Multi-Objective, Machine-Learning Assisted First-Principles Design of Transition Metal Complexes for Redox Couples

Jon Paul Janet [1] Heather Kulik [1]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology

Applications of Data Science in Molecular Sciences I

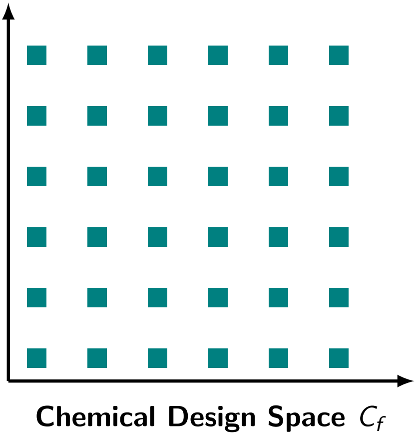
11.11.19

Motivation: chemical discovery

How can we design new materials using computers?

Motivation: chemical discovery

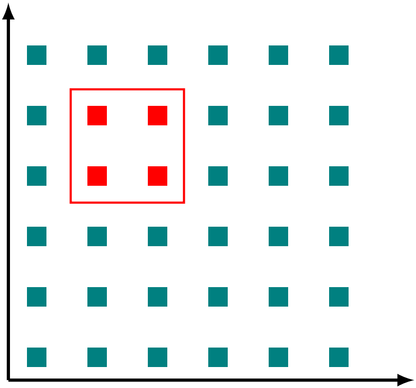
How can we design new materials using computers?



Motivation: chemical discovery

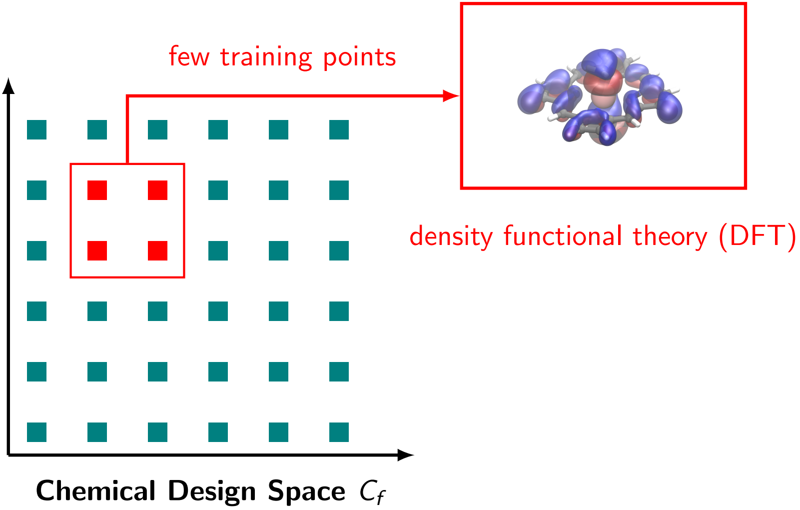
How can we design new materials using computers?

few training points

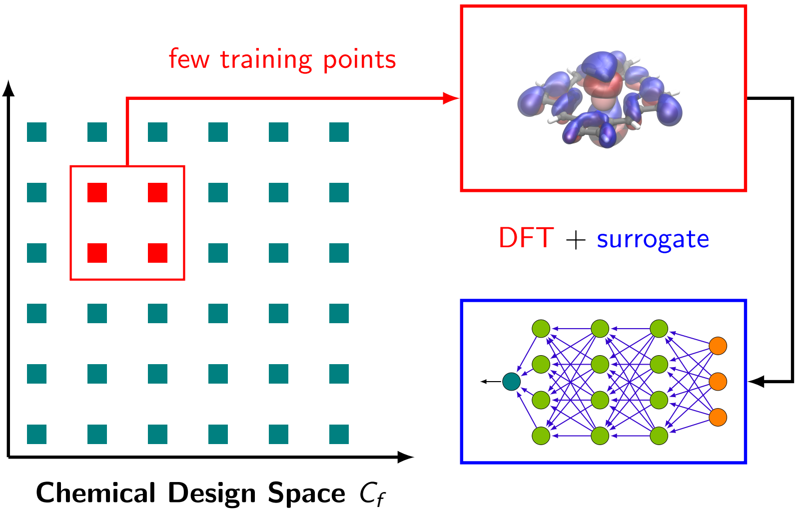


Chemical Design Space C_f

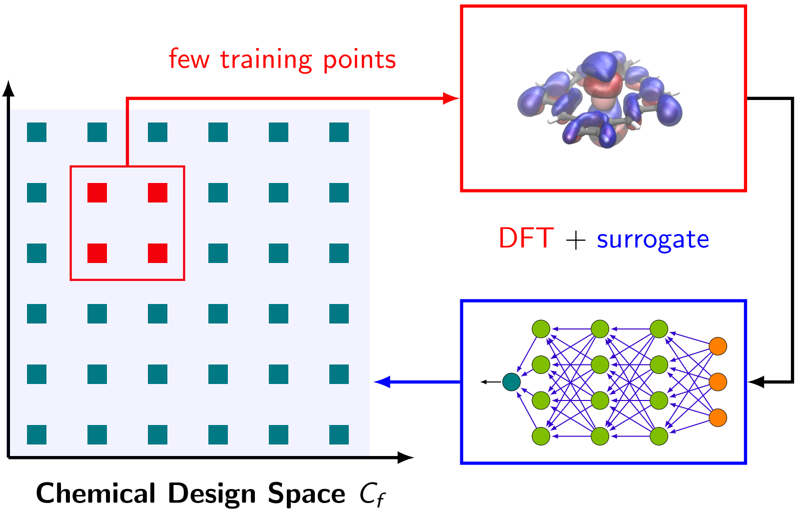
Motivation: chemical discovery



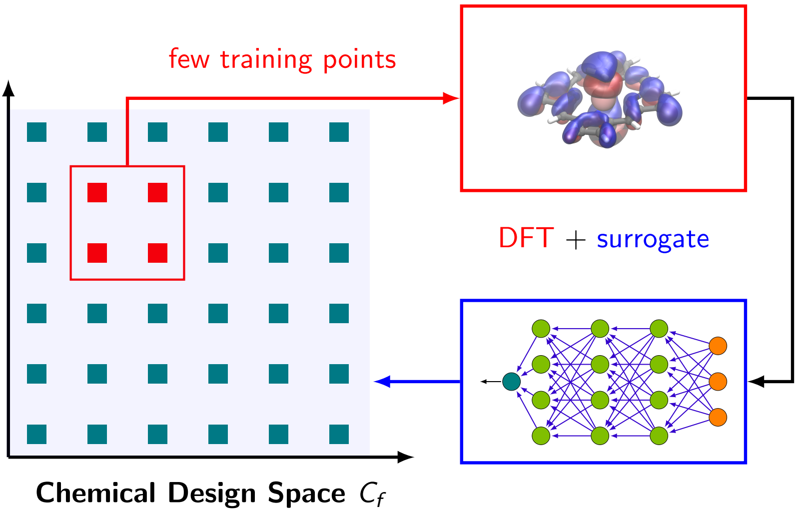
Motivation: chemical discovery



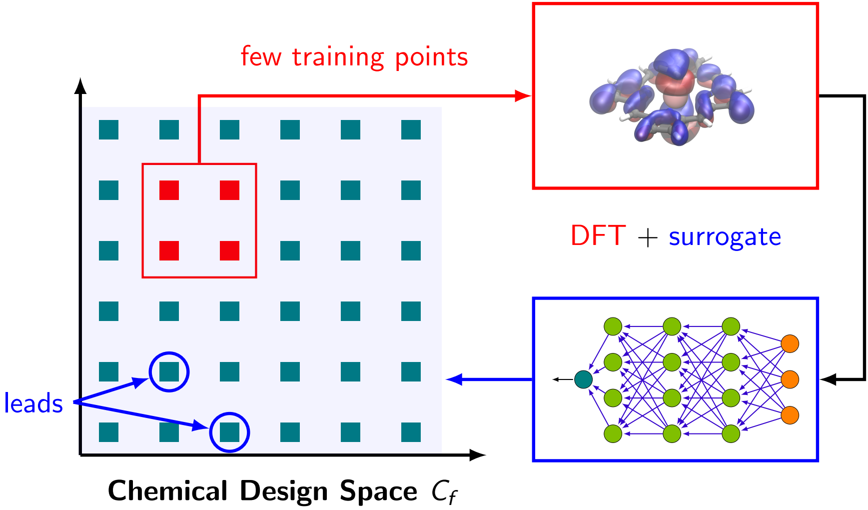
Motivation: chemical discovery



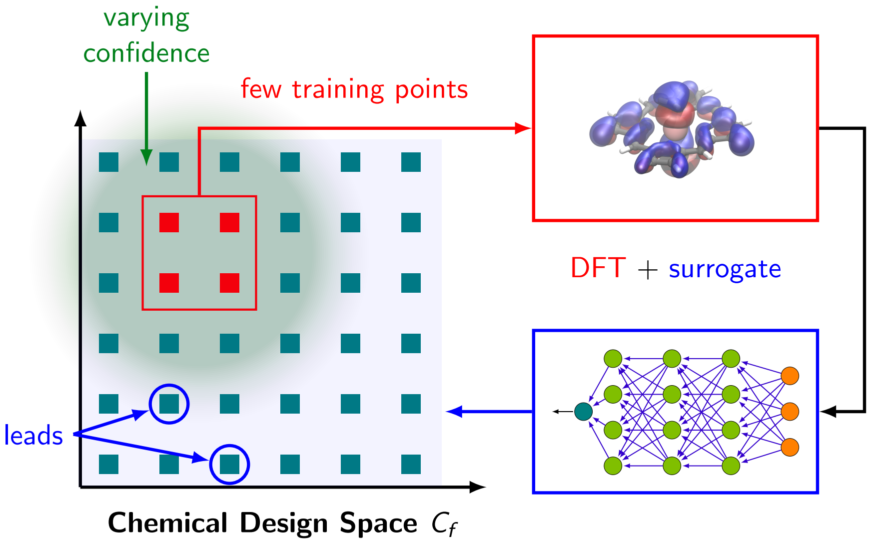
Motivation: chemical discovery



Motivation: chemical discovery

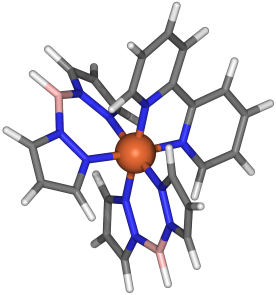


Motivation: chemical discovery



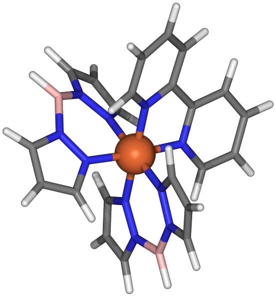
Transition metal complexes

Transition metal complexes



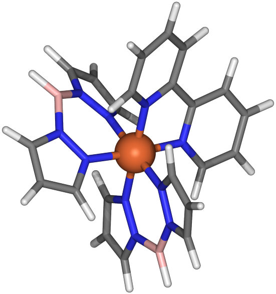
Transition metal complexes

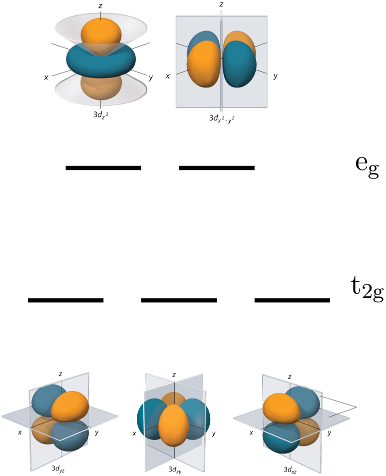
Difficult chemistry with multiple spin & oxidation states!



Transition metal complexes

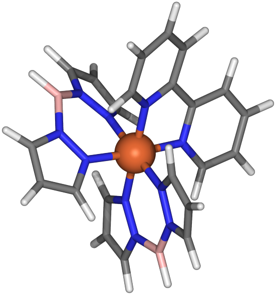
Difficult chemistry with multiple spin & oxidation states!

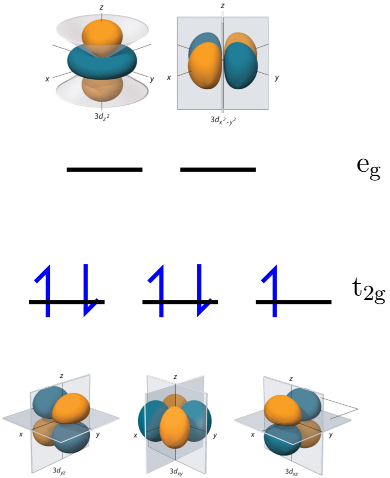




Transition metal complexes

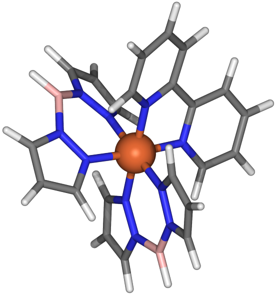
Difficult chemistry with multiple spin & oxidation states!

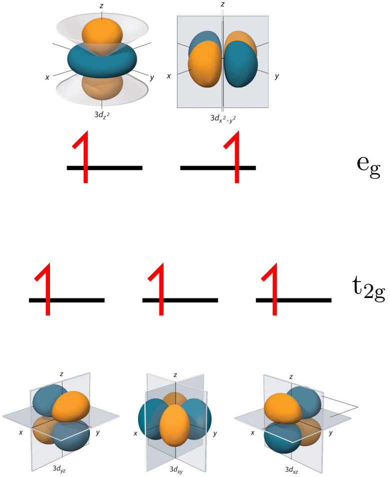




Transition metal complexes

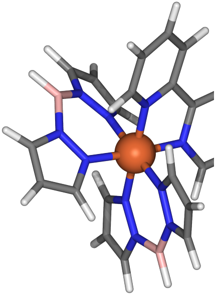
Difficult chemistry with multiple spin & oxidation states!

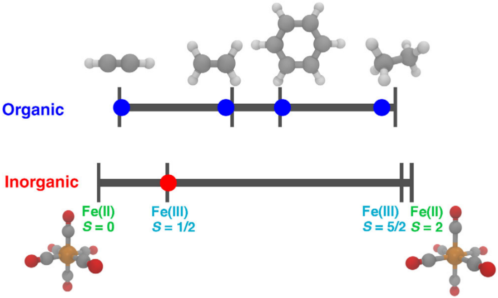




Transition metal complexes

Difficult chemistry with multiple spin & oxidation states!





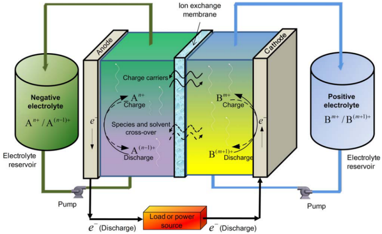
Kulik, HJ., *WIREs Comput Mol Sci.*, e1439, 2019

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:

Redox flow batteries

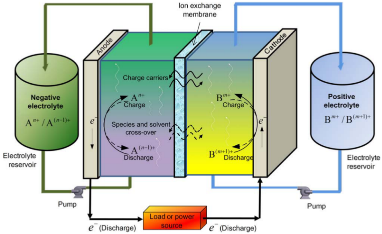
Redox flow batteries (RFBs) are a promising option for scalable energy storage:



Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:

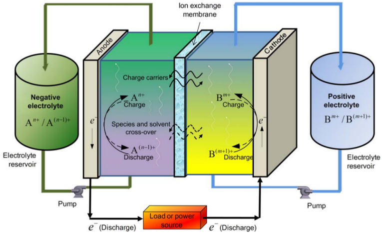


Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Transition metal complexes make attractive redox couples for RFBs

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:



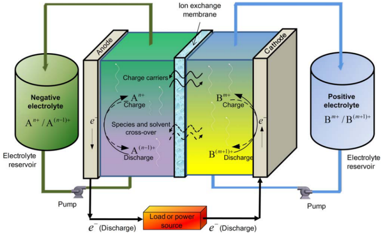
Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Transition metal complexes make attractive redox couples for RFBs

- good ion stability (compared to organics)

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:



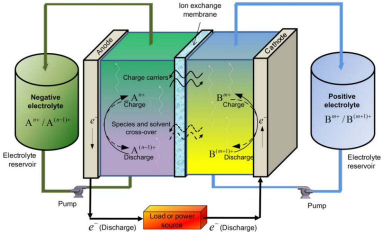
Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Transition metal complexes make attractive redox couples for RFBs

- good ion stability (compared to organics)
- good range of redox potentials available

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:



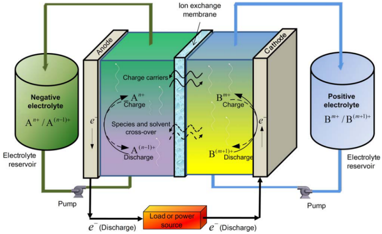
Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Transition metal complexes make attractive redox couples for RFBs

- good ion stability (compared to organics)
- good range of redox potentials available
- solubility is an issue!

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:



Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

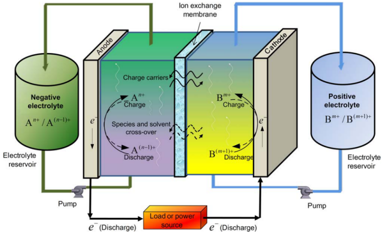
Transition metal complexes make attractive redox couples for RFBs

- good ion stability (compared to organics)
- good range of redox potentials available
- solubility is an issue!

$$E_{\text{cell}} = 0.5 \times \Delta G_{\text{solv}} \times C \times n \times F$$

Redox flow batteries

Redox flow batteries (RFBs) are a promising option for scalable energy storage:



Perry, M.L. and Adam, Z., *J. Electrochem. Soc.*, 163(1):A5064–A5067, 2018.

Transition metal complexes make attractive redox couples for RFBs

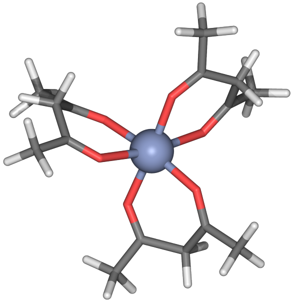
- good ion stability (compared to organics)
- good range of redox potentials available
- solubility is an issue!

$$E_{\text{cell}} = 0.5 \times \Delta G_{\text{solv}} \times C \times n \times F$$

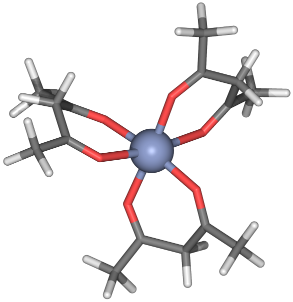
We need complexes that have high redox potential **and** good solubility

First principles methods

First principles methods



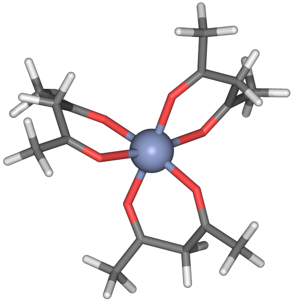
First principles methods



Metals:

Cr	Mn	Fe	Co

First principles methods

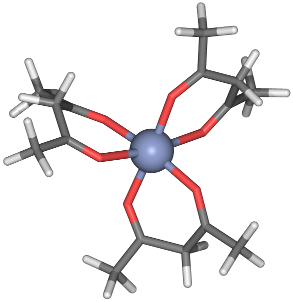


$$M(III)_s + e^- \xrightarrow[\Delta G_{solv}]{} M(II)_s$$

Metals:

Cr	Mn	Fe	Co

First principles methods

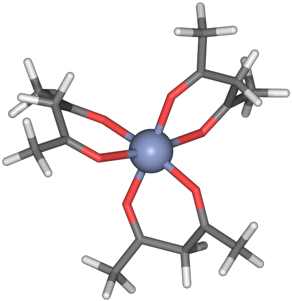


$$\begin{tikzcd}
\mathrm{M(III)}_g + e^- \arrow[r, "\Delta G_{gas}"] \arrow[d, "\Delta G_{s,III}"'] & \mathrm{M(II)}_g \arrow[d, "\Delta G_{s,II}"] \\
\mathrm{M(III)}_s + e^- \arrow[r, "\Delta G_{solv}"'] & \mathrm{M(II)}_s
\end{tikzcd}$$

Metals:

Cr	Mn	Fe	Co

First principles methods



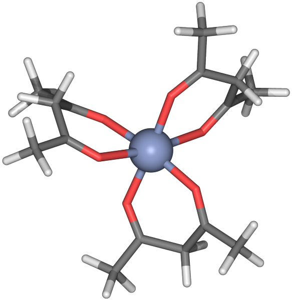
$$\begin{tikzcd}
M(III)_g + e^- \arrow[r, "\Delta G_{gas}"] \arrow[d, "\Delta G_{s,III}"'] & M(II)_g \arrow[d, "\Delta G_{s,II}"] \\
M(III)_s + e^- \arrow[r, "\Delta G_{solv}"'] & M(II)_s
\end{tikzcd}$$

$$\boxed{\log P \approx \log \frac{\Delta G_{s,II,octanol}}{\Delta G_{s,II,water}}}$$

Metals:

Cr	Mn	Fe	Co

First principles methods



$$\begin{array}{ccc} \mathrm{M(III)}_g + e^- & \xrightarrow{\Delta \mathrm{G_{gas}}} & \mathrm{M(II)}_g \\ \Delta \mathrm{G_{s,III}} \downarrow & & \downarrow \Delta \mathrm{G_{s,II}} \\ \mathrm{M(III)}_s + e^- & \xrightarrow[\Delta \mathrm{G_{solv}}]{} & \mathrm{M(II)}_s \end{array}$$

$$\boxed{\log \mathrm{P} \approx \log \frac{\Delta \mathrm{G_{s,II,octanol}}}{\Delta \mathrm{G_{s,II,water}}}}$$

Metals:

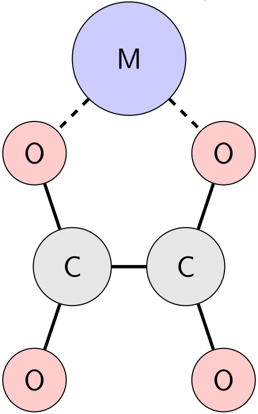
Cr	Mn	Fe	Co

- B3LYP-like DFT
- gas phase optimizatons
- LANL2DZ/6-31G*
- COSMO solvent, $\varepsilon = 78.39$ or 10.30
- high- and low-spin

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:

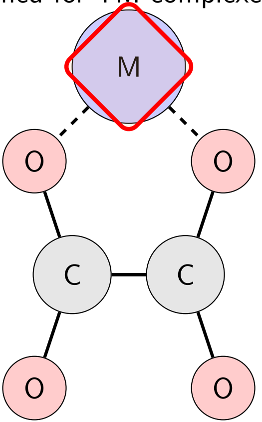


Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:

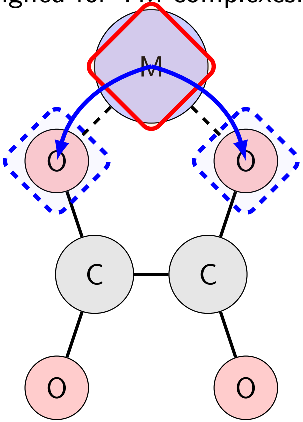


Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:



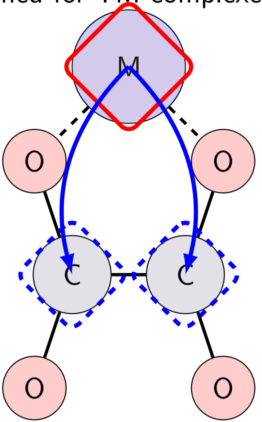
$$d_1 : \sum_{M,O} Z_M Z_O$$

Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:



$$d_1 : \sum_{M,O} Z_M Z_O$$

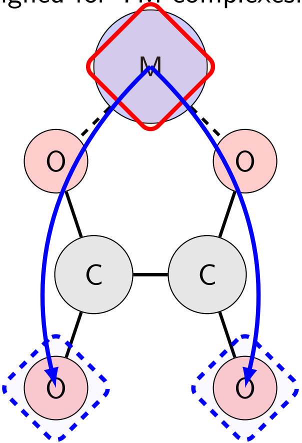
$$d_2 : \sum_{M,C} Z_M Z_C$$

Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:



$d_1 : \sum_{M,O} Z_M Z_O$

$d_2 : \sum_{M,C} Z_M Z_C$

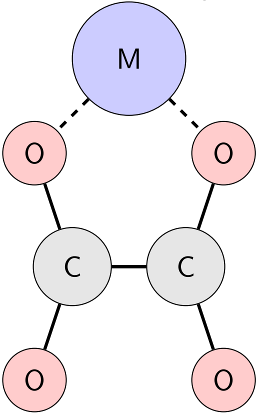
$d_3 : \sum_{M,O} Z_M Z_O$

Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:



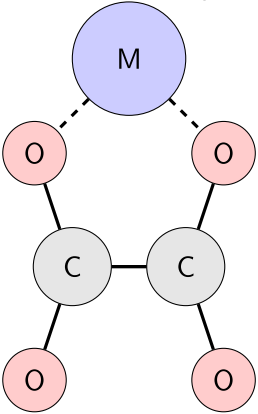
Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Regression:

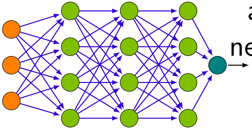
Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:



Regression:



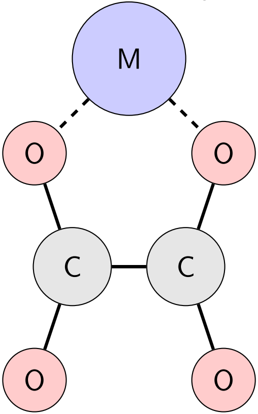
artifical neural networks (ANNs)

Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

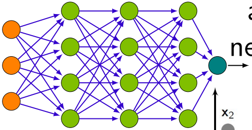
Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:

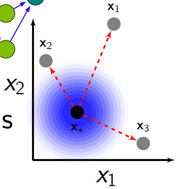


Regression:



artifical neural networks (ANNs)

Kernel methods (KRR)

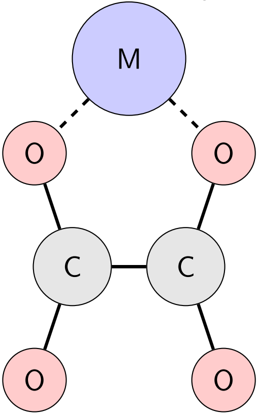


Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Machine learning methods

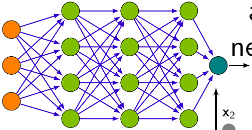
Featurization:

Graph-based features (RACs) designed for TM complexes:

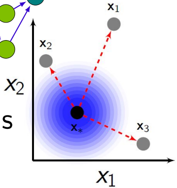


Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

Regression:

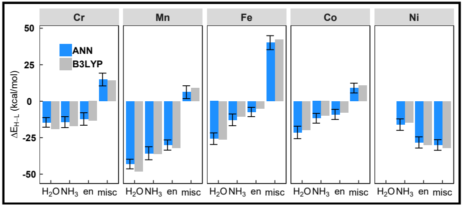


artifical neural networks (ANNs)



Kernel methods (KRR)

spin splitting energies

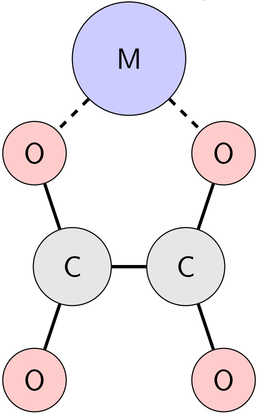


Janet, J.P. and Kulik, H.J., *Chem. Sci.*, 8:5137–5152, 2017.

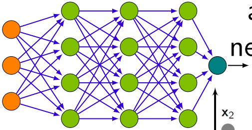
Machine learning methods

Featurization:

Graph-based features (RACs) designed for TM complexes:

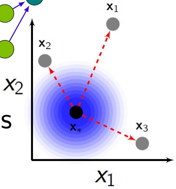


Regression:

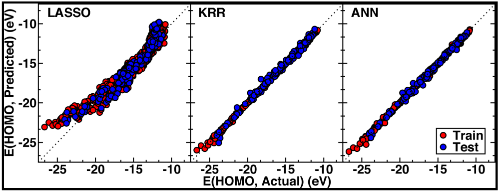


artifical neural networks (ANNs)

Kernel methods (KRR)



frontier orbital properties

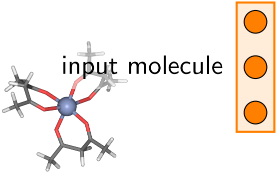


Janet, J.P., and Kulik, H.J., *J. Phys. Chem. A*, 121(46):8939–8954, 2017.

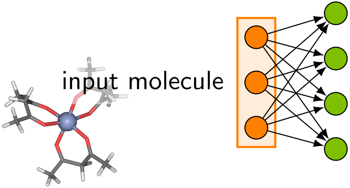
Nandy, A. *et al.*, *Ind. Eng. Chem. Res.*, 57(42):13973–13986, 2018.

Latent distance similarity

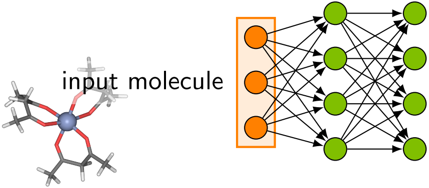
Latent distance similarity



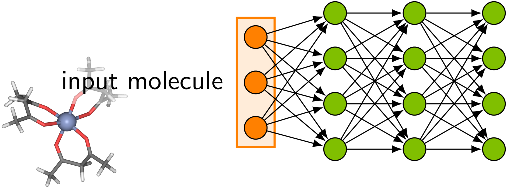
Latent distance similarity



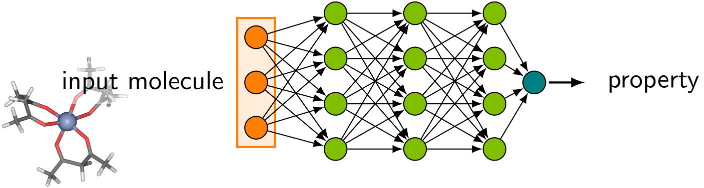
Latent distance similarity



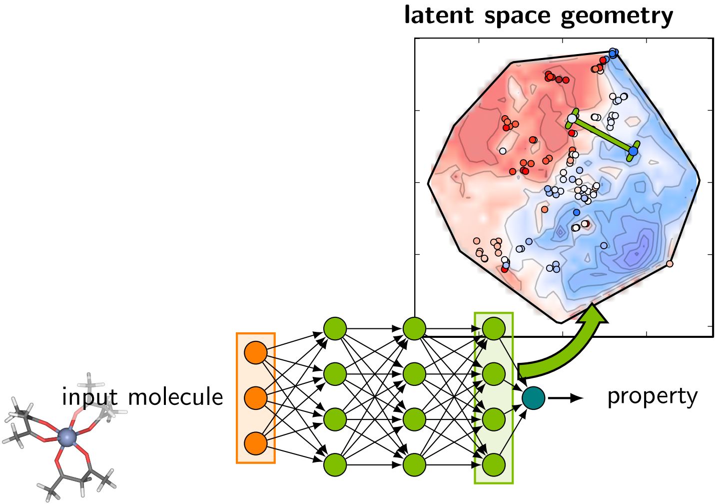
Latent distance similarity



Latent distance similarity



Latent distance similarity

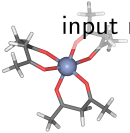


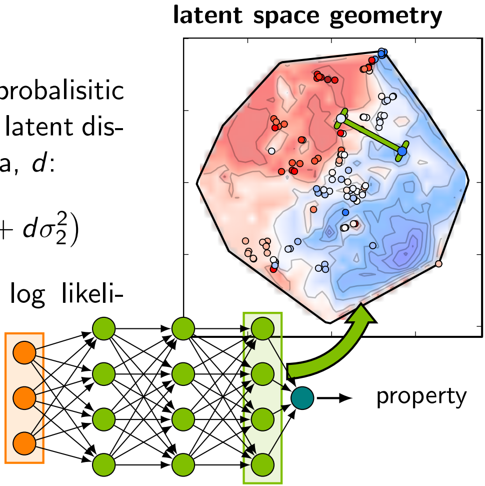
Latent distance similarity

Proposed a simple probalisitic error model based on latent distance to training data, d:

$$\varepsilon(d) \sim \mathcal{N}\left(0, \sigma_1^2 + d\sigma_2^2\right)$$

Estimate σ_1, σ_2 by log likelihood



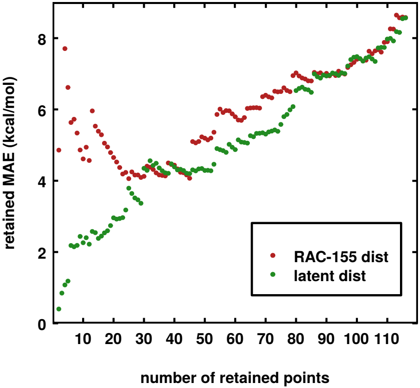


Latent distance similarity

Proposed a simple probalisitic error model based on latent distance to training data, d:

$$\varepsilon(d) \sim \mathcal{N}\left(0, \sigma_1^2 + d\sigma_2^2\right)$$

Estimate σ_1, σ_2 by log likelihood



Janet, J.P., *et al.*, *Chem. Sci.*, 10:7913–7922 , 2019.

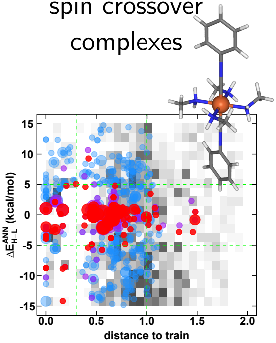
Using models for discovery

We have used evolutionary algorithms for uncertainty–aware design for TM complexes:

Using models for discovery

We have used evolutionary algorithms for uncertainty–aware design for TM complexes:

spin crossover complexes

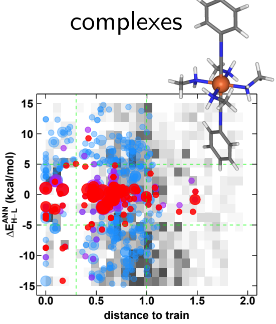


Janet, J.P., Chan, L. and Kulik, H.J., *J. Phys. Chem. Lett.*, 9(5):1064–1071, 2018.

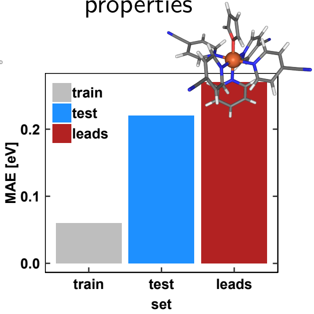
Using models for discovery

We have used evolutionary algorithms for uncertainty–aware design for TM complexes:

spin crossover complexes



frontier orbital properties

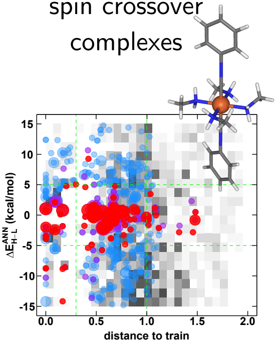


Janet, J.P., Chan, L. and Kulik, H.J., *J. Phys. Chem. Lett.*, 9(5):1064–1071, 2018.
Nandy, A. *et al.*, *Ind. Eng. Chem. Res.*, 57(42):13973–13986, 2018.

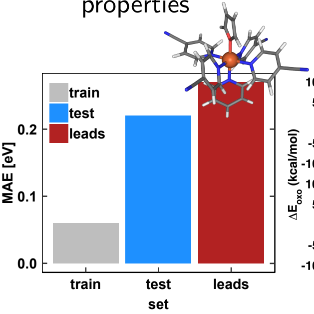
Using models for discovery

We have used evolutionary algorithms for uncertainty–aware design for TM complexes:

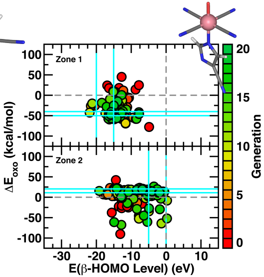
spin crossover complexes



frontier orbital properties



unusual catalytic reaction energies

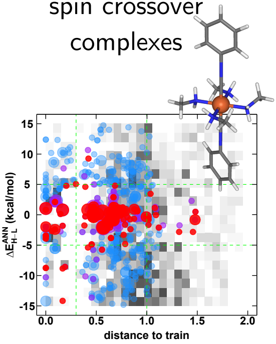


Janet, J.P., Chan, L. and Kulik, H.J., *J. Phys. Chem. Lett.*, 9(5):1064–1071, 2018.
Nandy, A. *et al.*, *Ind. Eng. Chem. Res.*, 57(42):13973–13986, 2018.
Nandy, A. *et al.*, *ACS Catal.*, 9(9):8243–8255, 2019.

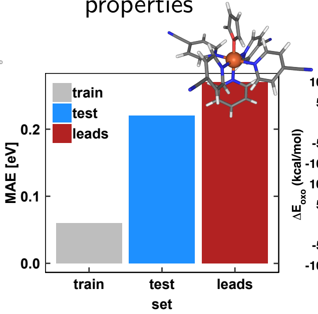
Using models for discovery

We have used evolutionary algorithms for uncertainty–aware design for TM complexes:

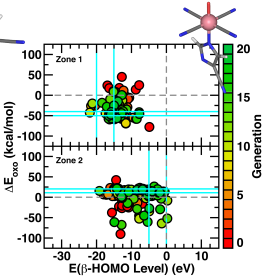
spin crossover complexes



frontier orbital properties



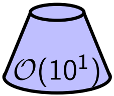
unusual catalytic reaction energies



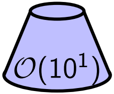
443e: Aditya Nandy: Wednesday, 9:12 (Peacock Spring)

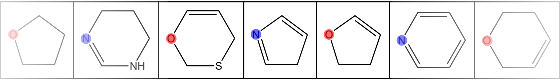
A design space for RFBs

A design space for RFBs



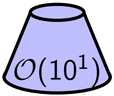
A design space for RFBs

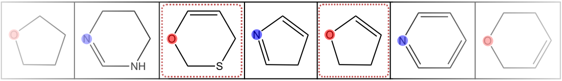




40 heterocycles

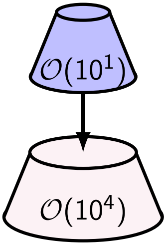
A design space for RFBs

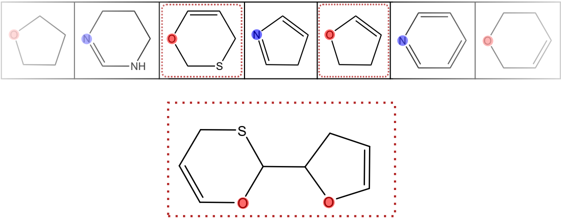




40 heterocycles

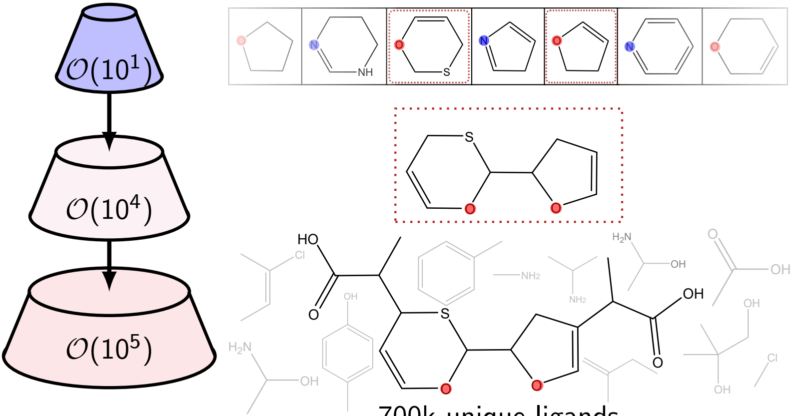
A design space for RFBs





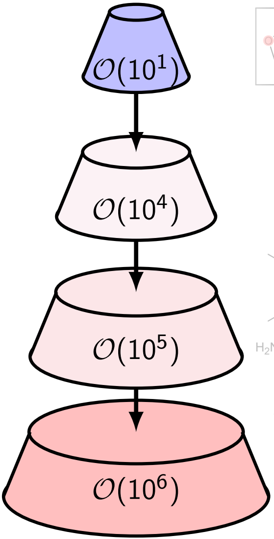
800 base ligands

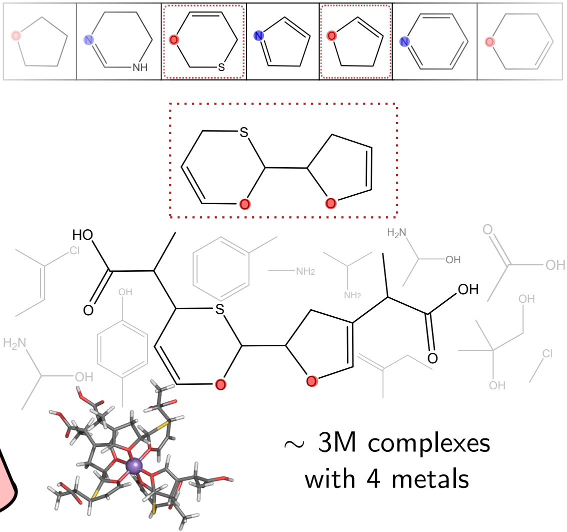
A design space for RFBs



700k unique ligands

A design space for RFBs





∼ 3M complexes
with 4 metals

Multiobjective optimization

We can predict quantites of interest for our RFBs:

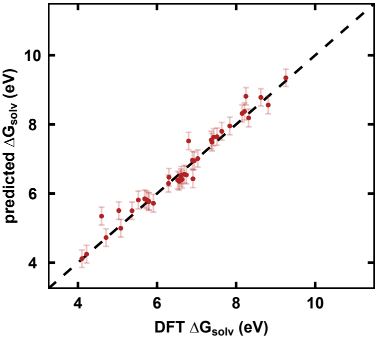
Janet, J.P., *et al.*, *in preparation*.

3×100 tanh nodes, multitask, fully connected

Multiobjective optimization

We can predict quantites of interest for our RFBs:

ΔG_{solv}



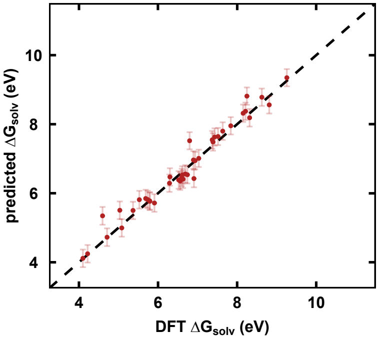
Janet, J.P., *et al.*, *in preparation*.

3×100 tanh nodes, multitask, fully connected

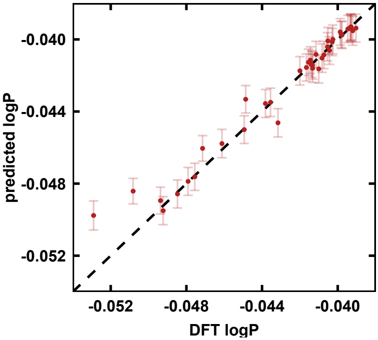
Multiobjective optimization

We can predict quantites of interest for our RFBs:

ΔG_{solv}



log P



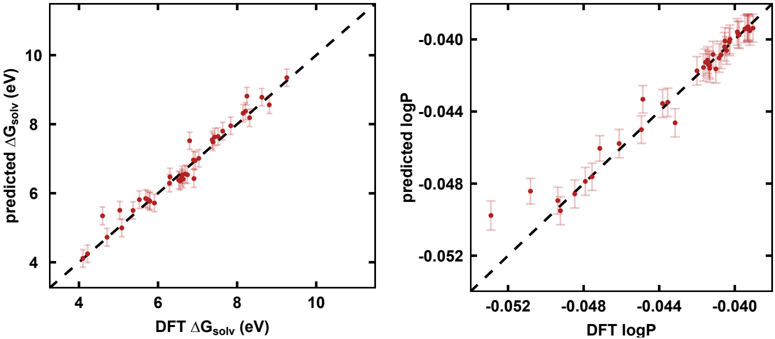
Janet, J.P., *et al.*, *in preparation*.

3 × 100 tanh nodes, multitask, fully connected

Multiobjective optimization

We can predict quantites of interest for our RFBs:

ΔG_{solv} log P



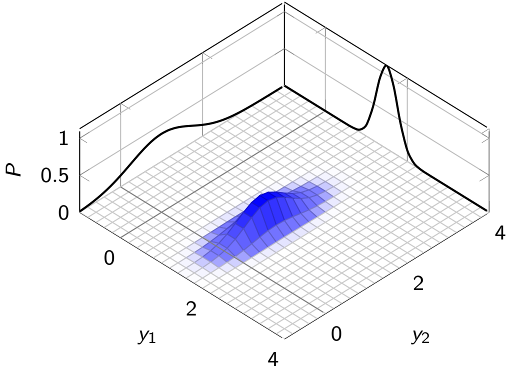
$$\begin{matrix}\Delta G_{solv} \\ \log P\end{matrix} = \begin{bmatrix}\hat{y}_1 \\ \hat{y}_2\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\hat{\mu}_1 \\ \hat{\mu}_2\end{bmatrix}, \begin{bmatrix}\hat{\sigma}_1^2 & 0 \\ 0 & \hat{\sigma}_2^2\end{bmatrix}\right)$$

Janet, J.P., *et al.*, *in preparation*.

3×100 tanh nodes, multitask, fully connected

Multiobjective optimization

We can predict quantites of interest for our RFBs:

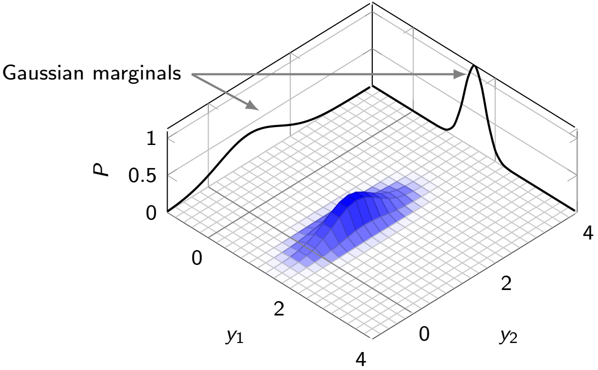


$$\begin{matrix}\Delta G_{\text{solv}} \\ \log \text{P}\end{matrix} = \begin{bmatrix}\hat{y}_1 \\ \hat{y}_2\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\hat{\mu}_1 \\ \hat{\mu}_2\end{bmatrix}, \begin{bmatrix}\hat{\sigma}_1^2 & 0 \\ 0 & \hat{\sigma}_2^2\end{bmatrix}\right)$$

Janet, J.P., *et al.*, *in preparation*.

Multiobjective optimization

We can predict quantites of interest for our RFBs:

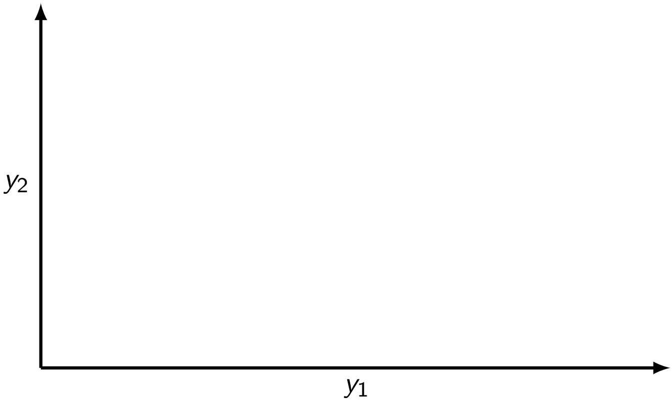


$$\begin{matrix}\Delta G_{\text{solv}} \\ \log \mathrm{P}\end{matrix} = \begin{bmatrix}\hat{y}_1 \\ \hat{y}_2\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\hat{\mu}_1 \\ \hat{\mu}_2\end{bmatrix}, \begin{bmatrix}\hat{\sigma}_1^2 & 0 \\ 0 & \hat{\sigma}_2^2\end{bmatrix}\right)$$

Janet, J.P., *et al.*, *in preparation*.

2D EGO Illustration

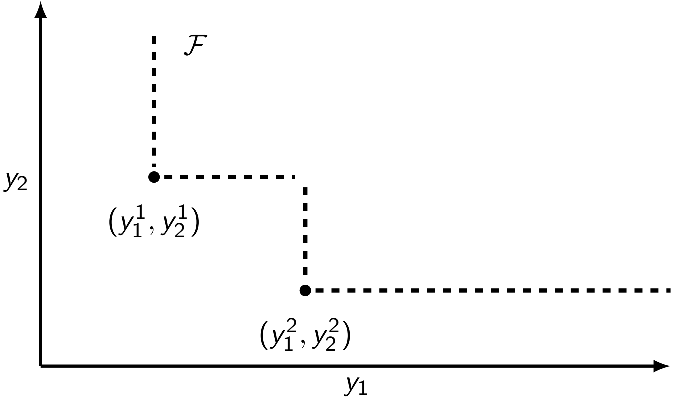
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

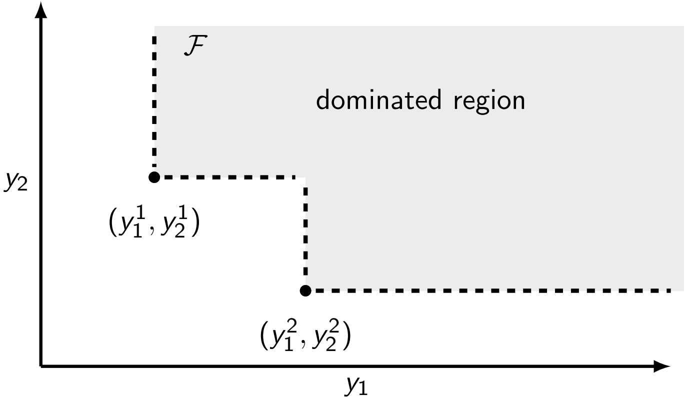
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

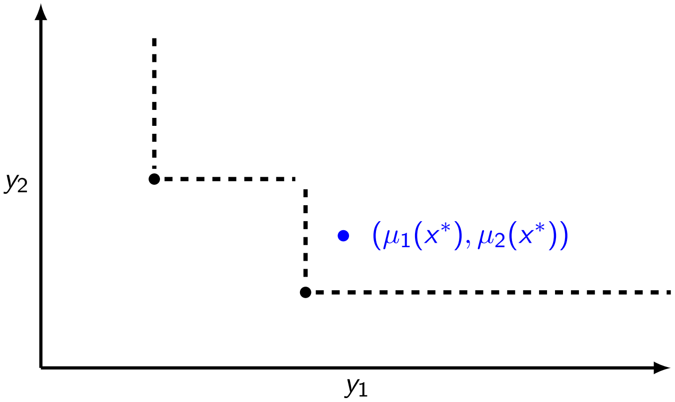
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

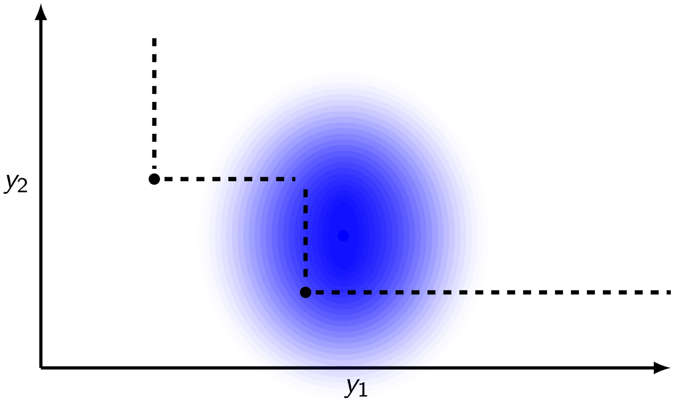
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

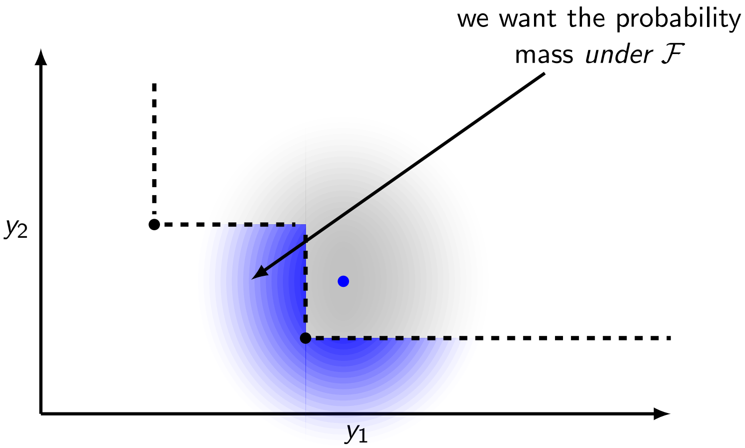
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

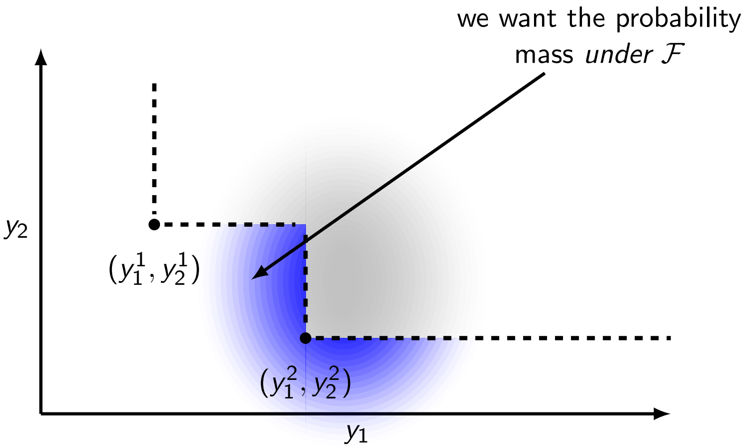
We will use a multiobjective expected improvement framework:



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

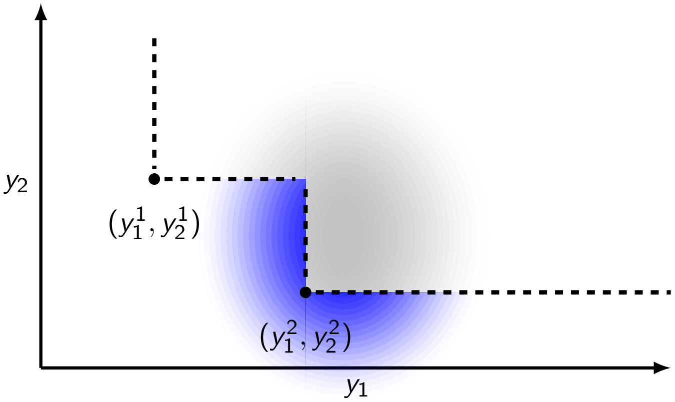


Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

$$P(I) =$$

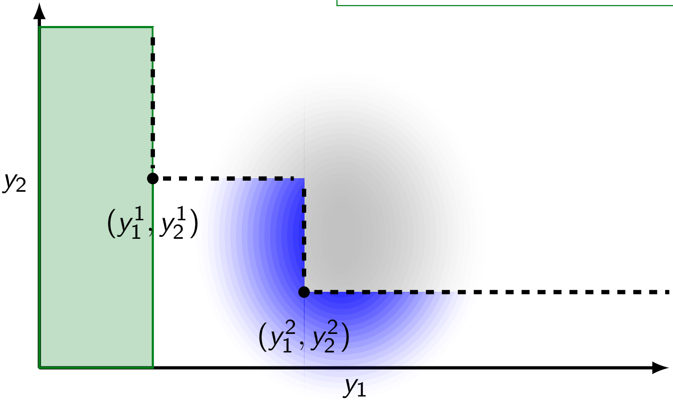


Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

$$P(I) = \boxed{\int_{-\infty}^{y_1^1}\int_{-\infty}^{\infty} p(\hat{y}_1(x^*), \hat{y}_2(x^*))\, dy_1\, dy_2}$$



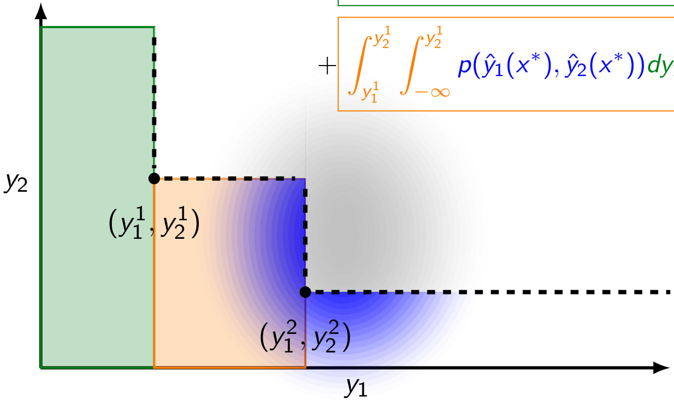
Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

$$P(I) = \boxed{\int_{-\infty}^{y_1^1}\int_{-\infty}^{\infty} p(\hat{y}_1(x^*), \hat{y}_2(x^*))dy_1 dy_2}$$

$$+ \boxed{\int_{y_1^1}^{y_2^1}\int_{-\infty}^{y_2^1} p(\hat{y}_1(x^*), \hat{y}_2(x^*))dy_1 dy_2}$$

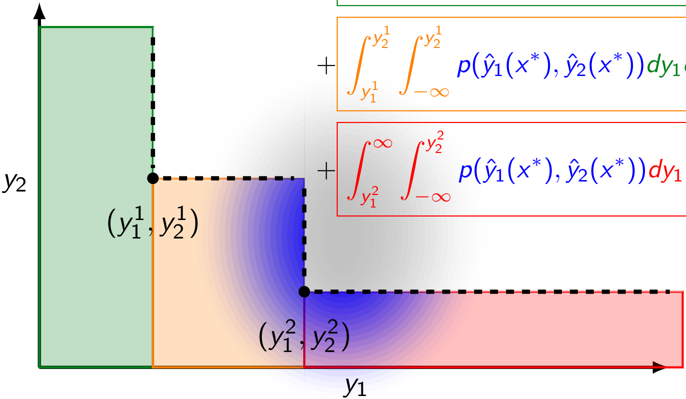


Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

$$P(I) = \boxed{\int_{-\infty}^{y_1^1}\int_{-\infty}^{\infty} p(\hat{y}_1(x^*), \hat{y}_2(x^*))dy_1 dy_2}$$
$$+ \boxed{\int_{y_1^1}^{y_2^1}\int_{-\infty}^{y_2^1} p(\hat{y}_1(x^*), \hat{y}_2(x^*))dy_1 dy_2}$$
$$+ \boxed{\int_{y_1^2}^{\infty}\int_{-\infty}^{y_2^2} p(\hat{y}_1(x^*), \hat{y}_2(x^*))dy_1 dy_2}$$



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

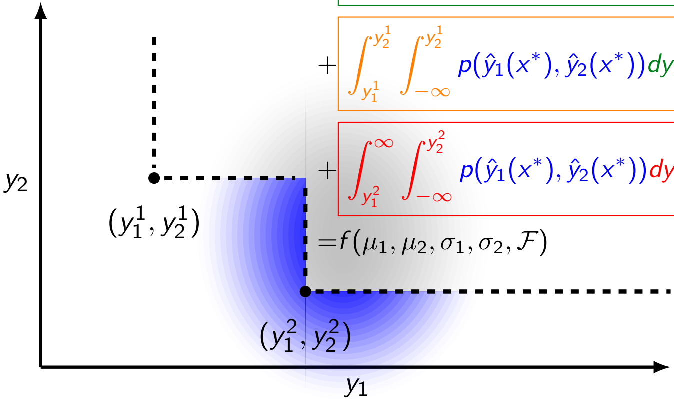
We will use a multiobjective expected improvement framework:

$$P(I) = \boxed{\int_{-\infty}^{y_1^1} \int_{-\infty}^{\infty} p(\hat{y}_1(x^*), \hat{y}_2(x^*)) dy_1 dy_2}$$

$$+ \boxed{\int_{y_1^1}^{y_2^1} \int_{-\infty}^{y_2^1} p(\hat{y}_1(x^*), \hat{y}_2(x^*)) dy_1 dy_2}$$

$$+ \boxed{\int_{y_1^2}^{\infty} \int_{-\infty}^{y_2^2} p(\hat{y}_1(x^*), \hat{y}_2(x^*)) dy_1 dy_2}$$

$$= f(\mu_1, \mu_2, \sigma_1, \sigma_2, \mathcal{F})$$

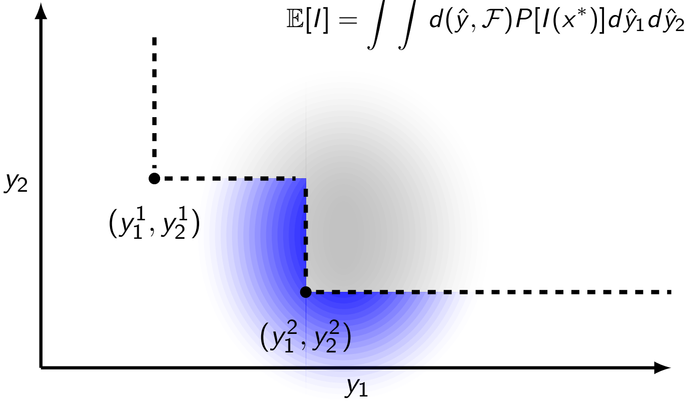


Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

2D EGO Illustration

We will use a multiobjective expected improvement framework:

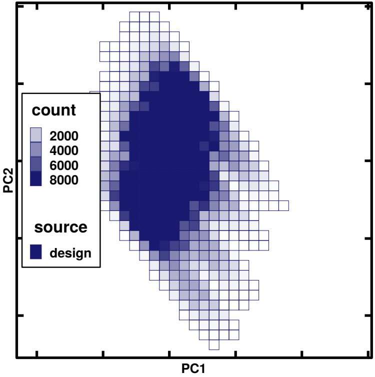
$$\mathbb{E}[I] = \int\int d(\hat{y}, \mathcal{F}) P[I(x^*)] d\hat{y}_1 d\hat{y}_2$$



Keane, A. J., AIAA Journal, 44(4):879–891, 2006.

Design space and clustering

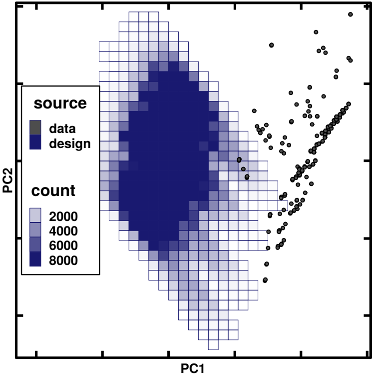
We jump start with diversity-oriented cluster:



Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

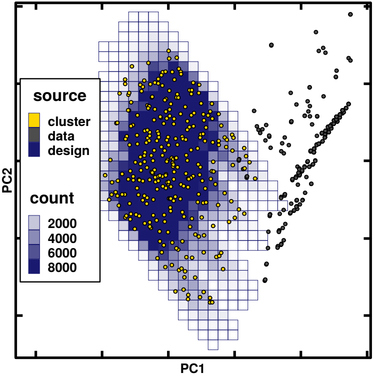
We jump start with diversity-oriented cluster:



Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

We jump start with diversity-oriented cluster:

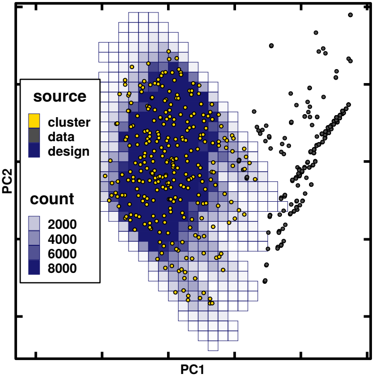


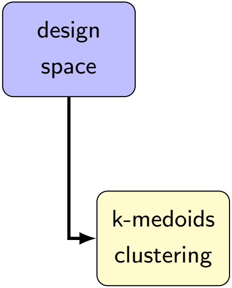
design space

Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

We jump start with diversity-oriented cluster:

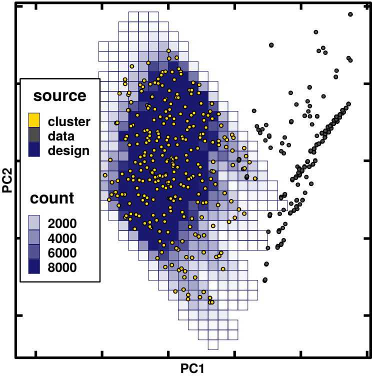


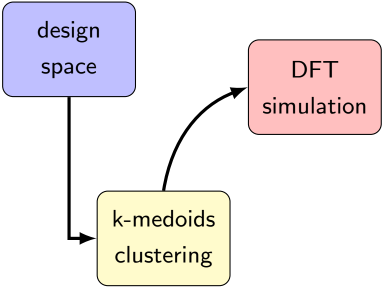


Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

We jump start with diversity-oriented cluster:

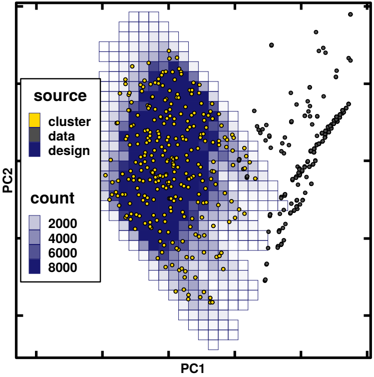


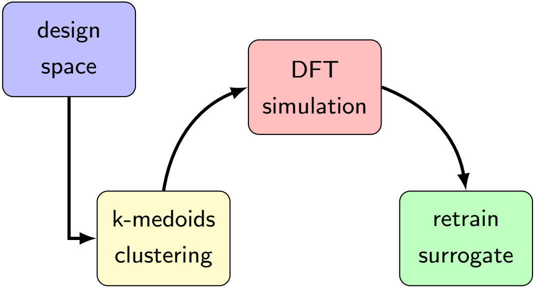


Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

We jump start with diversity-oriented cluster:

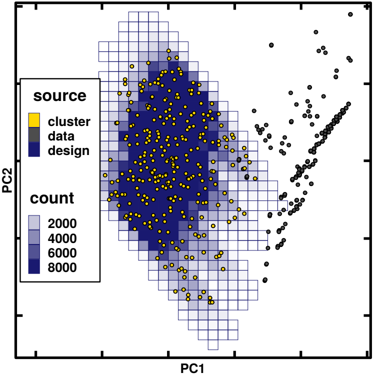


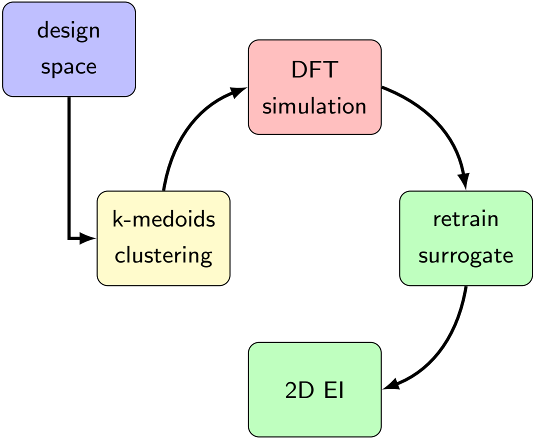


Janet, J.P., *et al.*, *in preparation*.

Design space and clustering

We jump start with diversity-oriented cluster:



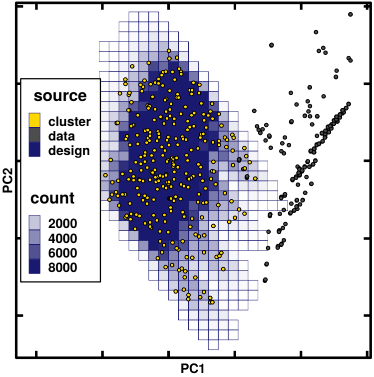


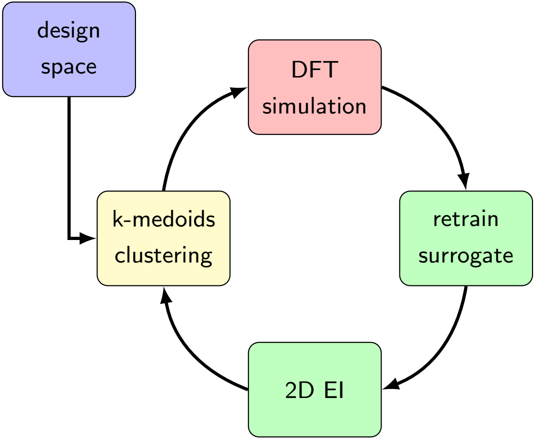
Janet, J.P., *et al.*, *in preparation*.

multitask: 2 × 100 tanh nodes, fully connected

Design space and clustering

We jump start with diversity-oriented cluster:



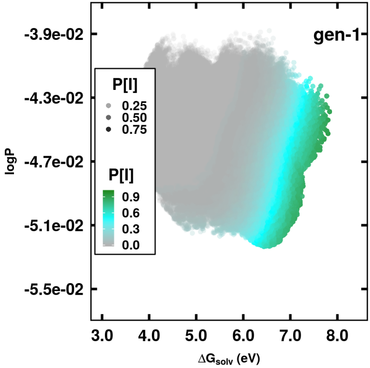


Janet, J.P., *et al.*, *in preparation*.

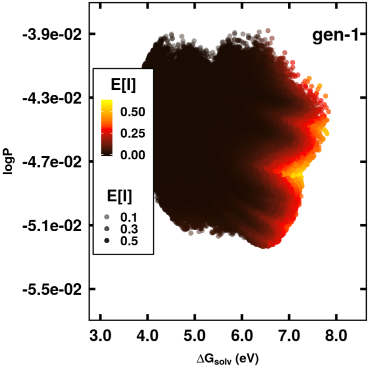
multitask: 2 × 100 tanh nodes, fully connected

EGO results

probability of improvement

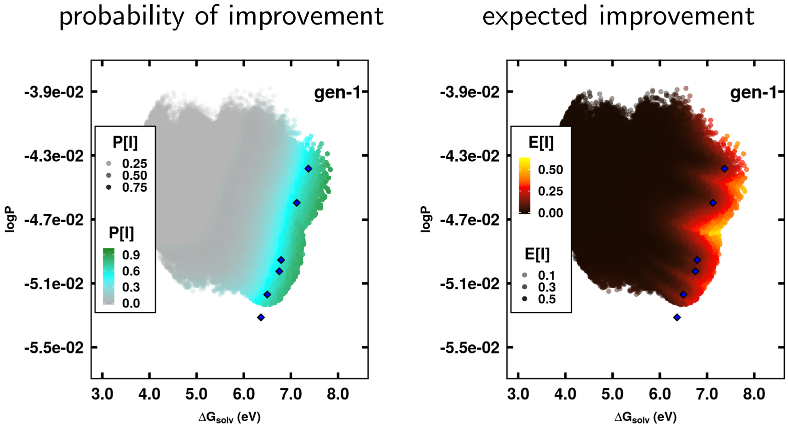


expected improvement



Janet, J.P., *et al.*, *in preparation*.

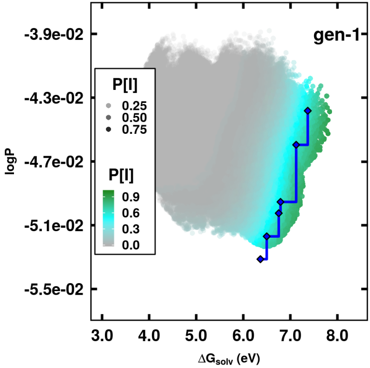
EGO results



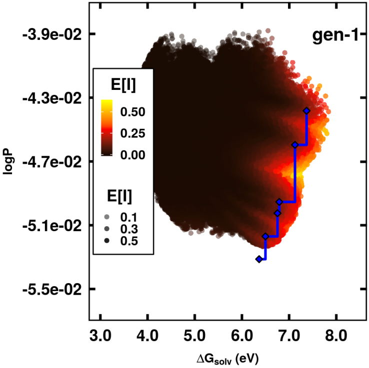
Janet, J.P., *et al.*, *in preparation*.

EGO results

probability of improvement

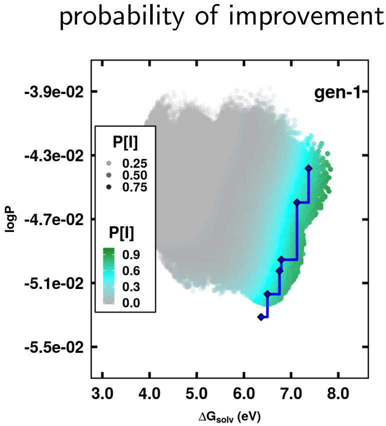


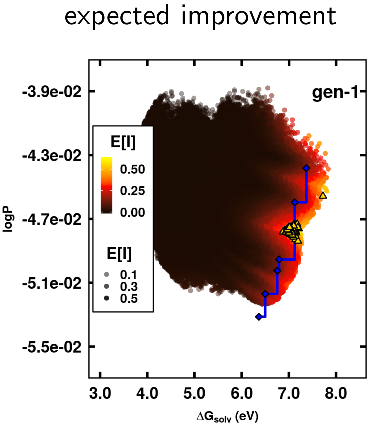
expected improvement



Janet, J.P., *et al.*, *in preparation*.

EGO results

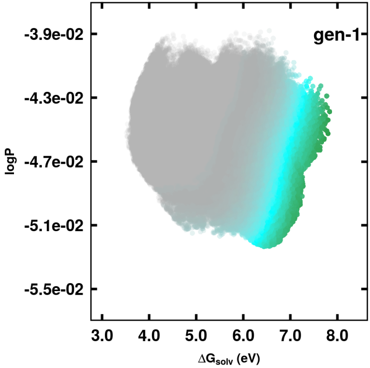




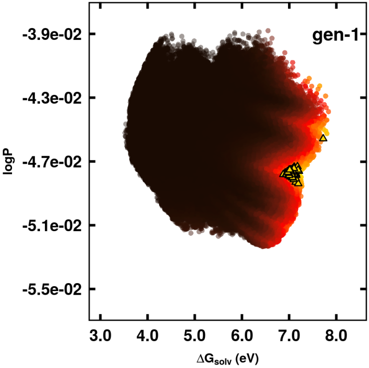
Janet, J.P., *et al.*, *in preparation*.

EGO results

probability of improvement

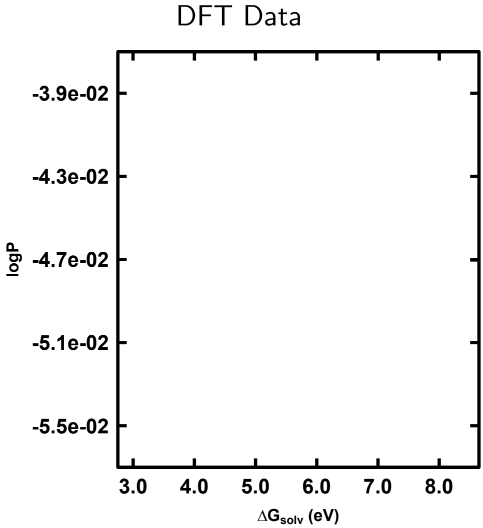


expected improvement



Janet, J.P., *et al.*, *in preparation*.

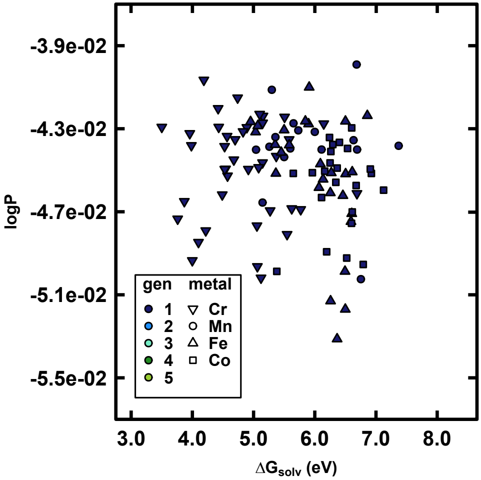
DFT results



Janet, J.P., *et al.*, *in preparation*.

DFT results

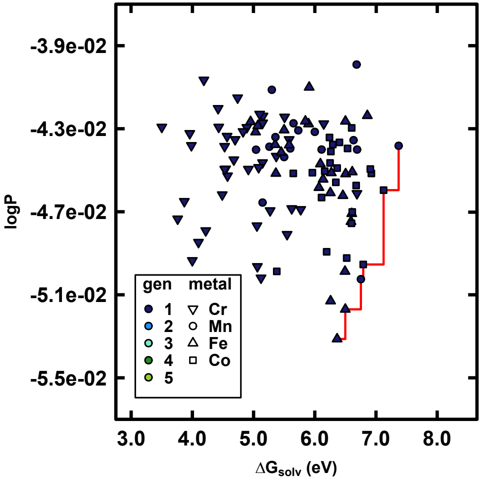
k-medoids points (generation 1)



Janet, J.P., *et al.*, *in preparation*.

DFT results

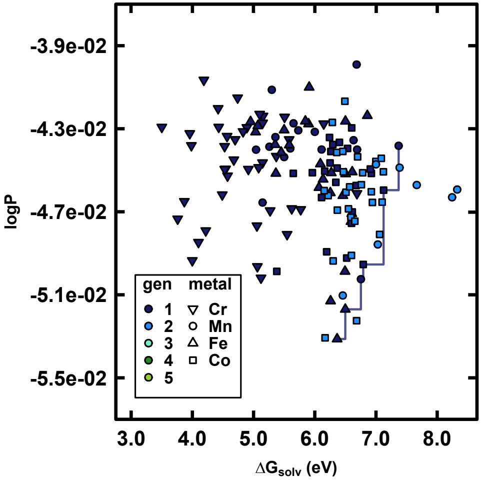
pareto front (generation 1)



Janet, J.P., *et al.*, *in preparation*.

DFT results

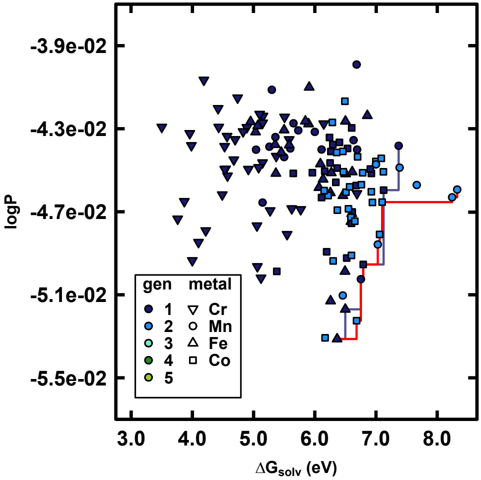
EI points (geneneration 2)



Janet, J.P., *et al.*, *in preparation*.

DFT results

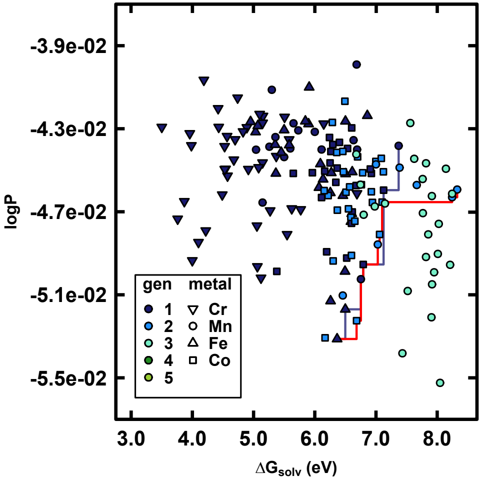
pareto front (generation 2)



Janet, J.P., *et al.*, *in preparation.*

DFT results

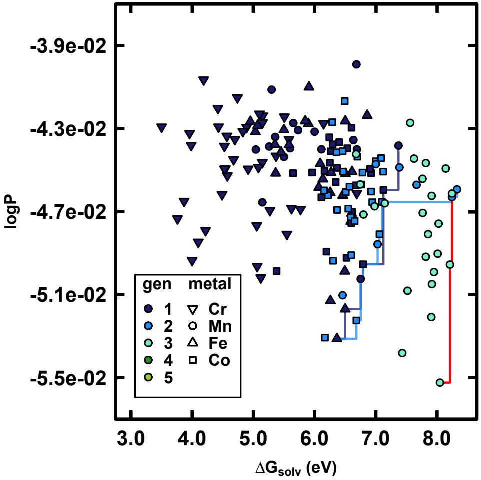
EI points (generation 3)



Janet, J.P., *et al.*, *in preparation*.

DFT results

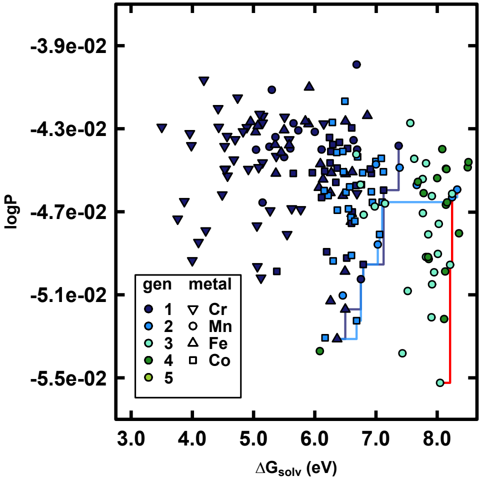
pareto front (generation 3)



Janet, J.P., *et al.*, *in preparation*.

DFT results

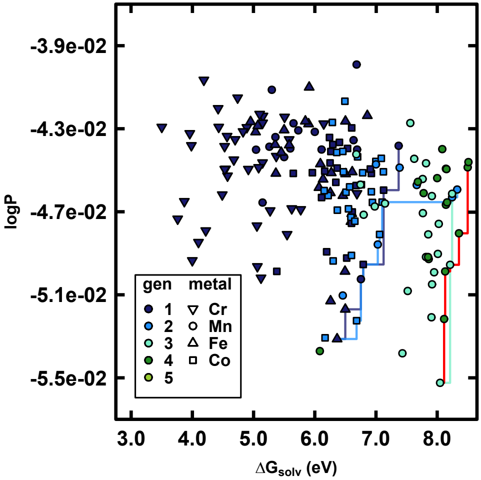
EI points (generation 4)



Janet, J.P., *et al.*, *in preparation*.

DFT results

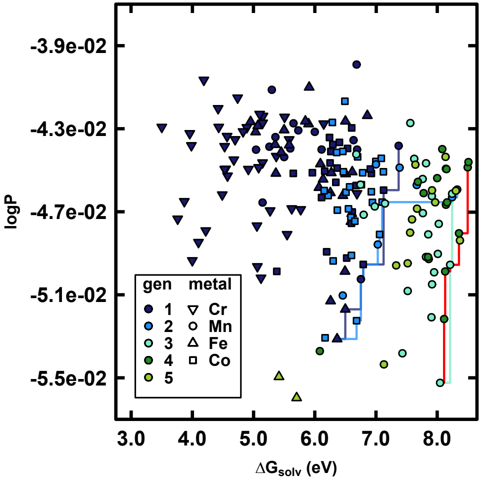
pareto front (generation 4)



Janet, J.P., *et al.*, *in preparation*.

DFT results

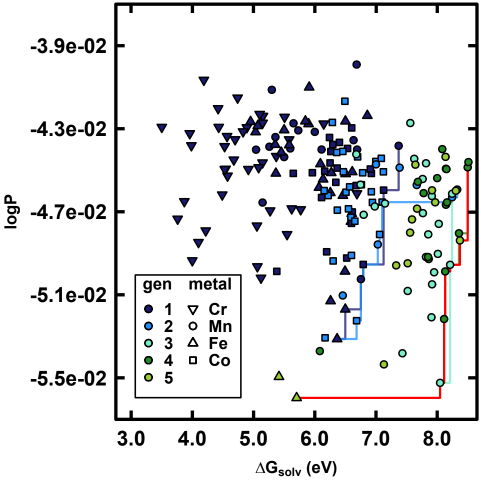
EI points (generation 5)



Janet, J.P., *et al.*, *in preparation*.

DFT results

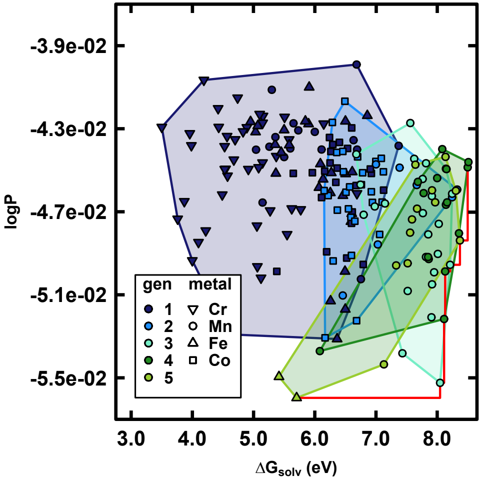
pareto front (generation 5)



Janet, J.P., *et al.*, *in preparation.*

DFT results

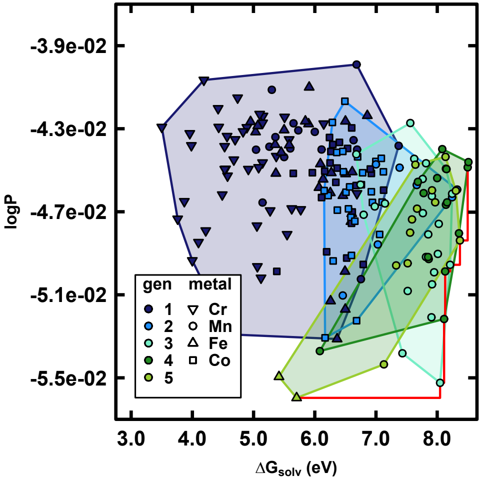
convex hulls



Janet, J.P., *et al.*, *in preparation*.

DFT results

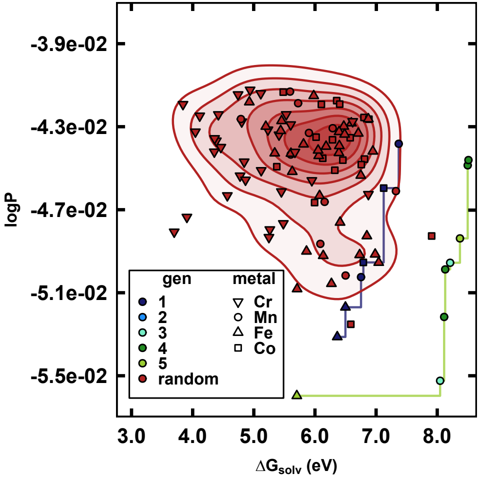
convex hulls



Janet, J.P., *et al.*, *in preparation*.

DFT results

comparison to random sampling

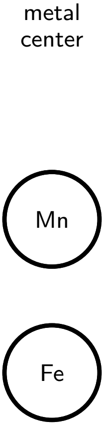


Janet, J.P., *et al.*, *in preparation*.

Final lead analysis

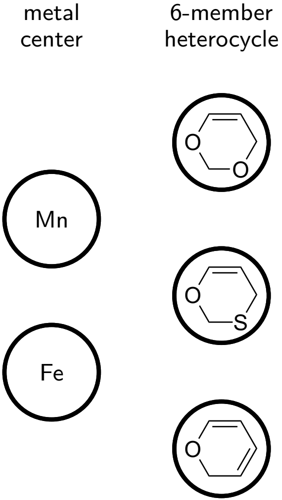
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



Janet, J.P., *et al.*, *in preparation*.

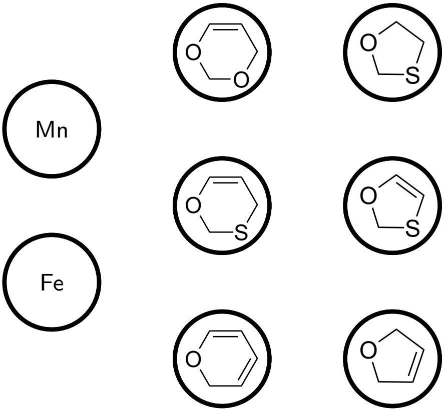
Final lead analysis



Janet, J.P., *et al.*, *in preparation*.

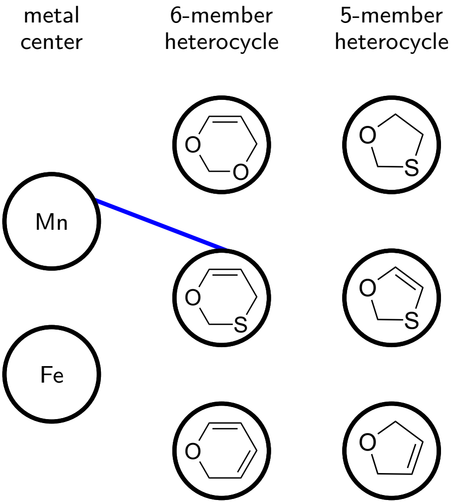
Final lead analysis

metal center	6-member heterocycle	5-member heterocycle



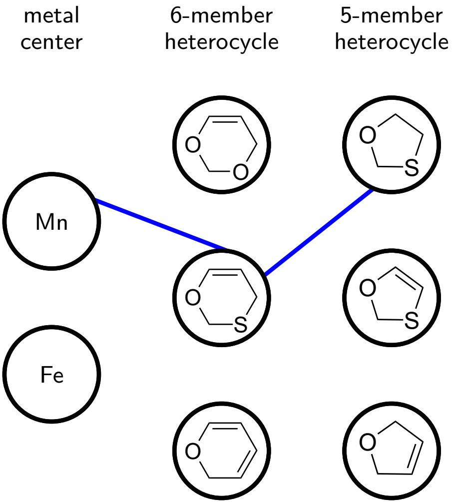
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



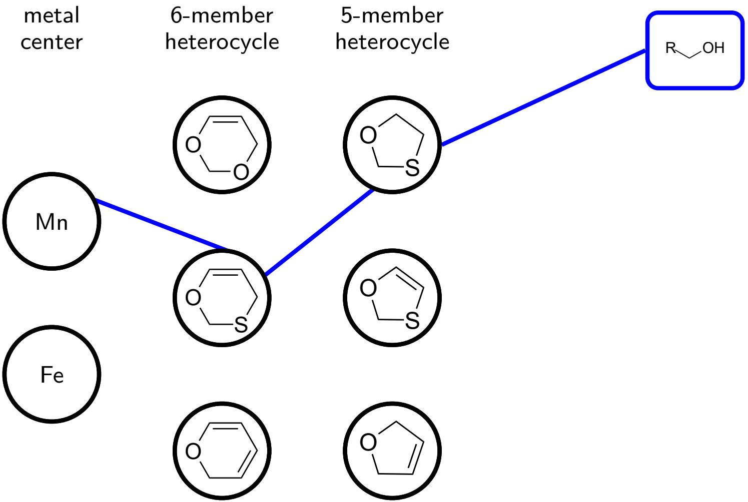
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



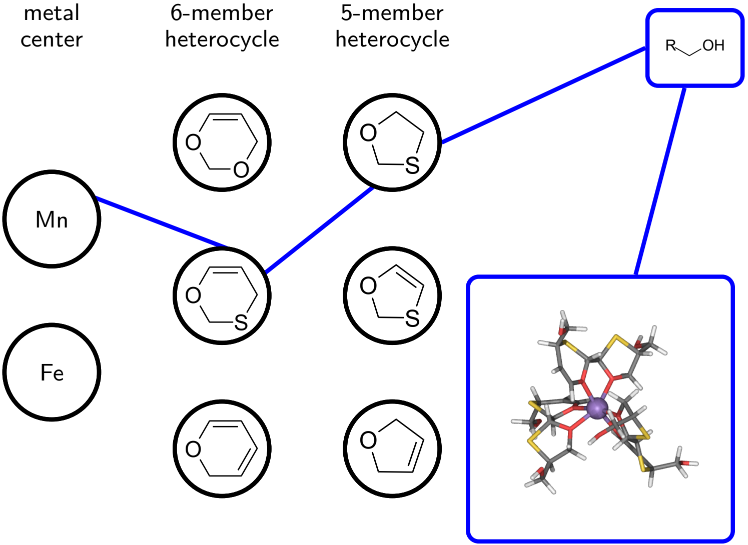
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



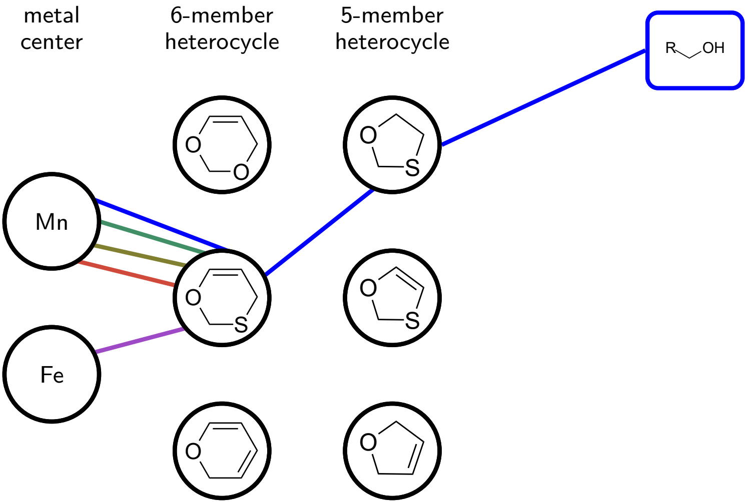
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



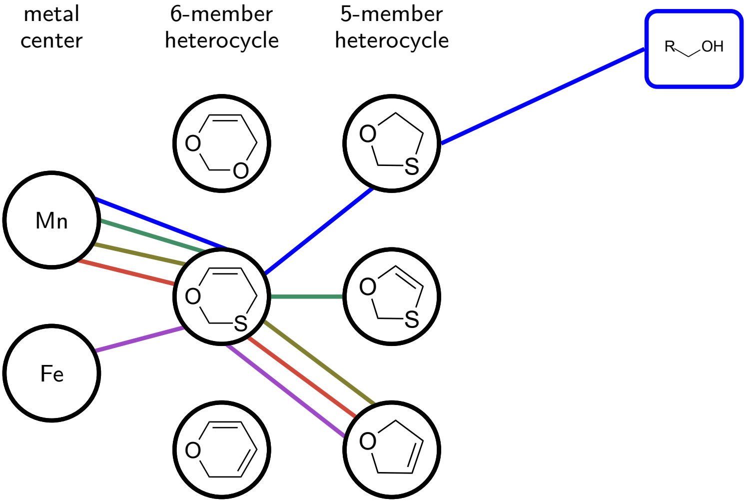
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



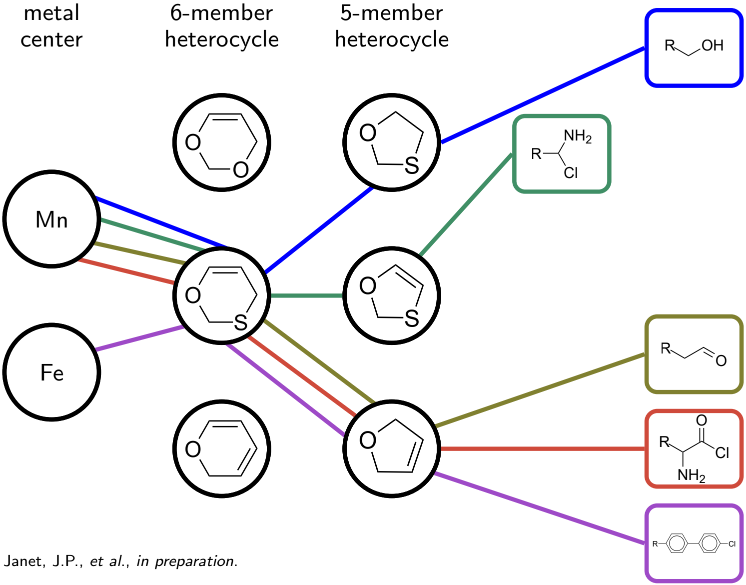
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



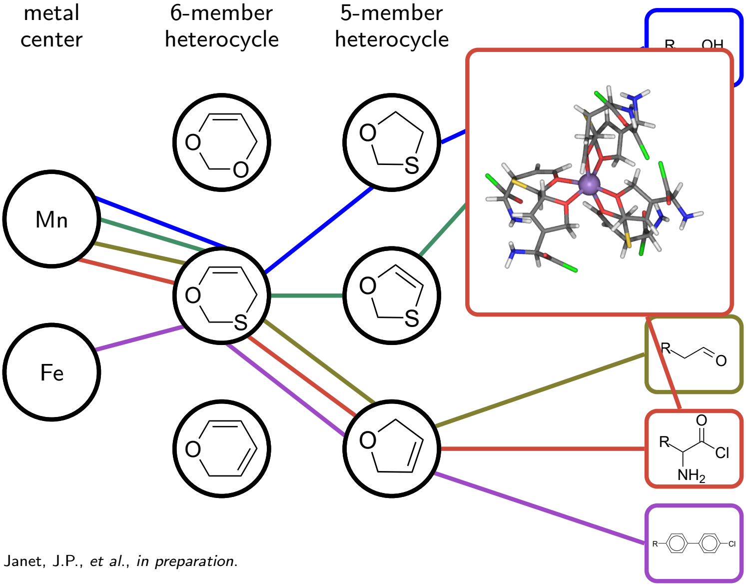
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



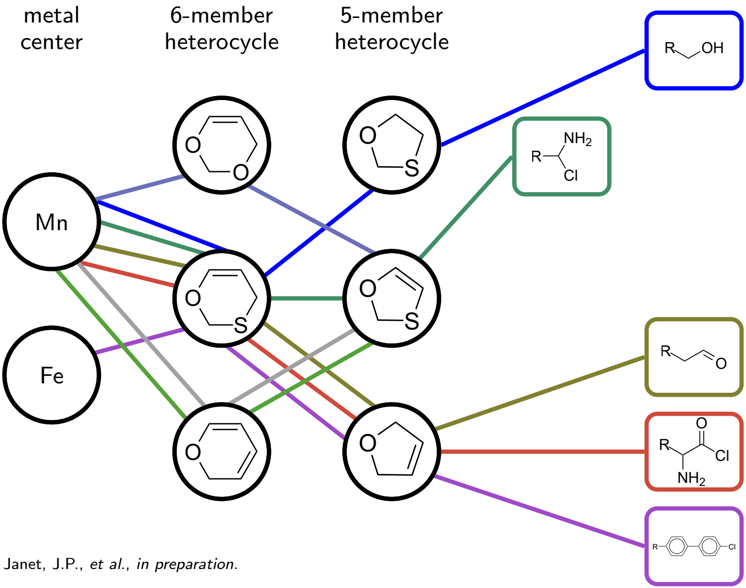
Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



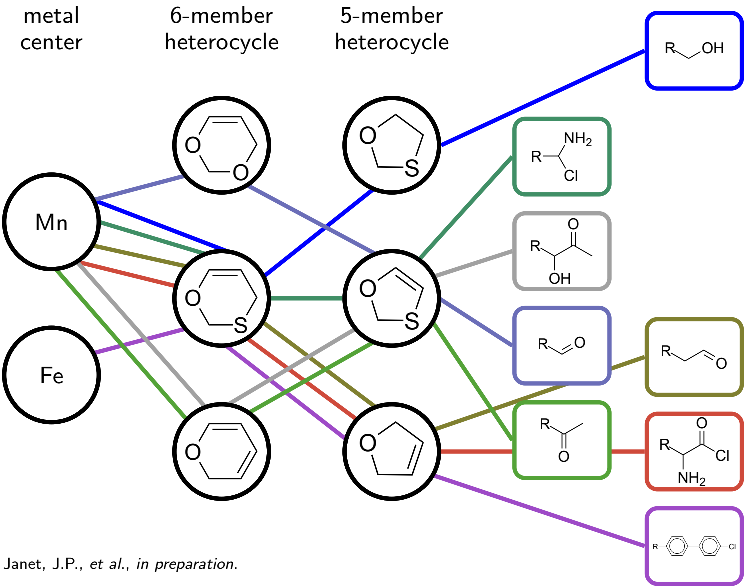
Janet, J.P., *et al.*, *in preparation.*

Final lead analysis



Janet, J.P., *et al.*, *in preparation*.

Final lead analysis



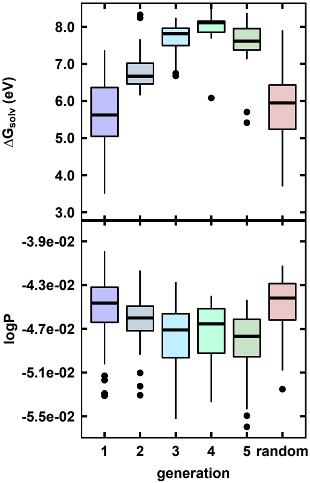
Janet, J.P., *et al.*, *in preparation*.

Case study conclusions

Janet, J.P., *et al.*, *in preparation*.

Case study conclusions

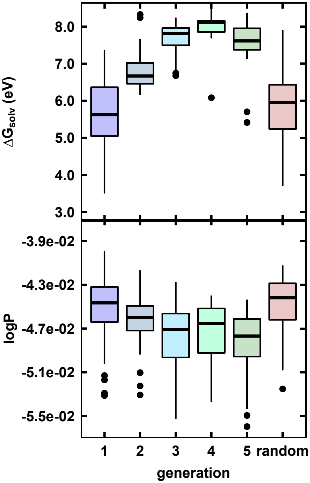
- EI framework provides high resolution in the region of interest (c.f. maximum uncertainty), converges quickly



Janet, J.P., *et al.*, *in preparation*.

Case study conclusions

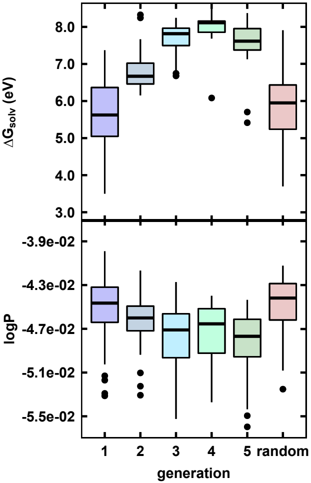
- EI framework provides high resolution in the region of interest (c.f. maximum uncertainty), converges quickly
- We are able to identify fruitful regions from large chemical spaces based on few DFT evaluations



Janet, J.P., *et al.*, *in preparation*.

Case study conclusions

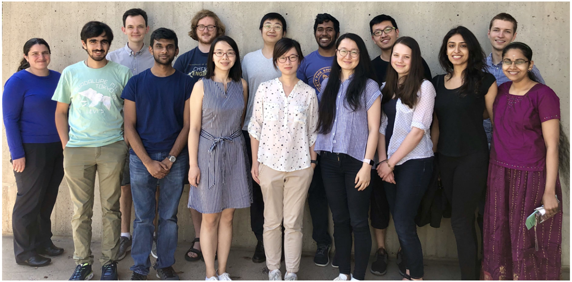
- EI framework provides high resolution in the region of interest (c.f. maximum uncertainty), converges quickly
- We are able to identify fruitful regions from large chemical spaces based on few DFT evaluations
- Multiobjective DFT optimization guided by data-driven method efficiency generates lead complexes



Janet, J.P., *et al.*, *in preparation.*

Acknowledgments

Thanks to the Kulik group and funding partners:



DARPA
MITei
Department of Energy, United States of America
eni
Office of Naval Research Science & Technology
NSF
BURROUGHS WELLCOME FUND