Integrating Data Science Tools into Molecular Design in Electrocatalysis and Energy Applications

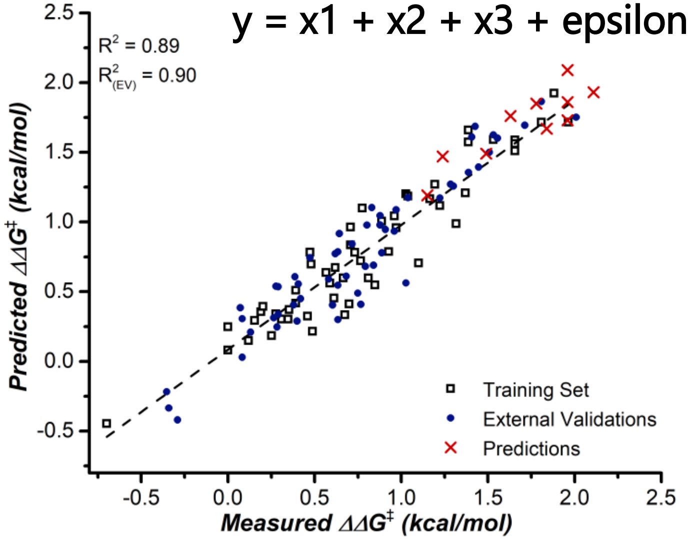
National Academy of Sciences
Workshop on Electrochemistry
November 19, 2019
@Sigman_Lab

DEPARTMENT OF
CHEMISTRY
THE UNIVERSITY OF UTAH

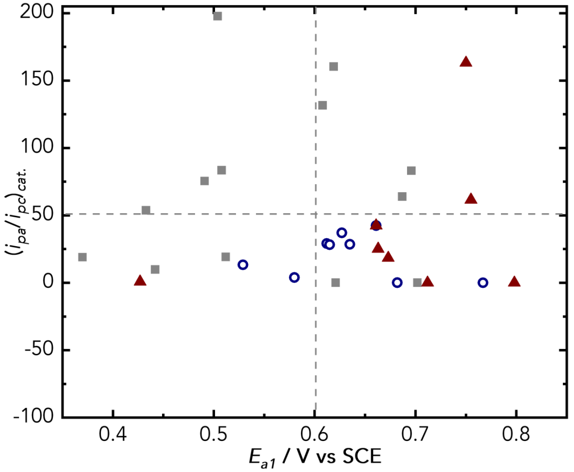
Outline of talk

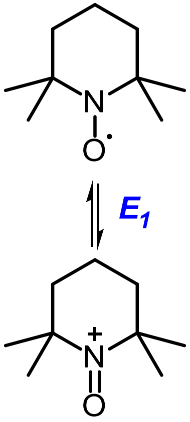
introduction to data science

y = x1 + x2 + x3 + epsilon

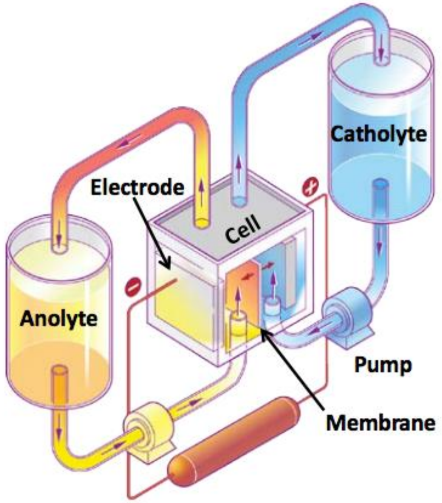


case study 1: electrocatalysis

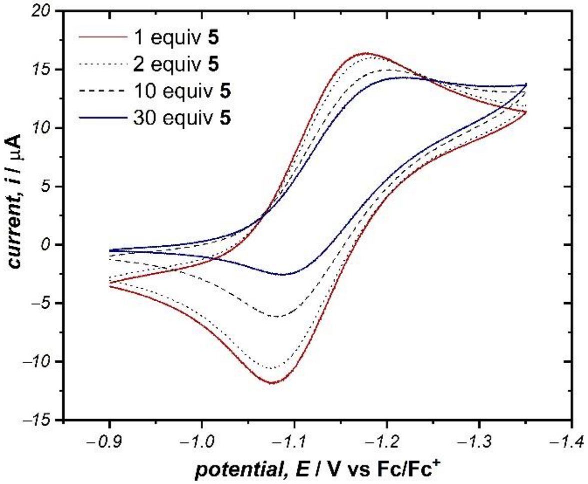




case study 2: flow batteries

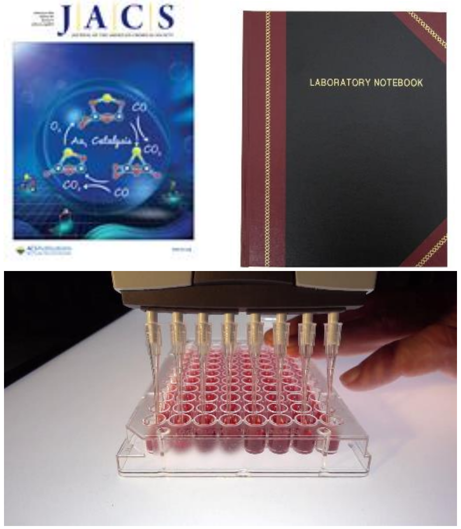


future applications and thoughts



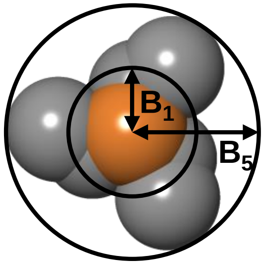
Data science meets physical organic chemistry

empirical data collection and mining

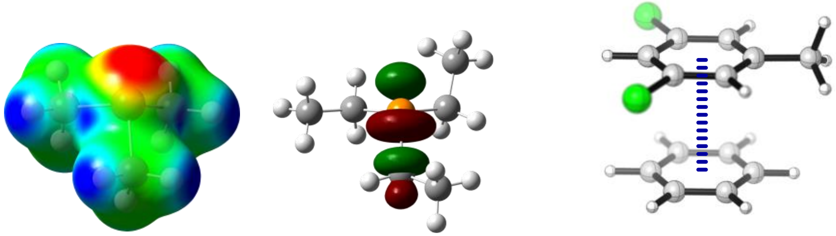


- No wasted data
- Use substrate & catalyst variation
- Data set size?
- Data set quality?

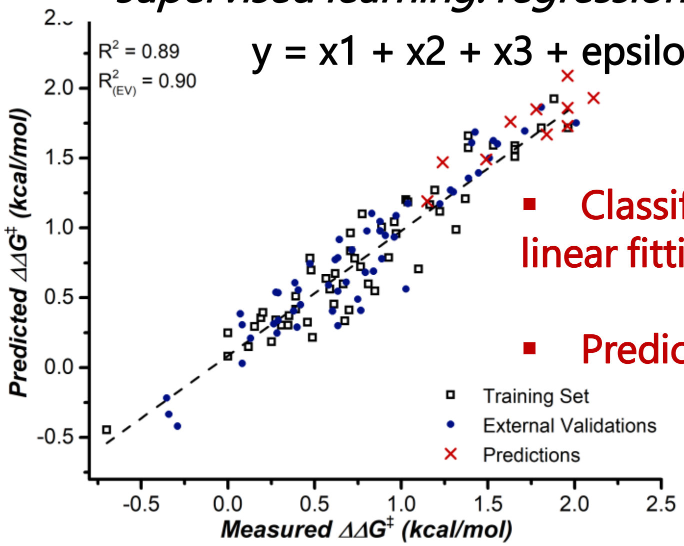
molecular featurization and property sets



- How do you best describe structure?
- Physical meaning?



supervised learning: regression tools



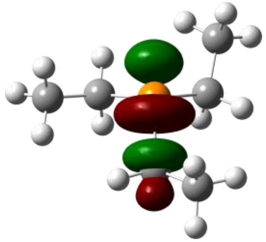
y = x1 + x2 + x3 + epsilon

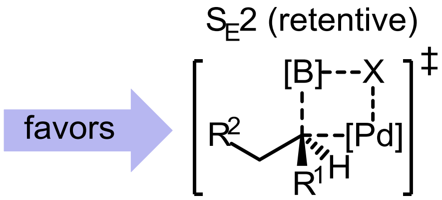
- Classification, linear fitting, or ML?
- Predictive?

transfer lSearning?

$E_{s*(P–C)avg}^{Boltz.}$: average energies of P–C antibonding orbitals

related to π-back bonding





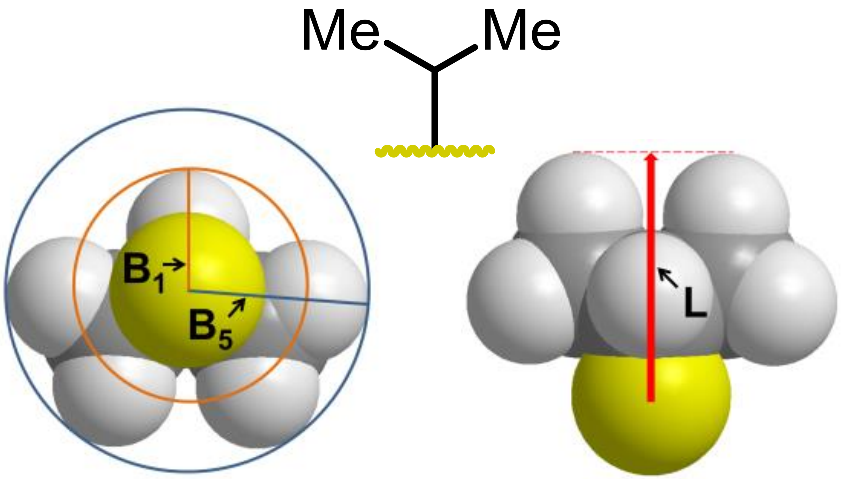
- Conversion of math to mechanism?
- Translate to new systems

Sigman, M. S. ; Harper, K. C.; Bess, E. N.; Milo, A. *Acc. Chem. Res.* **2016**, *49*, 1292–1301.
with Santiago, C. B.; Guo, J.-Y. *Chem. Sci.* **2018**, *9*, 2398-2412.
with Reid, J. L. *Nature Rev. Chem.* **2018**, *2*, 290-305.

Our typical molecular features

steric parameters used: Sterimol

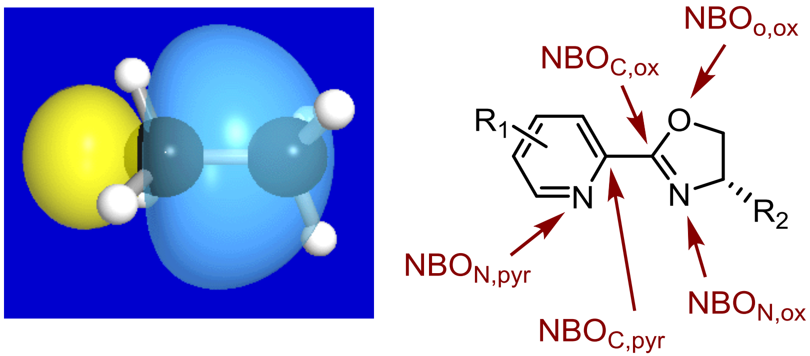
- Verloop developed in 1970's
- Three parameters
  - **L** is maximum length
  - **B**$_1$ is minimal radius perpendicular to **L**
  - **B**$_5$ is maximum radius perpendicular to **L**
  - **L** & **B**$_5$ are not independent



- **Other parameters**
  - Torsion angles, distances, % Vbur etc…

with Harper, K. C.; Bess, E. N. Nature Chem. **2012**, 366

standard electronic parameters



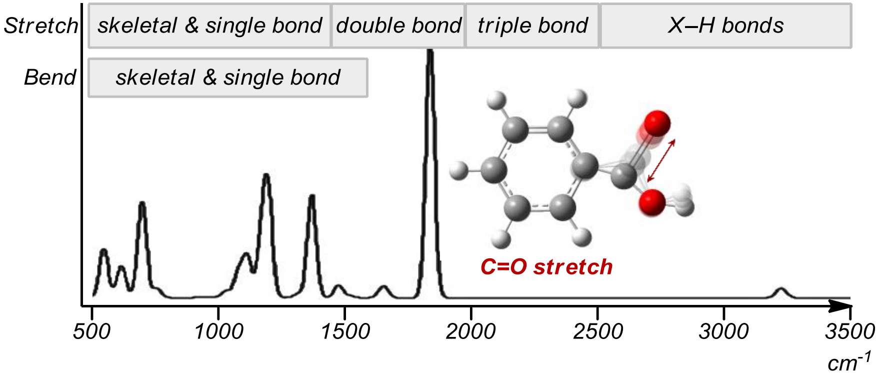
NBO

'natural bonding orbital'

Reed, A. E.; Weinhold, F. *J. Chem. Phys.*, **1983**, *78*, 4066

hybrid parameters

$$\nu = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu_{red}}}$$



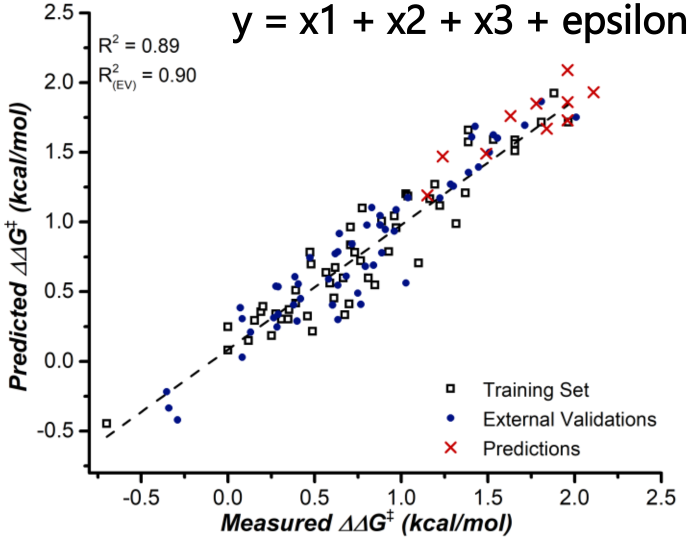
With Milo, A.; Bess, E. N *Nature* **2014**, *507*, 210

Jones, R. N.; et. Al. *Can. J. Chem.* **1957**, *35*, 504

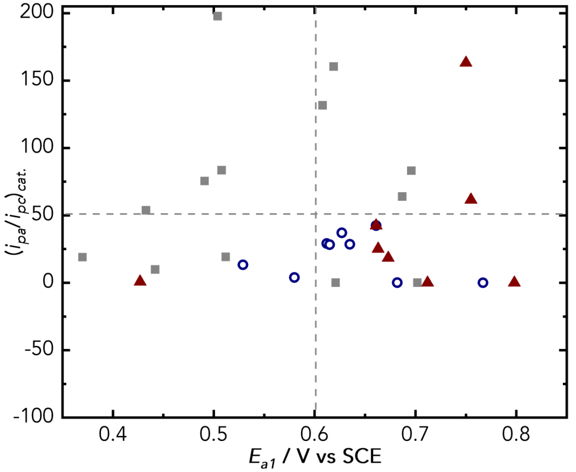
Outline of talk

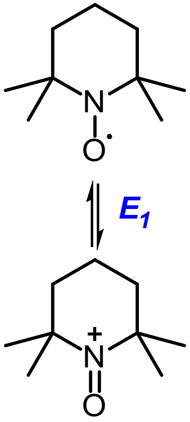
introduction to data science

y = x1 + x2 + x3 + epsilon

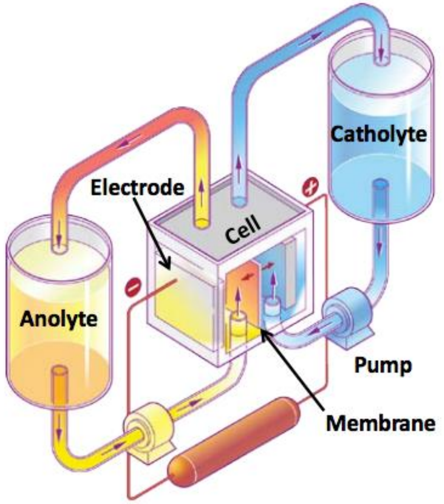


case study 1: electrocatalysis

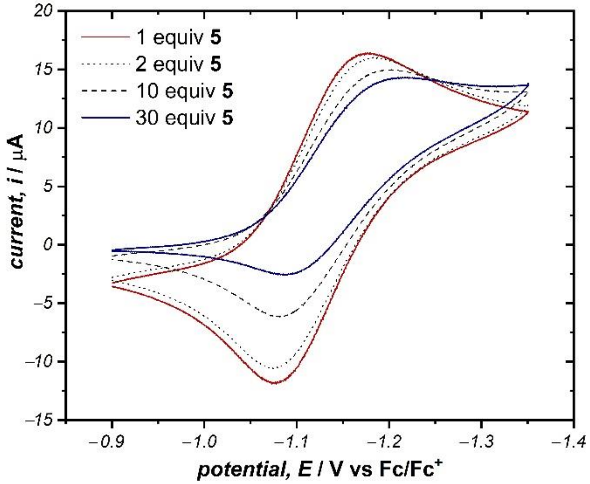




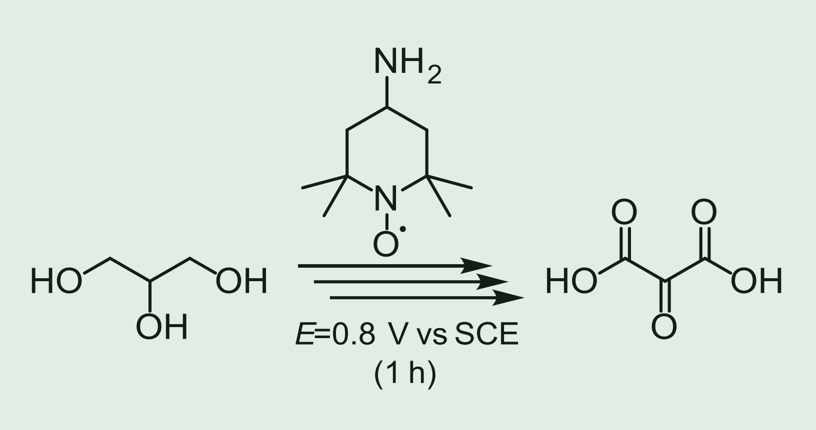
case study 2: flow batteries



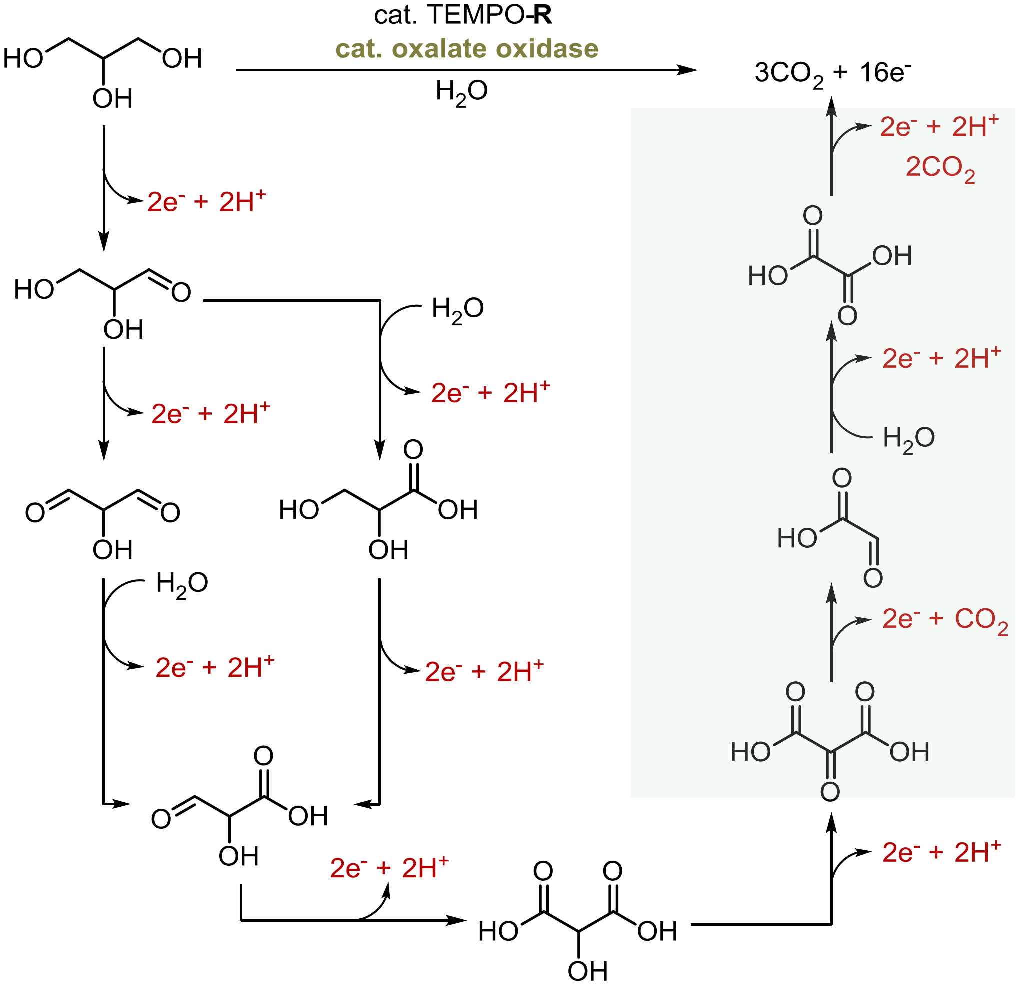
future applications and thoughts



Designing electrocatalytic oxidation cascades

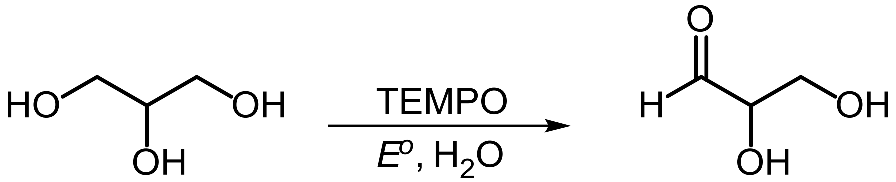


- Complete oxidation of glycerol to CO_2 was achieved using a hybrid catalytic system
- As many as 16 electrons were collected per molecule of glycerol
- *TEMPO-NH_2 was found to exhibit anomalous catalytic properties that allow for expanded pH range*

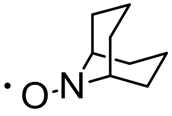


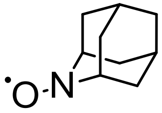
Hickey, D. P.; McCammant, M. S.; Giroud, F.; Sigman, M. S.; Minteer, S. D. *J. Am. Chem. Soc.* **2014**, *136*, 15917.

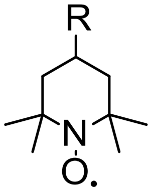
TEMPO-NH_2 and how do we adjust properties



	Electrochemical Potential $E_{1/2}$ / V vs SCE	Catalytic Activity J_{max} / μA cm^{-2}
⇨ AZADO	0.466	855
⇨ ABNO	0.485	801
TEMPO	0.499	282
TEMPO-COOH	0.548	224
TEMPO-OH	0.583	507
⇨ TEMPO-NH_2	0.623	1453
TEMPO=O	0.685	0
TEMPO-NHC=OMe	0.603	274
TEMPO-OMe	0.493	297
Me-AZADO	0.403	313


ABNO

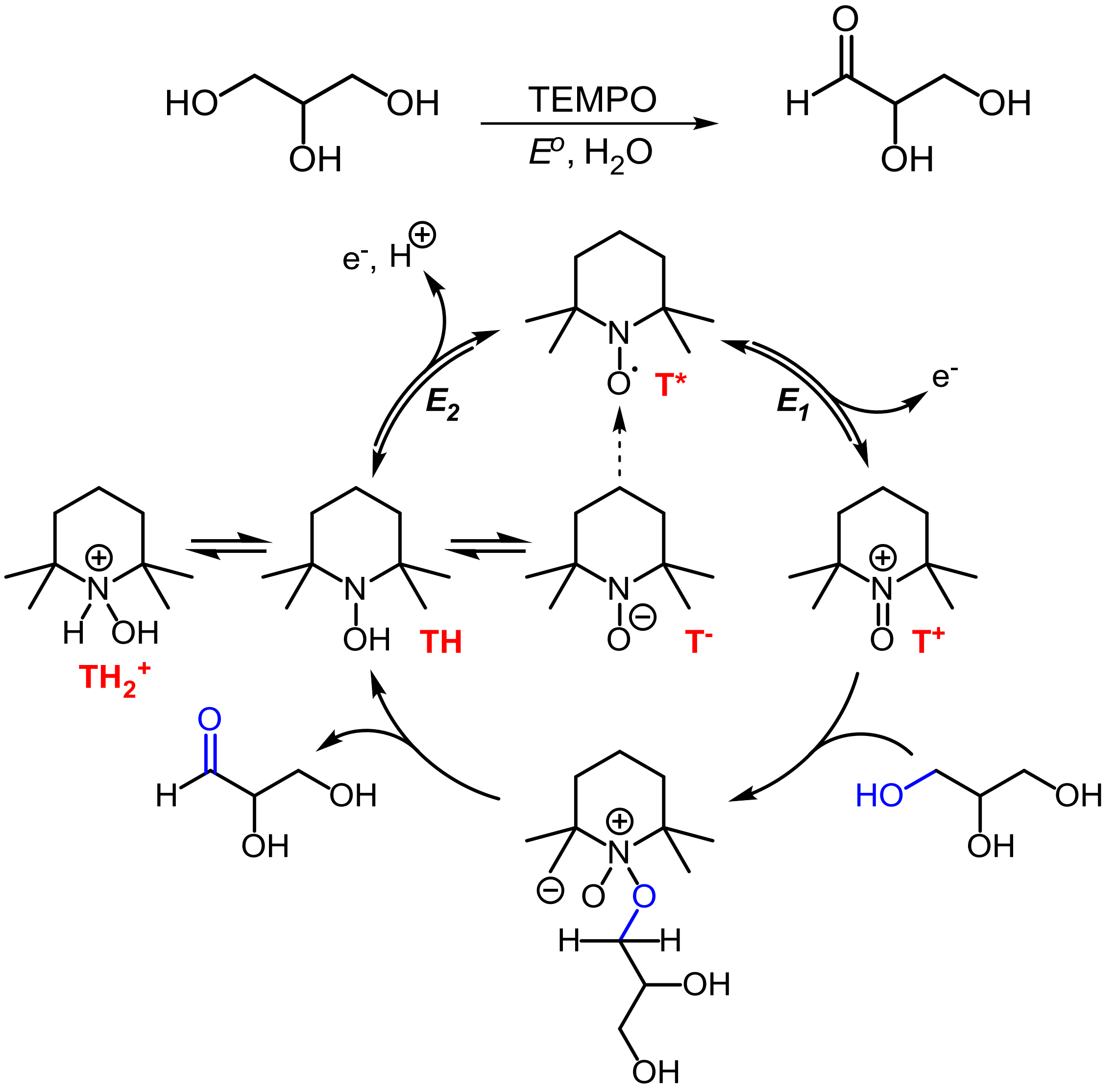

AZADO


TEMPO-R

GOAL: predict catalytic activity, eliminate the need for catalyst screening

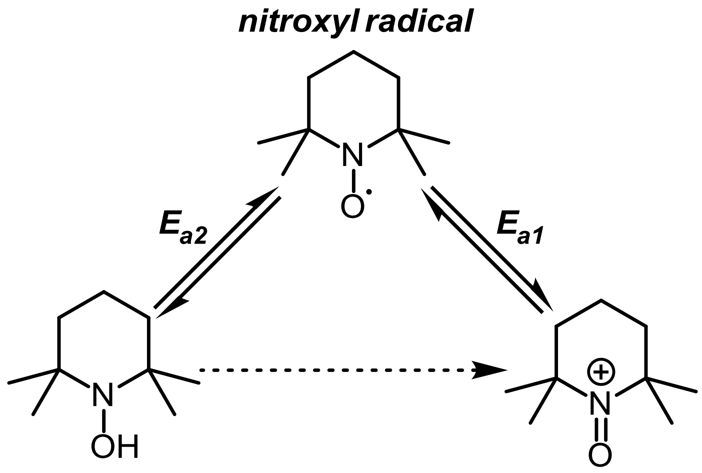
- need a better understanding of structure-function relationships
- enable design of TEMPO derivatives for various diverse applications

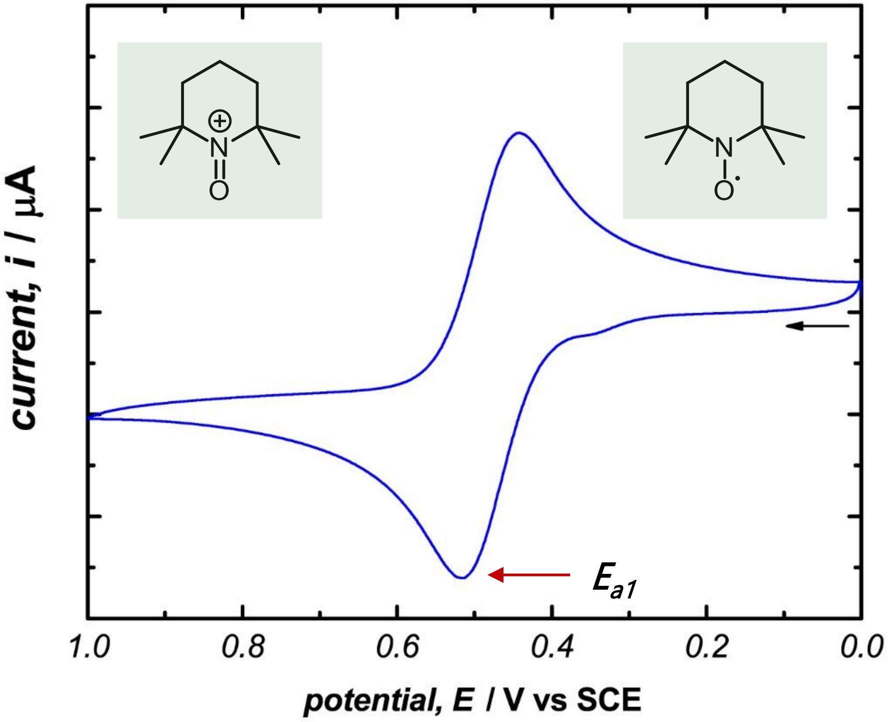
Initiate study by considering the mechanism

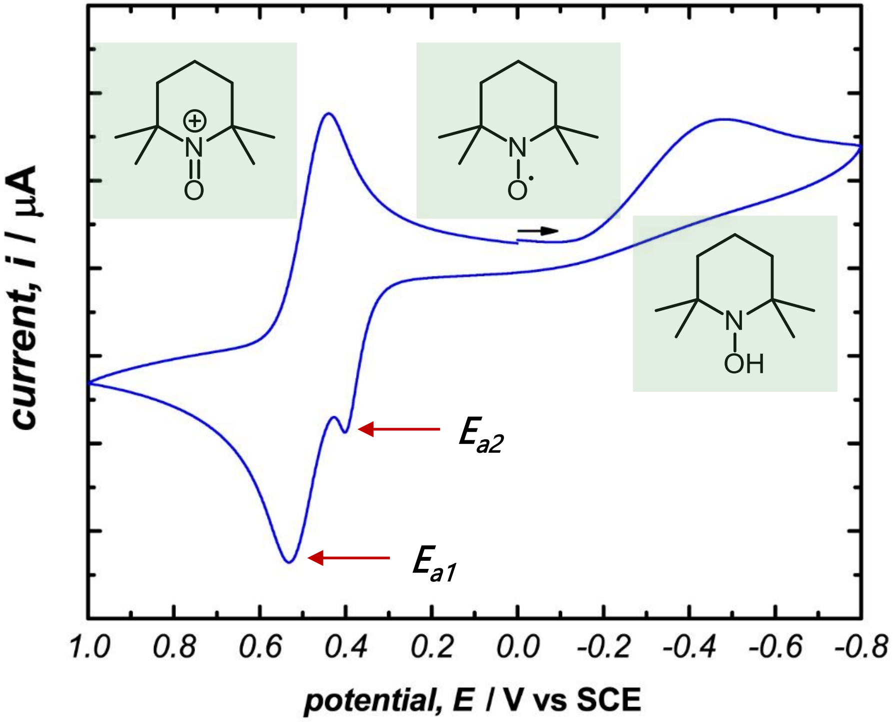


- ➢ The nature of E_2 is not well understood
- ➢ E_1 and E_2 can be accessed electrochemically

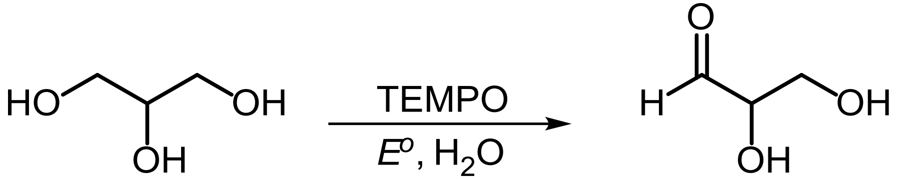
TEMPO exists in three oxidation states

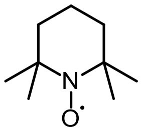




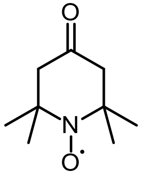


Evaluation of a broad range of TEMPO derivatives

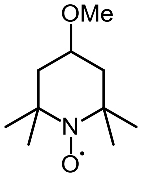




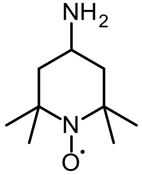
E_{a1}= 0.53
j_{max}= 239



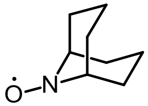
E_{a1}= 0.67
j_{max}= 33



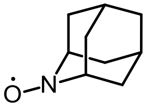
E_{a1}= 0.62
j_{max}= 371



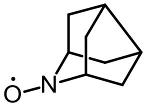
E_{a1}= 0.67
j_{max}= 1200



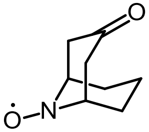
E_{a1}= 0.51
j_{max}= 1217



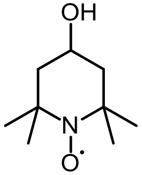
E_{a1}= 0.50
j_{max}= 1200



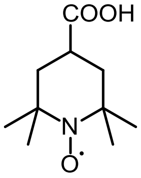
E_{a1}= 0.50
j_{max}= 2169



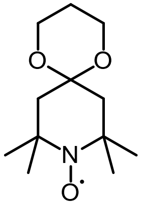
E_{a1}= 0.70
j_{max}= 6469



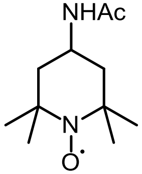
E_{a1}= 0.61
j_{max}= 515



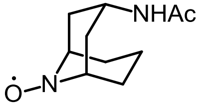
E_{a1}= 0.58
j_{max}= 159



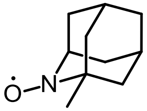
E_{a1}= 0.63
j_{max}= 484



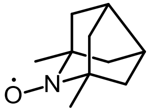
E_{a1}= 0.64
j_{max}= 518



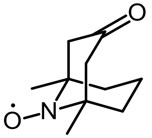
E_{a1}= 0.61
j_{max}= 1188



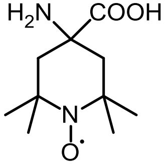
E_{a1}= 0.43
j_{max}= 657



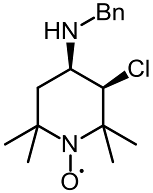
E_{a1}= 0.37
j_{max}= 303



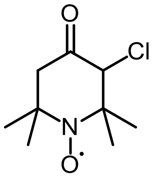
E_{a1}= 0.62
j_{max}= 184



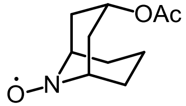
E_{a1}= 0.43
j_{max}= 56



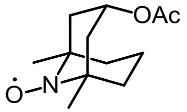
E_{a1}= 0.70
j_{max}= 55



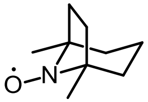
E_{a1}= 0.77
j_{max}= 31



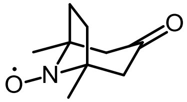
E_{a1}= 0.62
j_{max}= 2595



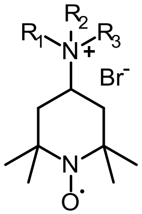
E_{a1}= 0.51
j_{max}= 458



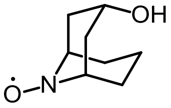
E_{a1}= 0.59
j_{max}= 639



E_{a1}= 0.70
j_{max}= 0

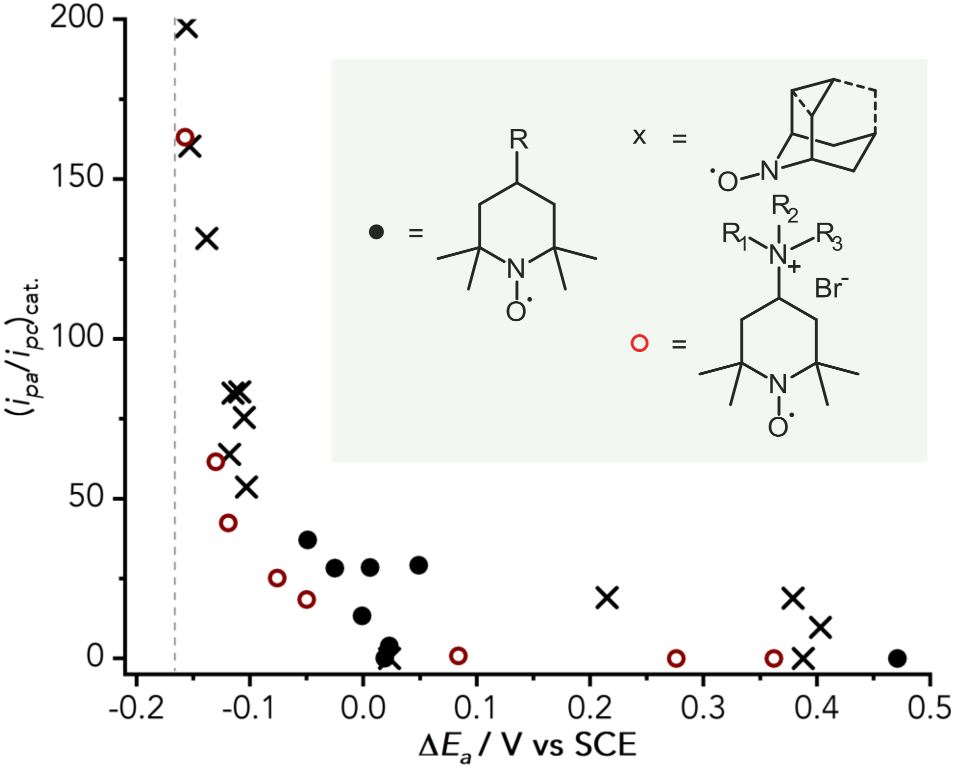


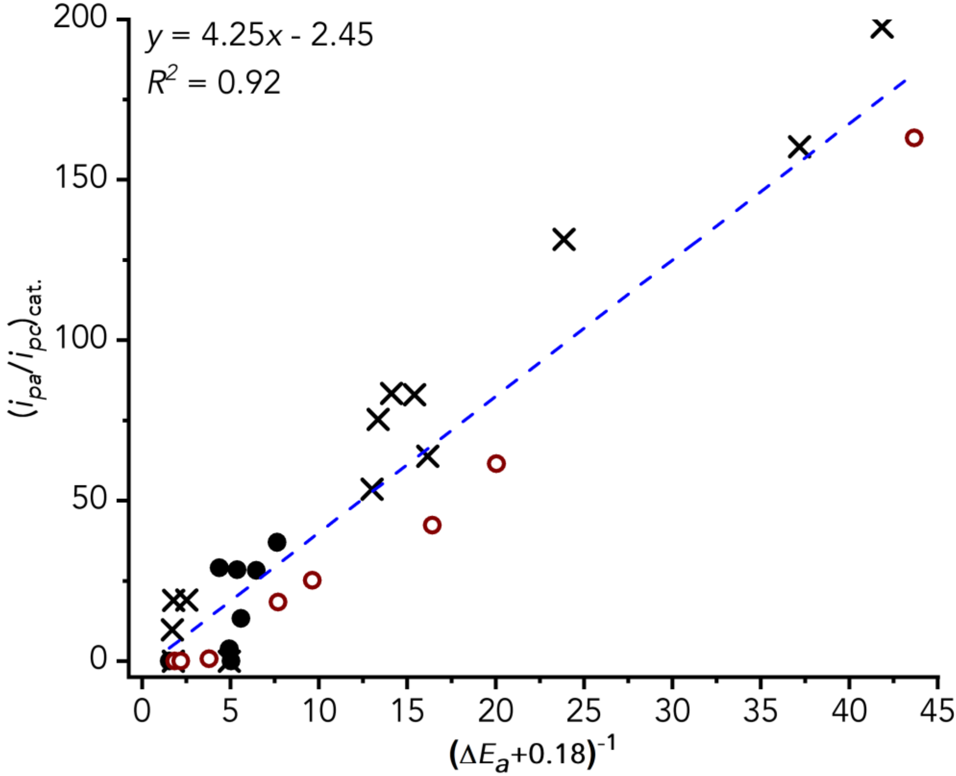
R^1	R^2	R^3	E_{a1}	j_{max}
Pyr	Pyr'	-	0.71	7
Bn	H	-	0.68	-
Me	Me	-	0.72	1487
Me	Me	nPr	0.75	2765
Me	Me	Bn	0.78	2309



E_{a1}= 0.69
j_{max}= 6373

Correlation found between E_2-E_1 and catalysis

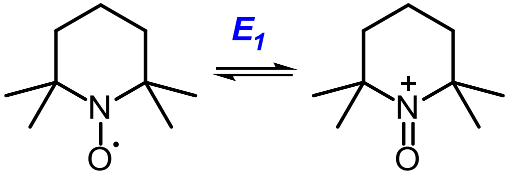


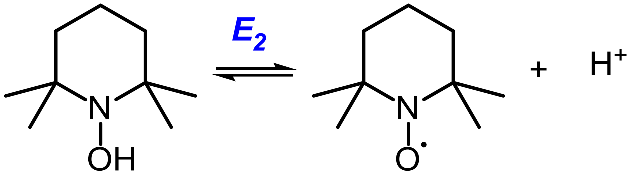


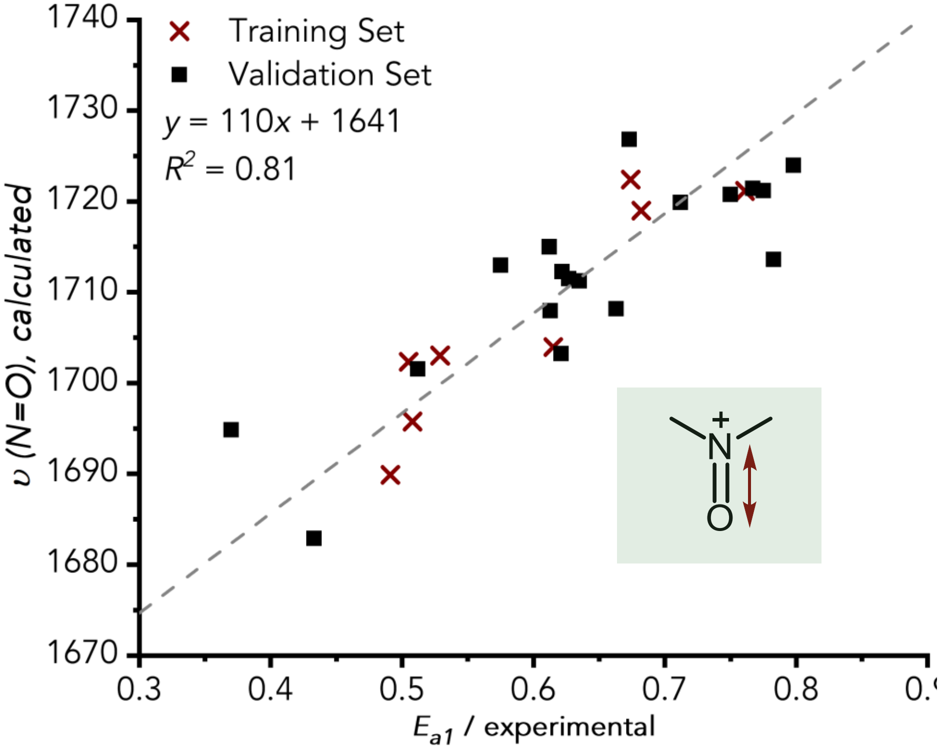
Activity increases if:

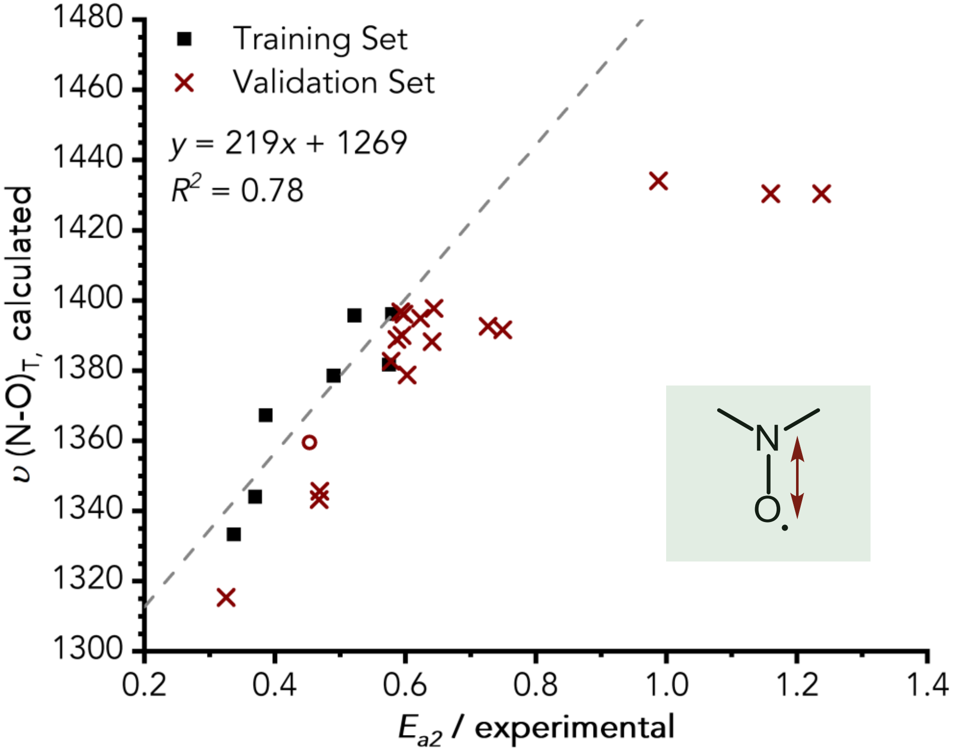
E_{a1} ⬆ *or* E_{a2} ⬇

Correlating both E_1 and E_2 with molecular features

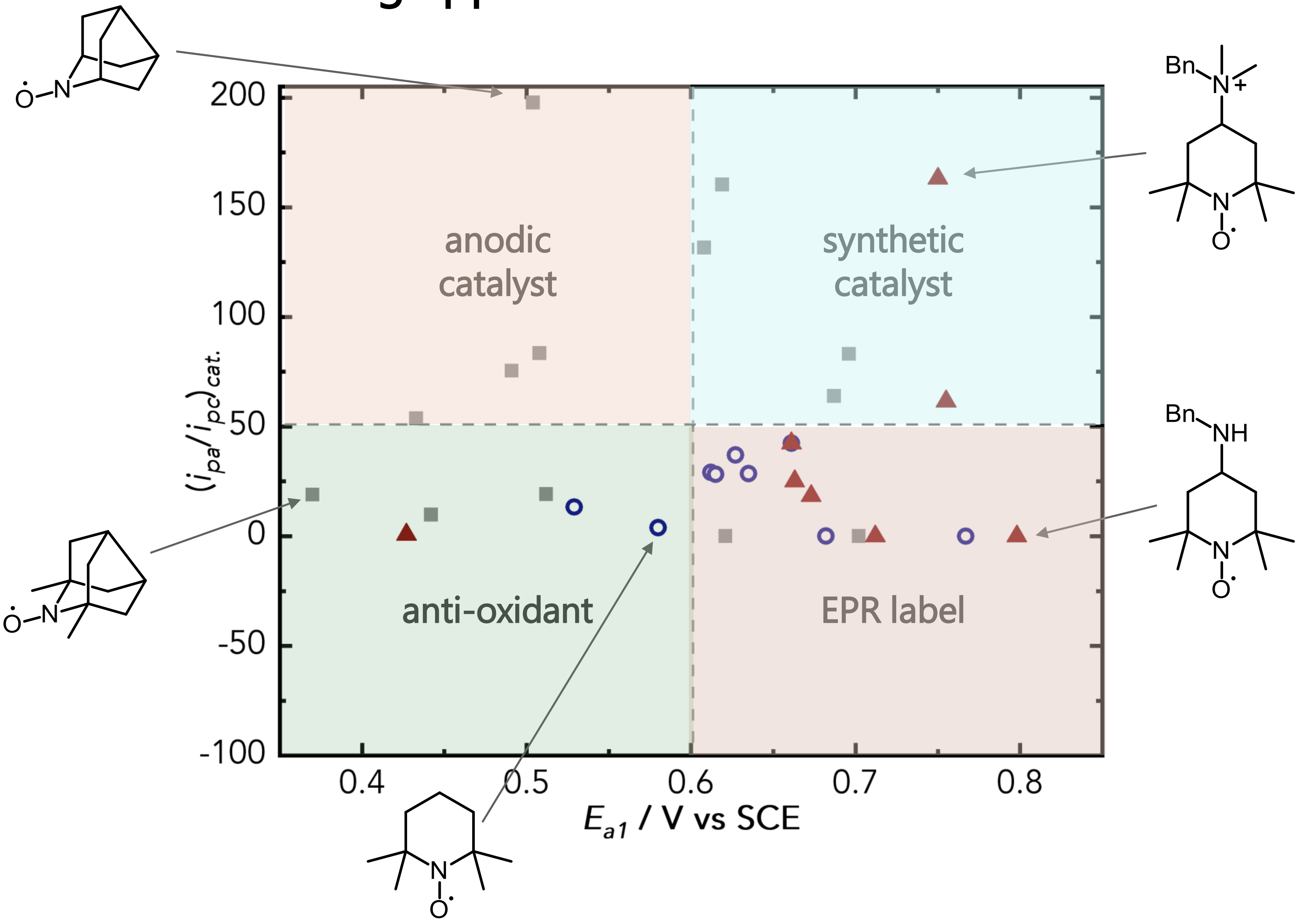








Predicting applications of TEMPO derivatives

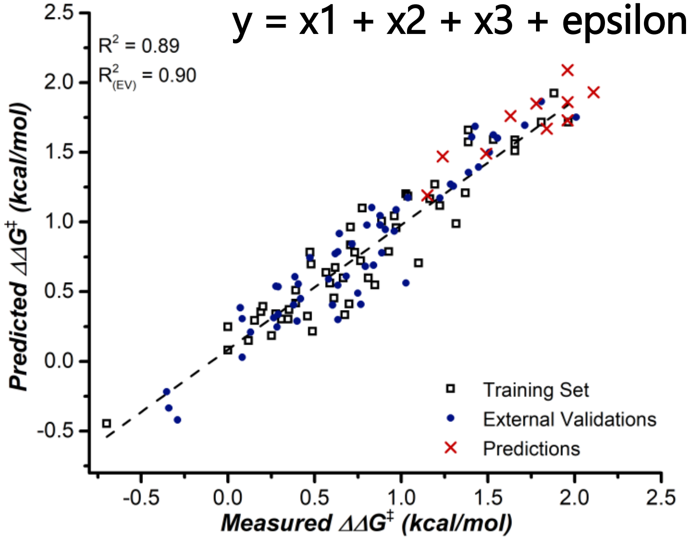


Applications: Lancaster, L.; Hickey, D. P.; Sigman, M. S.; Minteer, S. D.; Wheeldon, I. *Chem. Commun.* **2018**, *54*, 491;
F. Macazo, D.P. Hickey, S. Abdellaoui, M.S. Sigman, S.D. Minteer *Chem. Comm.*, **2017**, *53*, 10310;
D.P. Hickey, R. D. Milton, D. Chen, M. Sigman, S.D. Minteer, *ACS Catalysis*, **2015**, *5*, 5519.

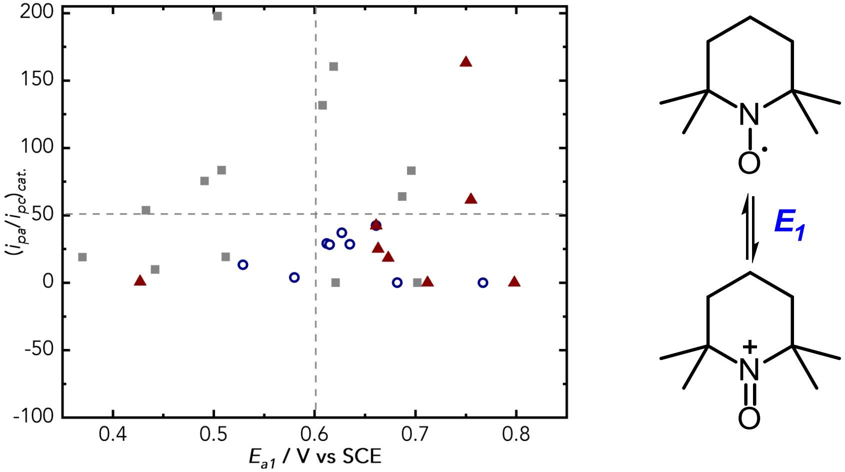
Outline of talk

introduction to data science

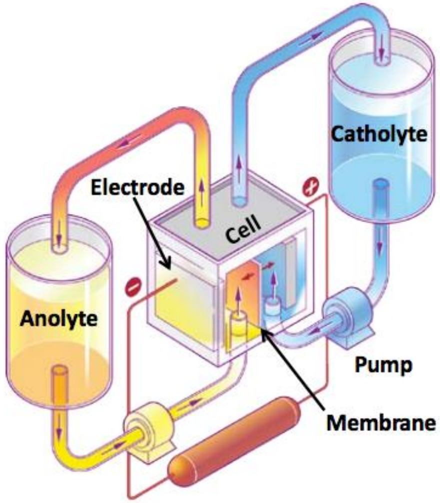
y = x1 + x2 + x3 + epsilon



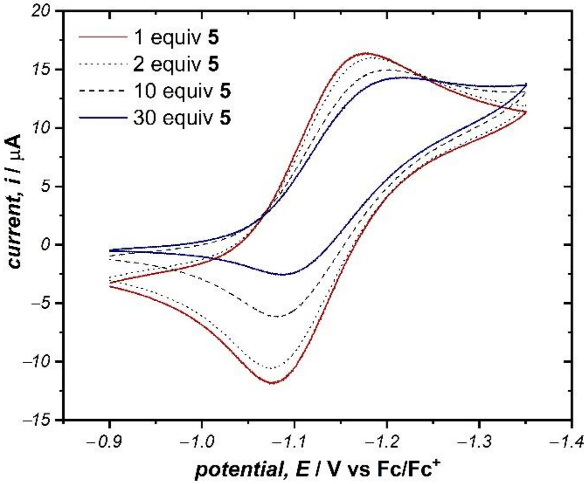
case study 1: electrocatalysis



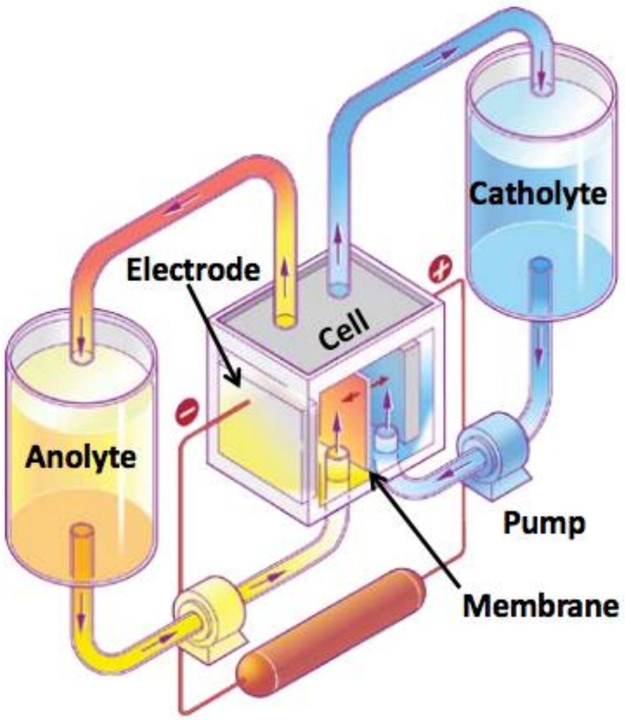
case study 2: flow batteries



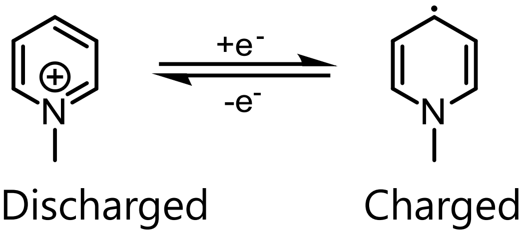
future applications and thoughts



Designing improved electrolytes for flow batteries

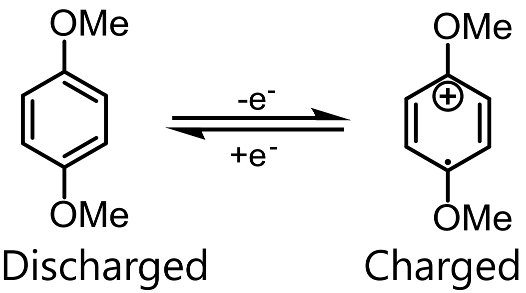


Anolyte:



$E_{1/2}$ = -1.4 V vs Ag/Ag $^{+}$

Catholyte:

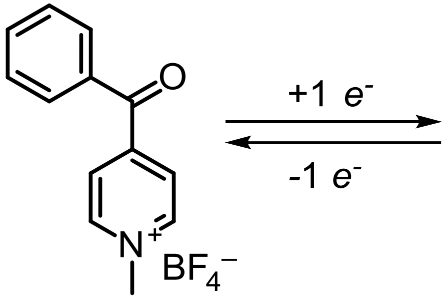


$E_{1/2}$ = 0.8 V vs Ag/Ag $^{+}$

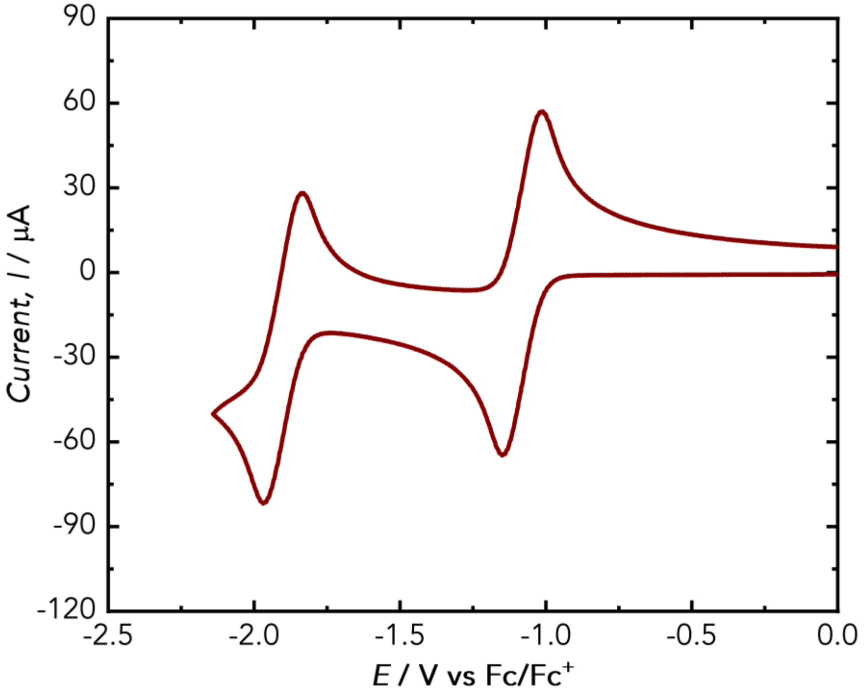
Essential requirements for anolyte/catholyte:

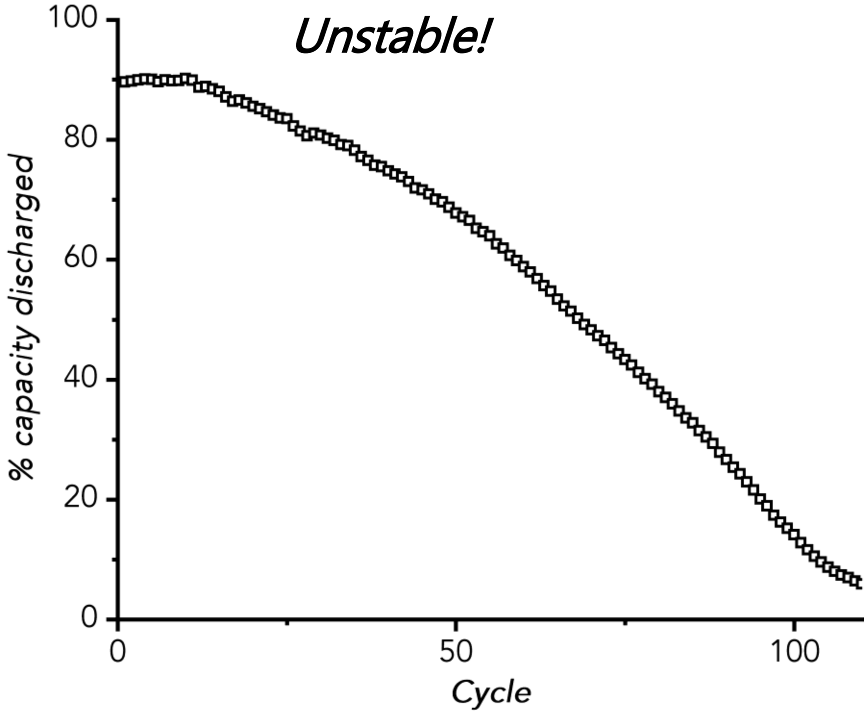
- *Stable in charged and discharged states*
- *Molecules need to store a significant amount of energy*

Identification of a promising anolyte candidate



Low potential (1.12V) reversible redox

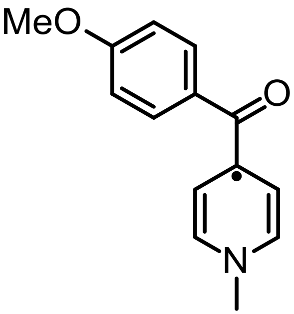



Unstable!

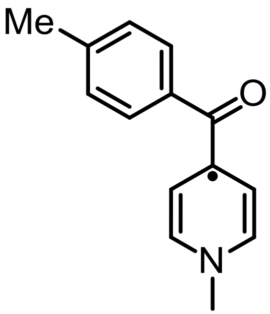
Christo Sevov and Melanie Sanford (U. Michigan) *J. Am. Chem. Soc.* **2015**, *137*,14465

Training set design

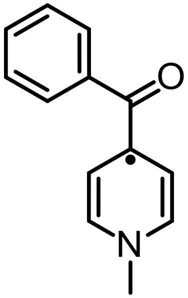
Decreased Energy



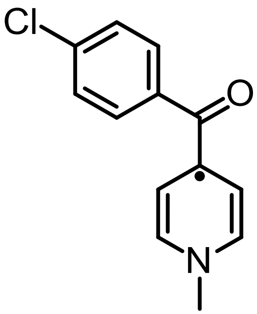
$E_{1/2}$: -1.13
k_{rel}: 117



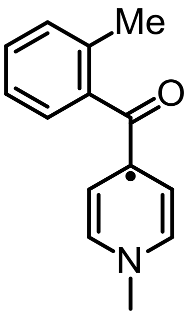
$E_{1/2}$: -1.11
k_{rel}: 106



$E_{1/2}$: -1.08
k_{rel}: 46



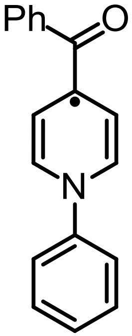
$E_{1/2}$: -1.04
k_{rel}: 19



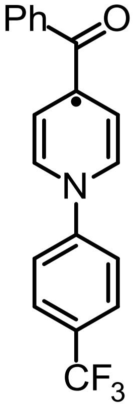
$E_{1/2}$: -1.03
k_{rel}: 18

Increased Persistence

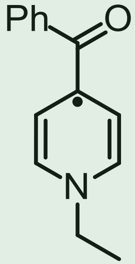
Mechanistic and structural hint



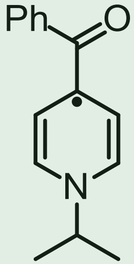
$E_{1/2}$: -0.94
k_{rel}: 1.8



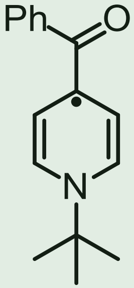
$E_{1/2}$: -0.93
k_{rel}: 1.0



$E_{1/2}$: -1.09
k_{rel}: 34

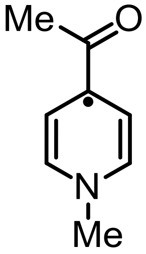


$E_{1/2}$: -1.09
k_{rel}: 2.7

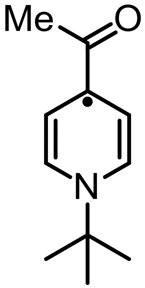


$E_{1/2}$: -1.09
k_{rel}: 0.49

similiar potential
increased persistence



$E_{1/2}$: -1.12
k_{rel}: 42



$E_{1/2}$: -1.12
k_{rel}: 2.1

With Christo Sevov, David Hickey, Sophia Robinson, Sanford group and Minteer group

Developing a statistical model of stability

Parameters evaluated:

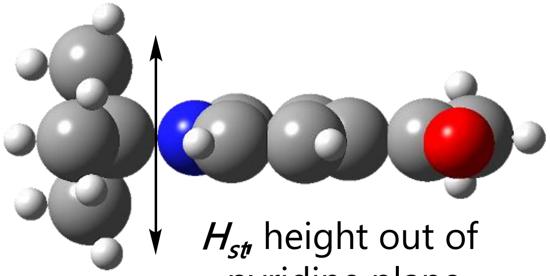
- *IR frequencies, NBO charges,*
- *Sterimol steric parameters*
- *Charton steric parameters*
- *Electrochemical potential (computed)*

Christo Sevov
David Hickey
Sophia Robinson

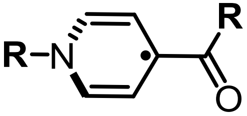
Developing a statistical model of stability

Parameters evaluated:

- *IR frequencies, NBO charges,*
- *Sterimol steric parameters*
- *Charton steric parameters*
- *Electrochemical potential (computed)*



H_{st}, height out of pyridine plane



$$\Delta G_{decomp} = 0.49\ E_{1/2(computed)} + 1.17\ H_{st} + 0.07$$

electronic/ thermodynamic

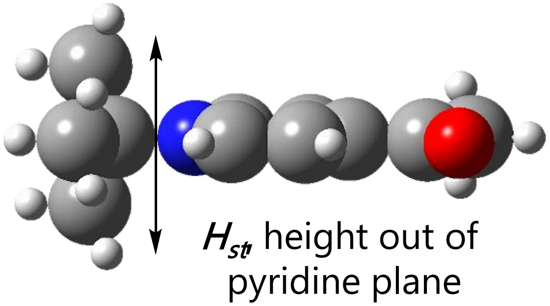
steric/ kinetic

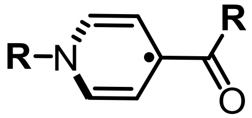
Christo Sevov
David Hickey
Sophia Robinson

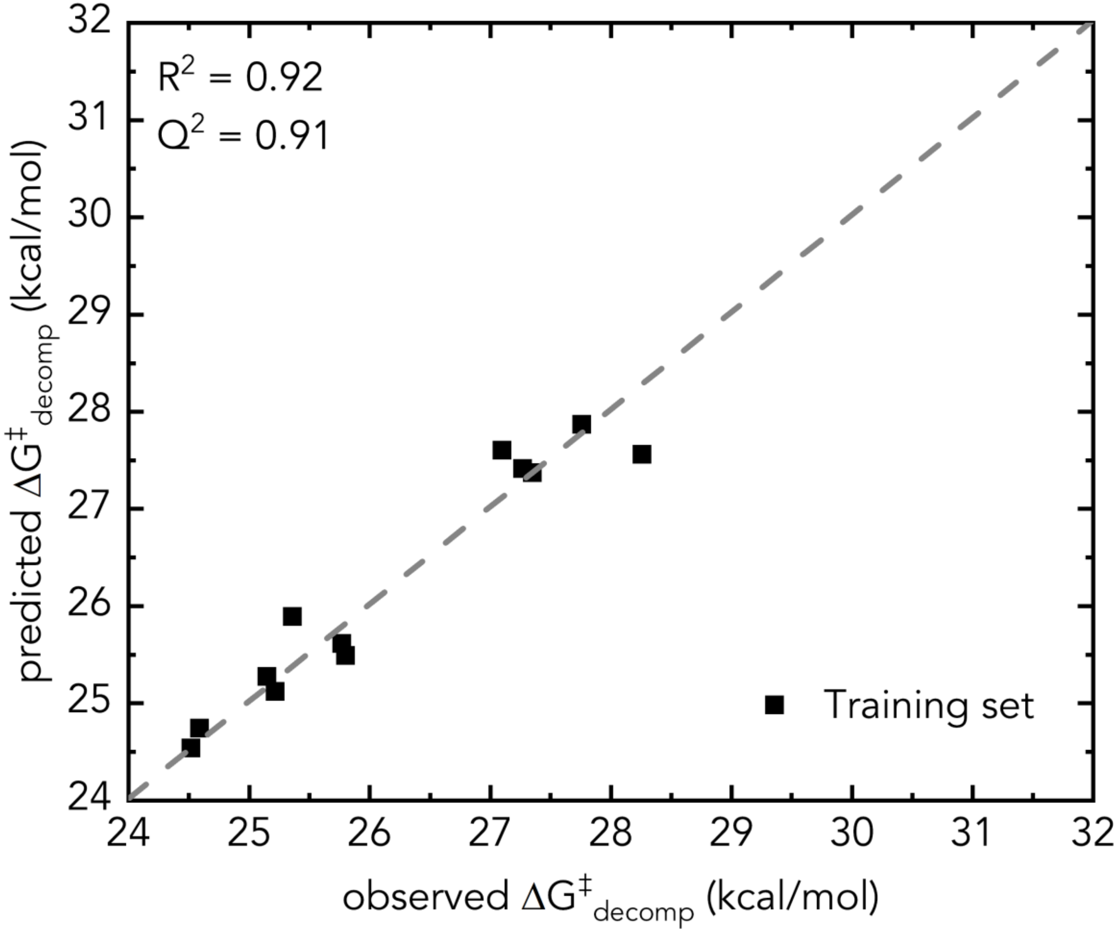
Developing a statistical model of stability

Parameters evaluated:

- *IR frequencies, NBO charges,*
- *Sterimol steric parameters*
- *Charton steric parameters*
- *Electrochemical potential (computed)*







$$\Delta G_{decomp} = 0.49\ E_{1/2(computed)} + 1.17\ H_{st} + 0.07$$

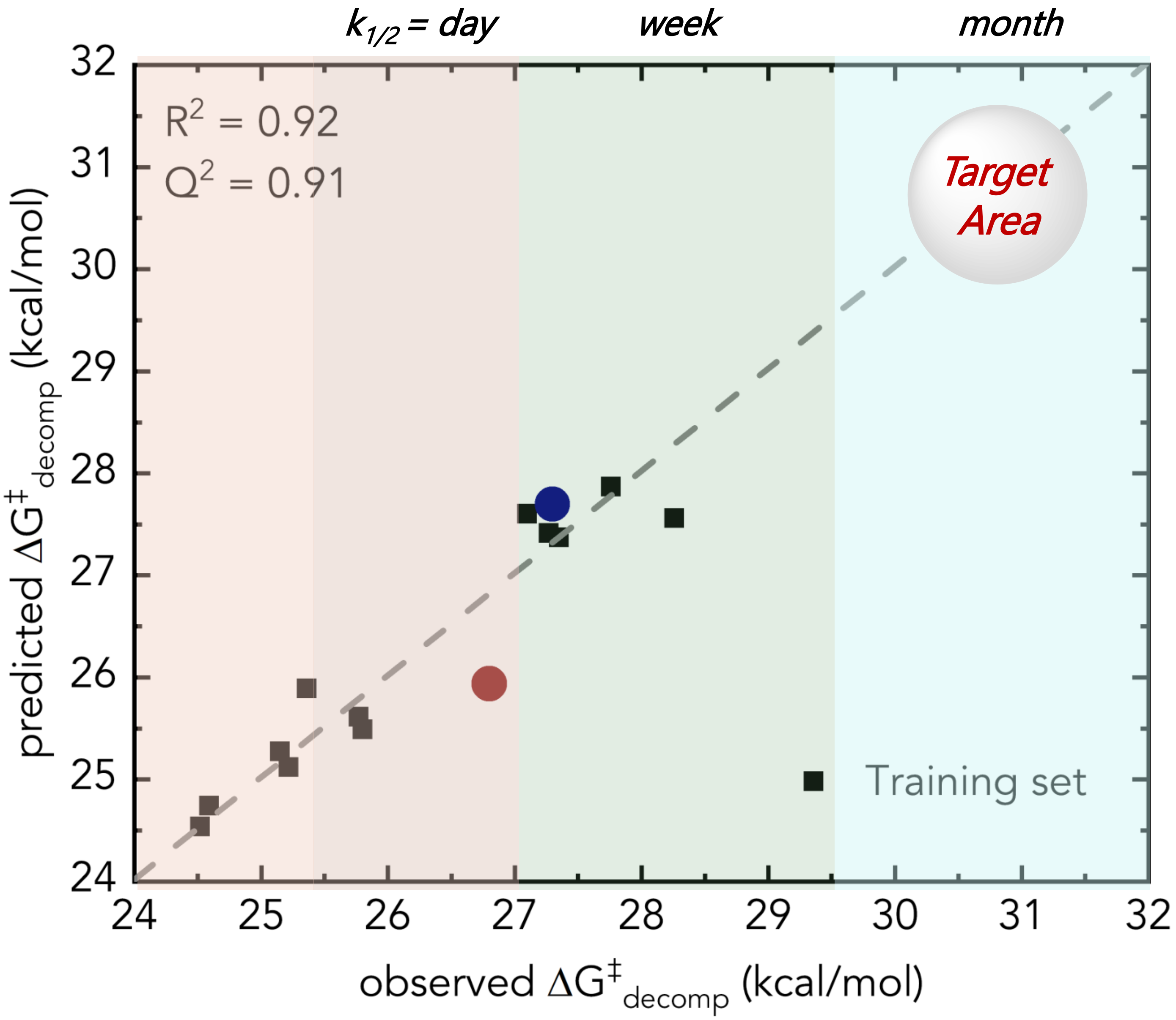
electronic/ thermodynamic

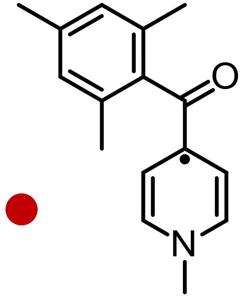
steric/ kinetic

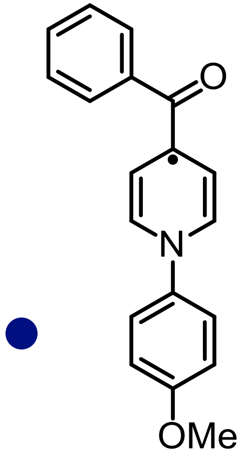
Christo Sevov
David Hickey
Sophia Robinson

Interpolations

Initial predictions: interpolations

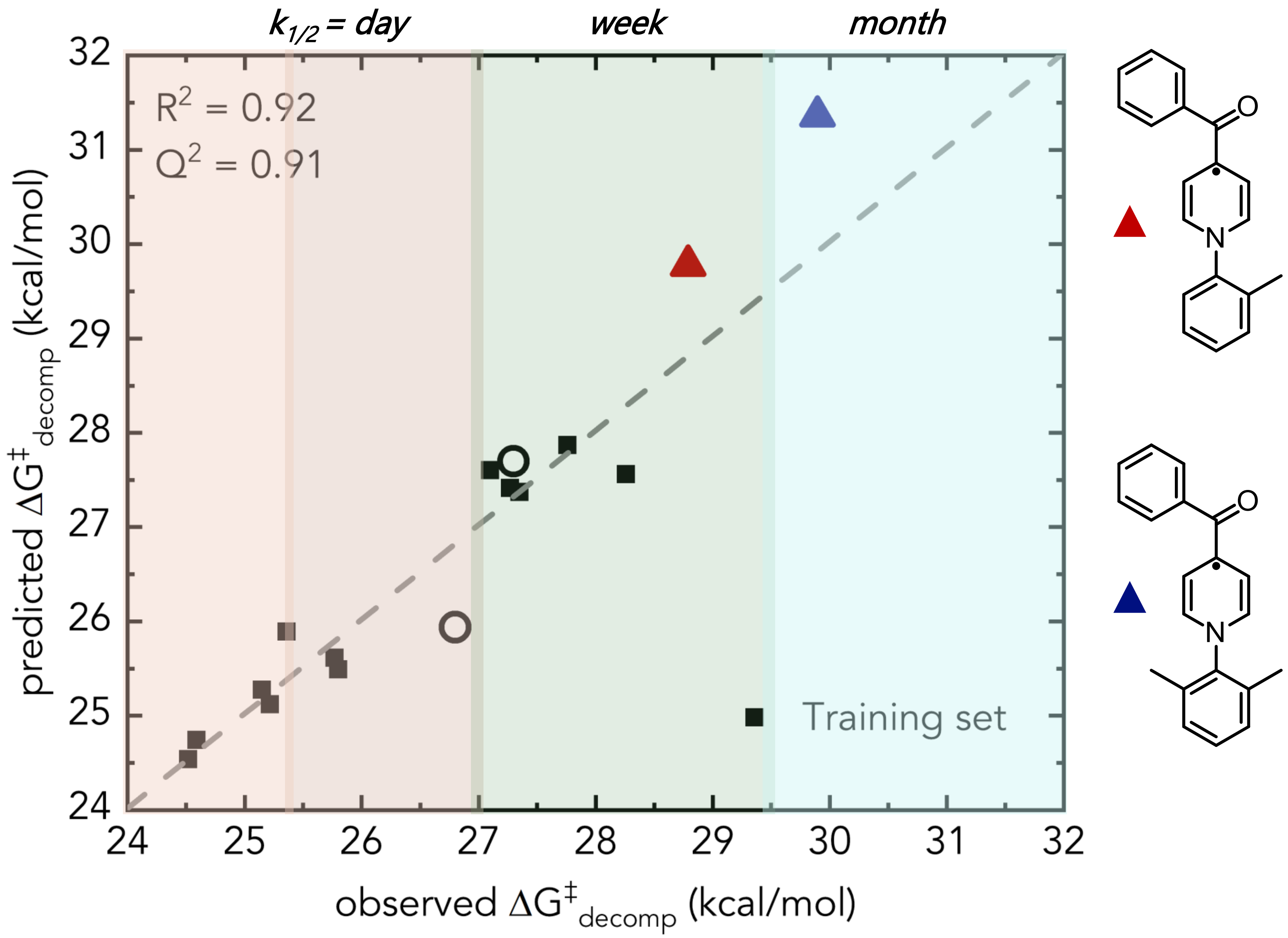






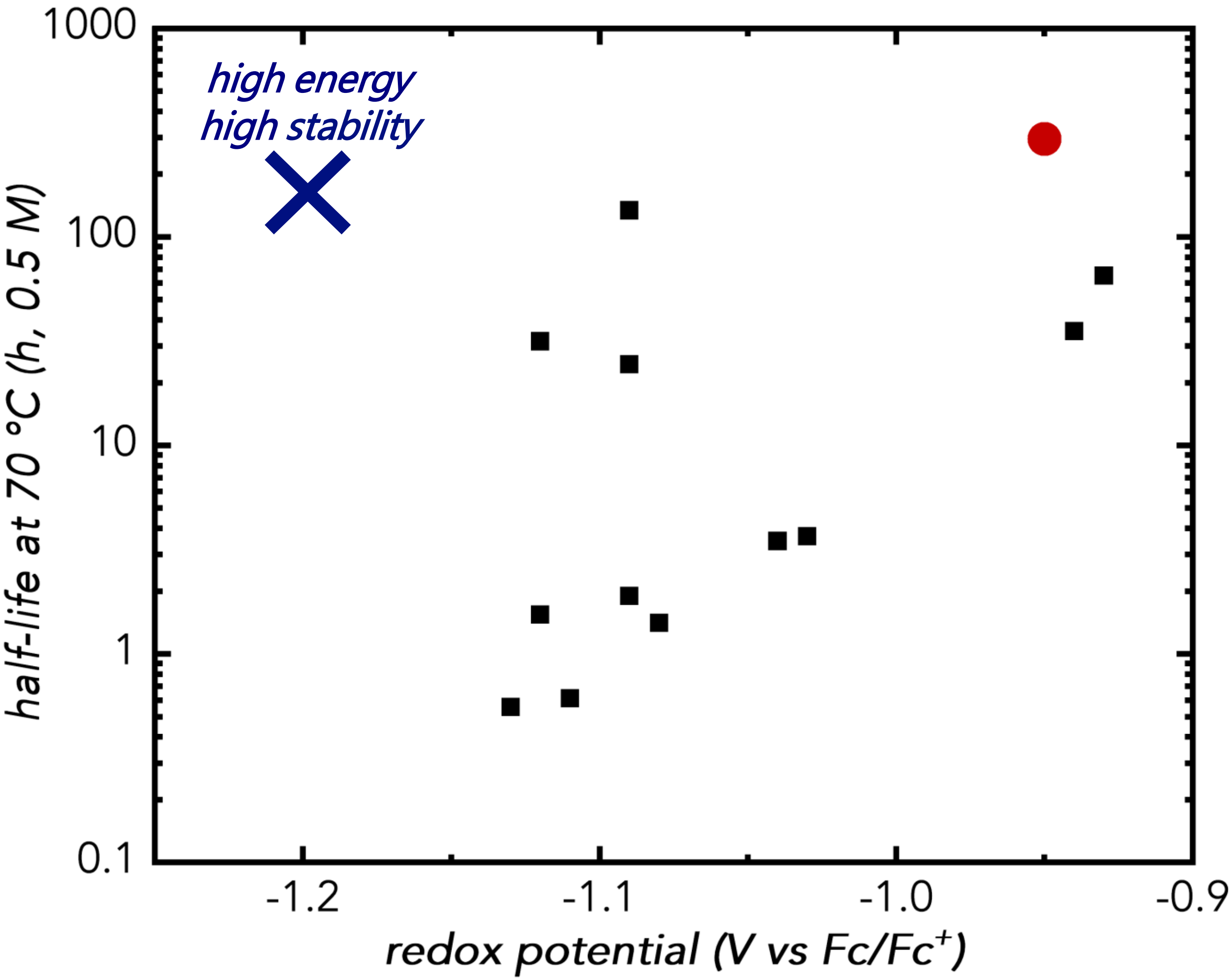
Christo Sevov
David Hickey
Sophia Robinson

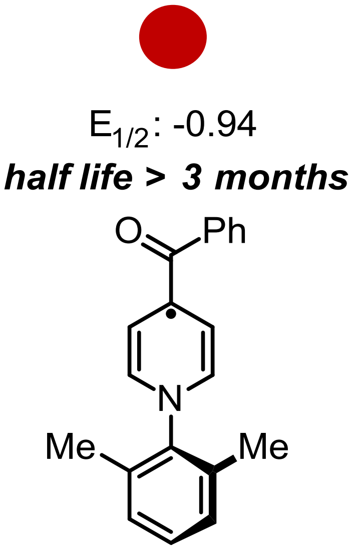
Predictions



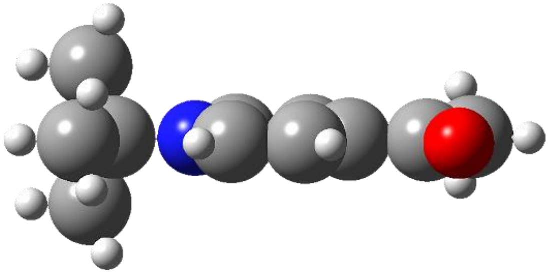
Predicted and validated 1000-fold increase in stability!!!

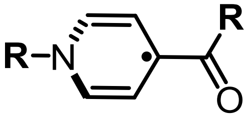
Stability versus potential: still correlated

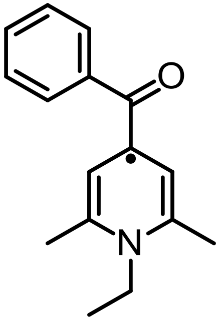




Stability versus potential: still correlated







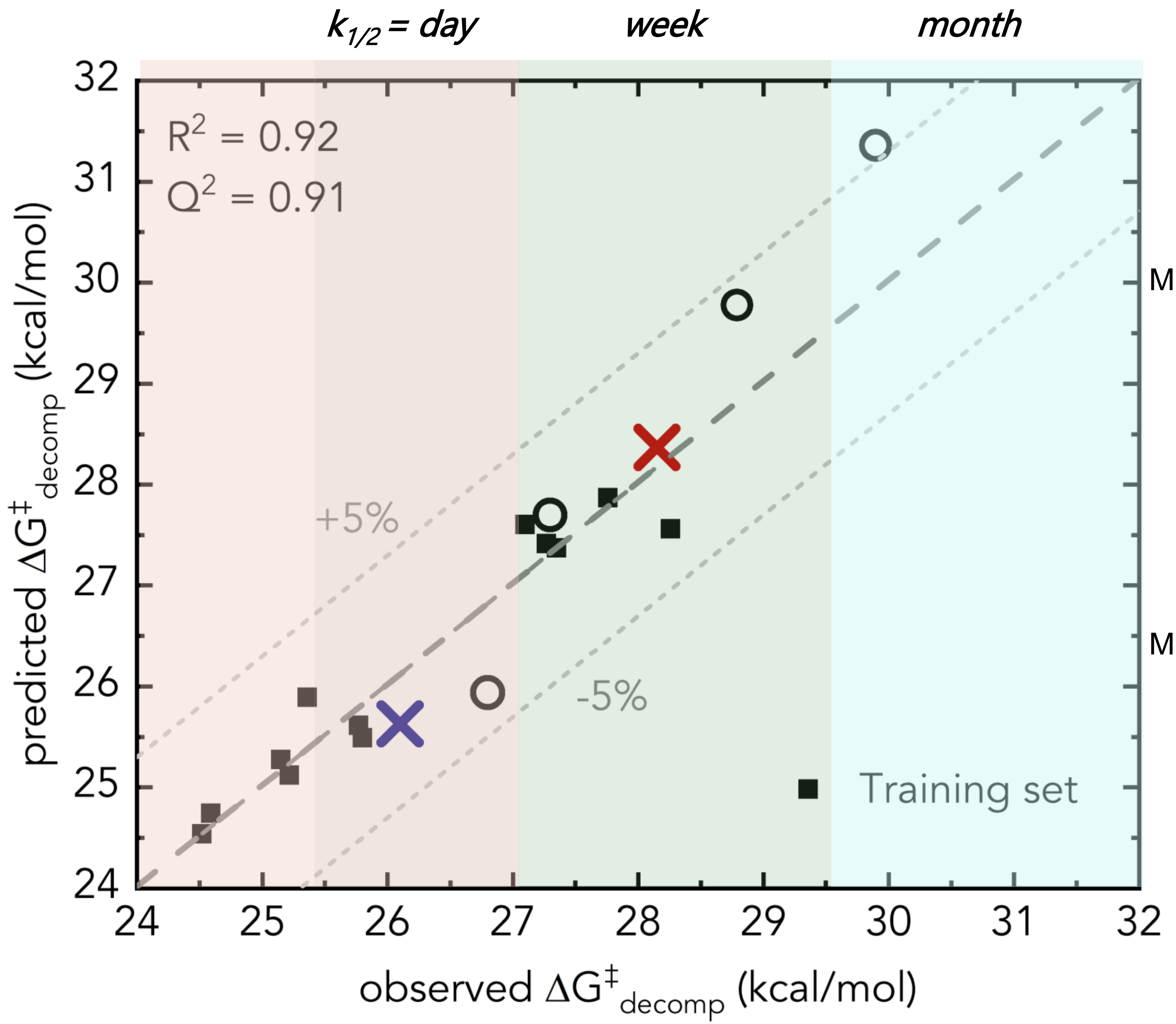
predicted potential: -1.23 V

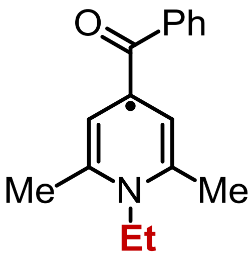
predicted $\Delta G^{\ddagger}$ =28.4 kcal/mol

Protect the 2,6-postions through substitution:

Model predicts good stability and higher redox potential

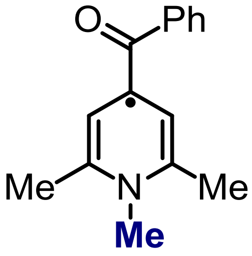
A bit more about the model





$E_{1/2}$: -1.21
k_{rel}: 0.56
persistent
& low potentia

✕



$E_{1/2}$: -1.21
k_{rel}: 12.0

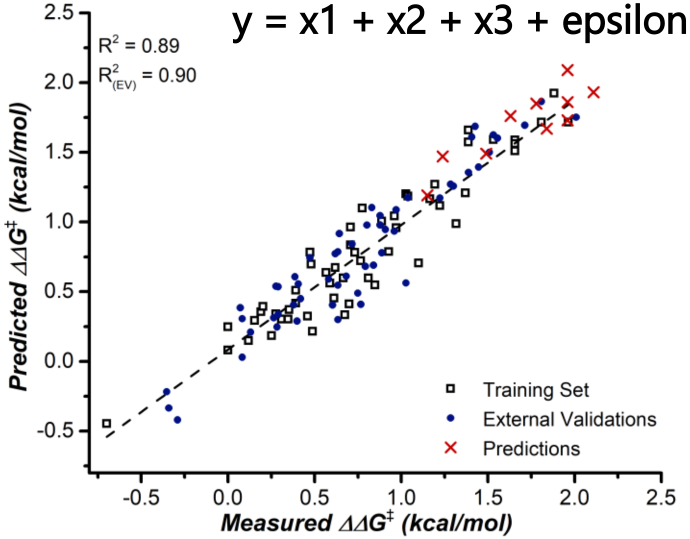
✕

Sevov, C. S.; Hickey, D. P.; Cook, M. E.; Robinson, S. G.; Barnett, S.; Minteer, S. D.; Sigman, M. S.; Sanford, M. S. *JACS* **2017**, *139*, 2924.

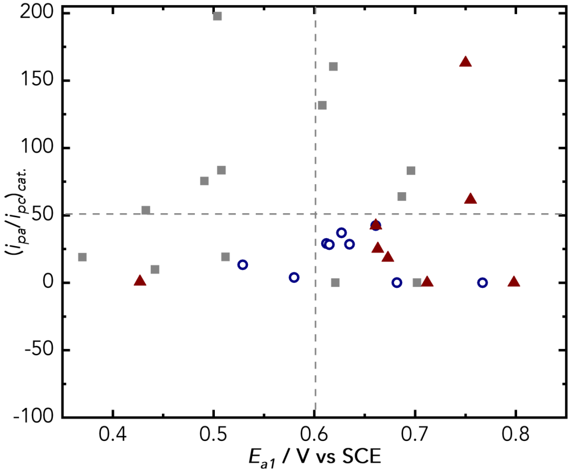
Outline of talk

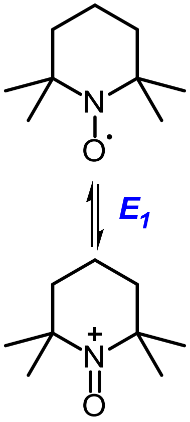
introduction to data science

y = x1 + x2 + x3 + epsilon

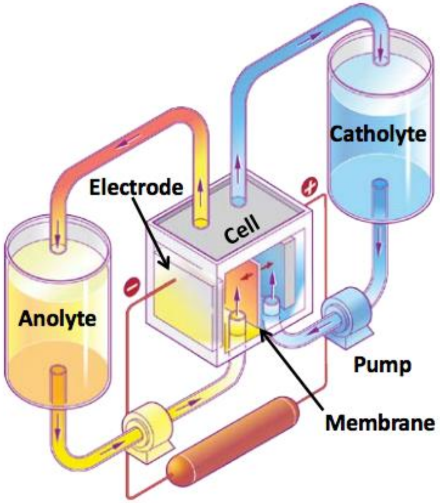


case study 1: electrocatalysis

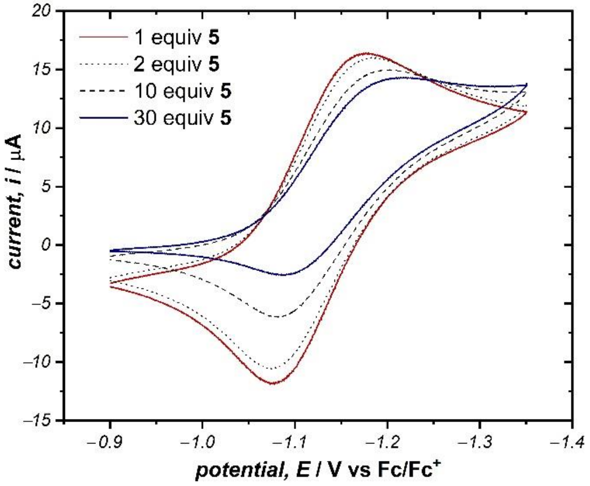




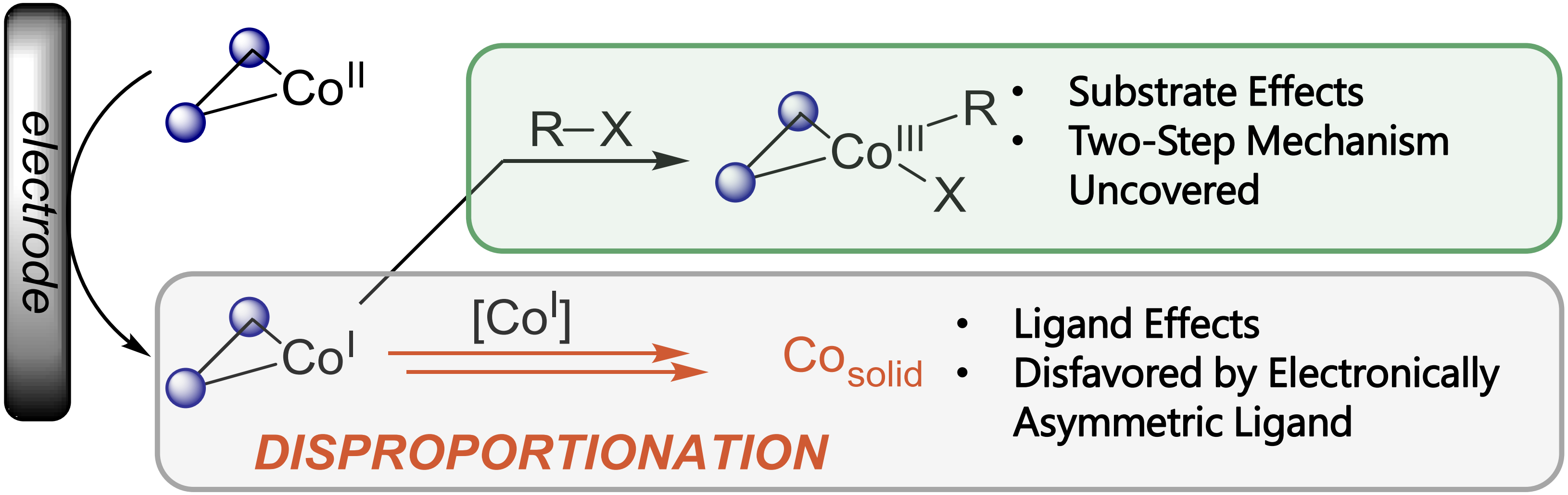
case study 2: flow batteries



future applications and thoughts

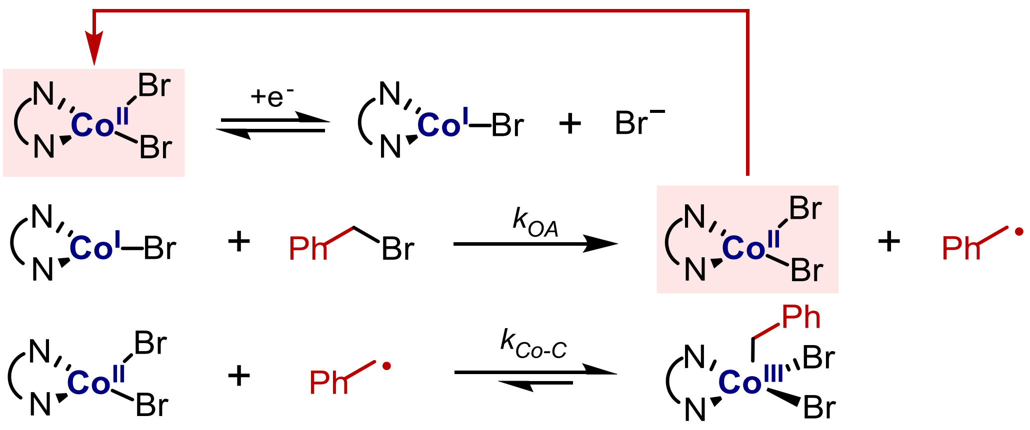


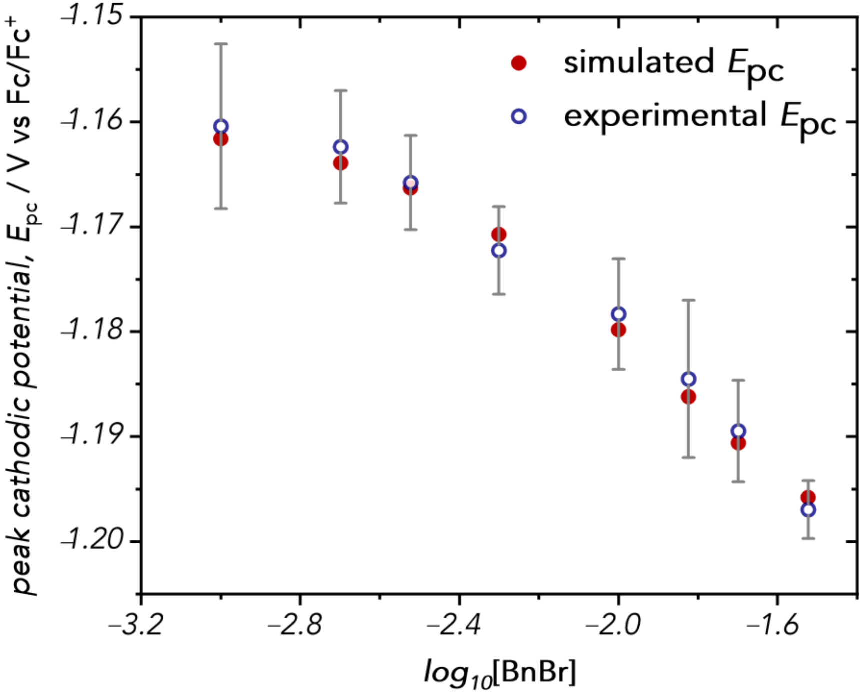
Electroanalytical tools in organometallic electrocatalysis



- Substrate Effects
- Two-Step Mechanism Uncovered

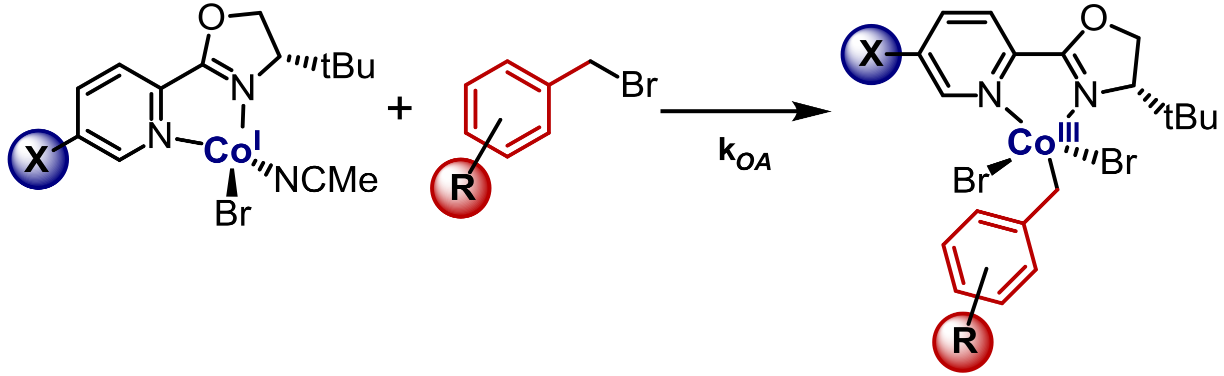
- Ligand Effects
- Disfavored by Electronically Asymmetric Ligand

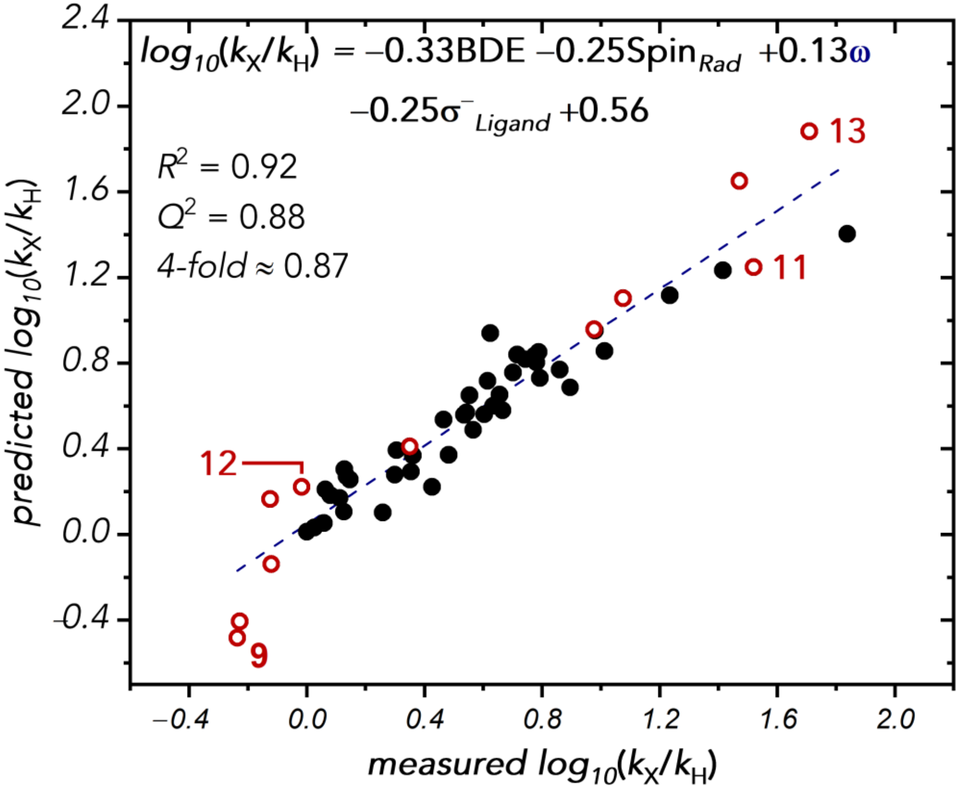


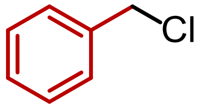


Hickey, D. P.; Sandford, C.; Rhodes, Z.; Gensch, T.; Fries, L. R.; Sigman, M. S.; Minteer, S. D. *J. Am. Chem. Soc.* **2019**, *141*, 1382.
Sandford, C.; Fries, L. R.; Ball, T.; Minteer, S. D. Sigman, M. S. *J. Am. Chem. Soc.* **2019**, *141*, ASAP

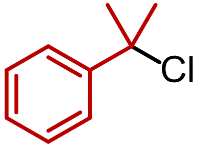
Electroanalytical tools enable data science



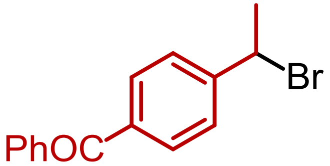




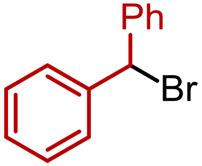
9
(with ligand **2**)



12
(with ligand **2**)



11
(with ligand **1**)



13
(with ligand **1**)

Hickey, D. P.; Sandford, C.; Rhodes, Z.; Gensch, T.; Fries, L. R.; Sigman, M. S.; Minteer, S. D. *J. Am. Chem. Soc.* **2019**, *141*, 1382.
Sandford, C.; Fries, L. R.; Ball, T.; Minteer, S. D. Sigman, M. S. *J. Am. Chem. Soc.* **2019**, *141*, ASAP

Final thoughts and data science in organic chemistry

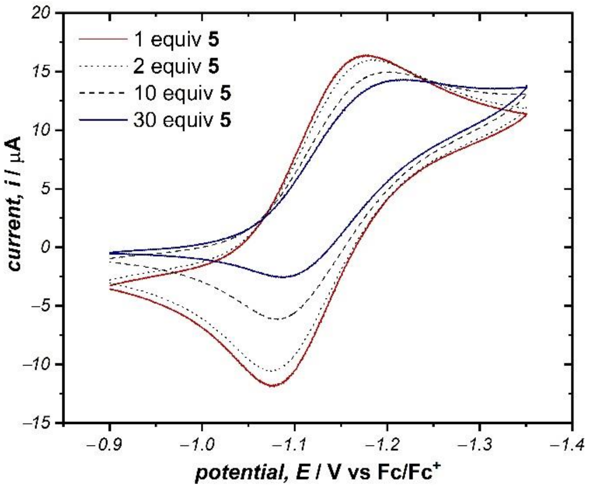
possible applications

- structure function relationships
- synthetic route planning
- reaction optimization protocols
- autonomous reaction optimization
- novel reaction prediction
- feature predictions

the issues

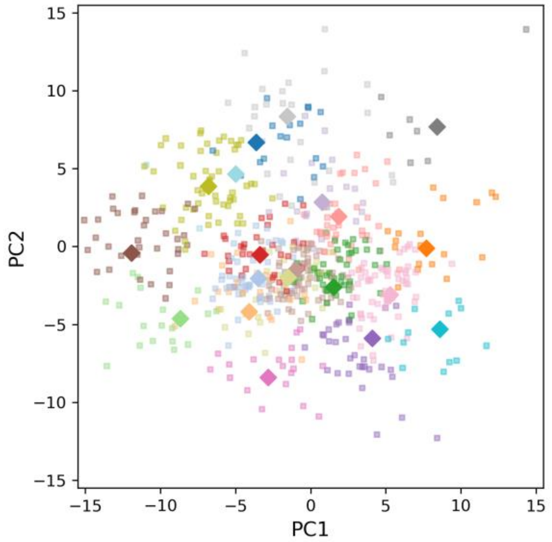
- data quality & quantity
- data bias
- feature selection
- not trained in data science
- AI versus ML versus data science
- OVERHYPED!

Data starved for data science



- electroanalytic tools are data rich
- kinetics, thermodynamics, etc.

Design of data sets



- use molecular features
- clustering to statistically orient data

SPRING 2019

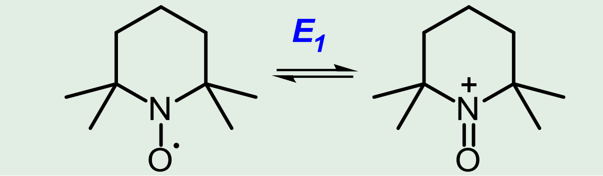
Active Collaborators

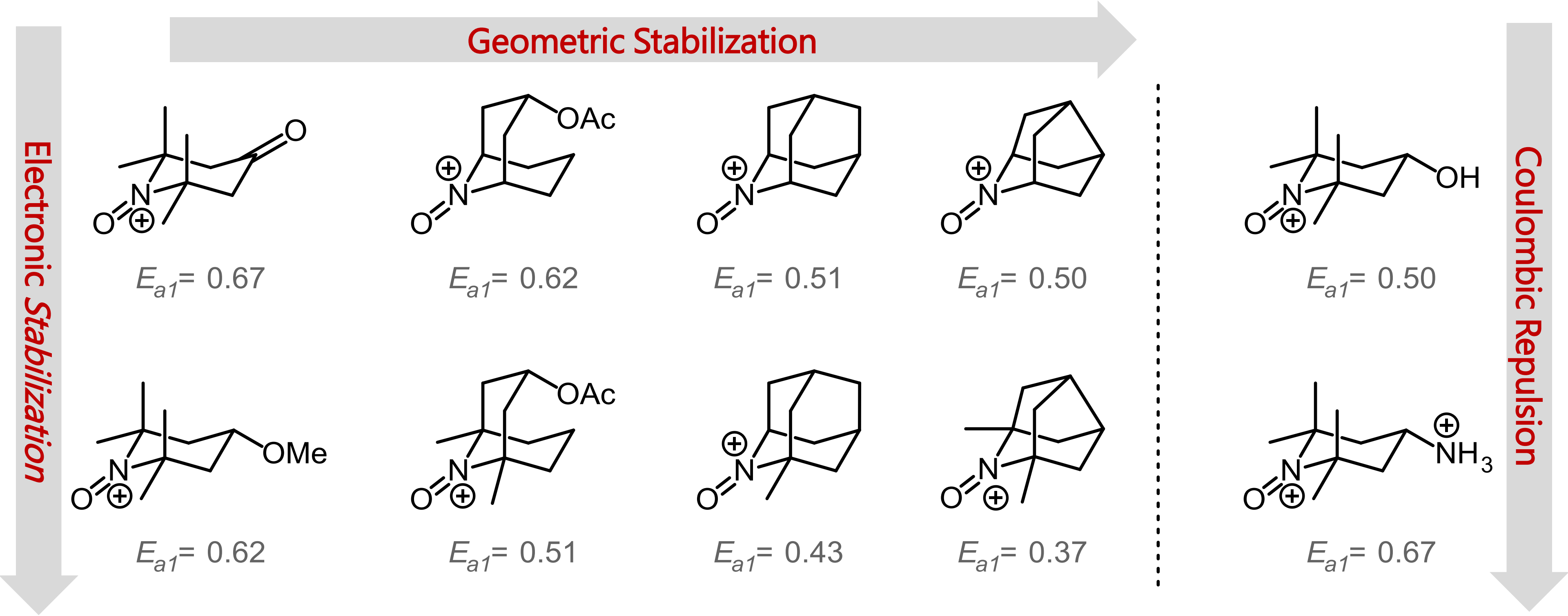
Alán Aspuru-Guzik (Toronto)
Mark Biscoe (CCNY)
Christophe Copéret (ETH)
Miquel Costas (iQCC)
Huw Davies (Emory)
Abigail Doyle (Princeton)
Justin Du Bois (Stanford)
Jason Hein (UBC)
Todd Hyster (Princeton)
Eric Jacobsen (Harvard)
Song Lin (Cornell)
Neal Mankad (UIC)
Scott Miller (Yale)
Shelley Minteer (Utah)
Olaf Wiest (Norte Dame)
Rob Paton (CSU)
Robert Phipps (Cambridge)
Sarah Reisman (Caltech)
Melanie Sanford (Michigan)
Corinna Schindler (Michigan)
Dean Toste (UC Berkeley)

Financial Support

NSF (CHE-1925607 – C-CAS)
NSF (CHE-1763436)
NIH (NIGMS, GM121383 with Toste/Miller)
NIH (NIGMS, GM63540)
DOE (JCSER) with Minteer/Sanford
ARO (MURI) with Shelley Minteer
NSF (CHE-1205646 – NSF CCHF)
Merck Research Labs, Millipore Sigma
Novartis (NIBR), Pfizer, Genentech

Mechanistic insight





Hickey, D. P.; Schiedler, D. A.; Matanovic, I.; Doan, P. V. §; Atanassov, P.;
Minteer, S. D.; Sigman, M. S. *J. Am. Chem. Soc.* **2015**, *137*, 16179-16186.