887; (c) G. D. De La Garza, A. P. Kaur, I. A. Shkrob, L. A. Robertson, S. A. Odom and A. J. McNeil, *J. Mater. Chem. A*, 2022, **10**, 18745–18752.

