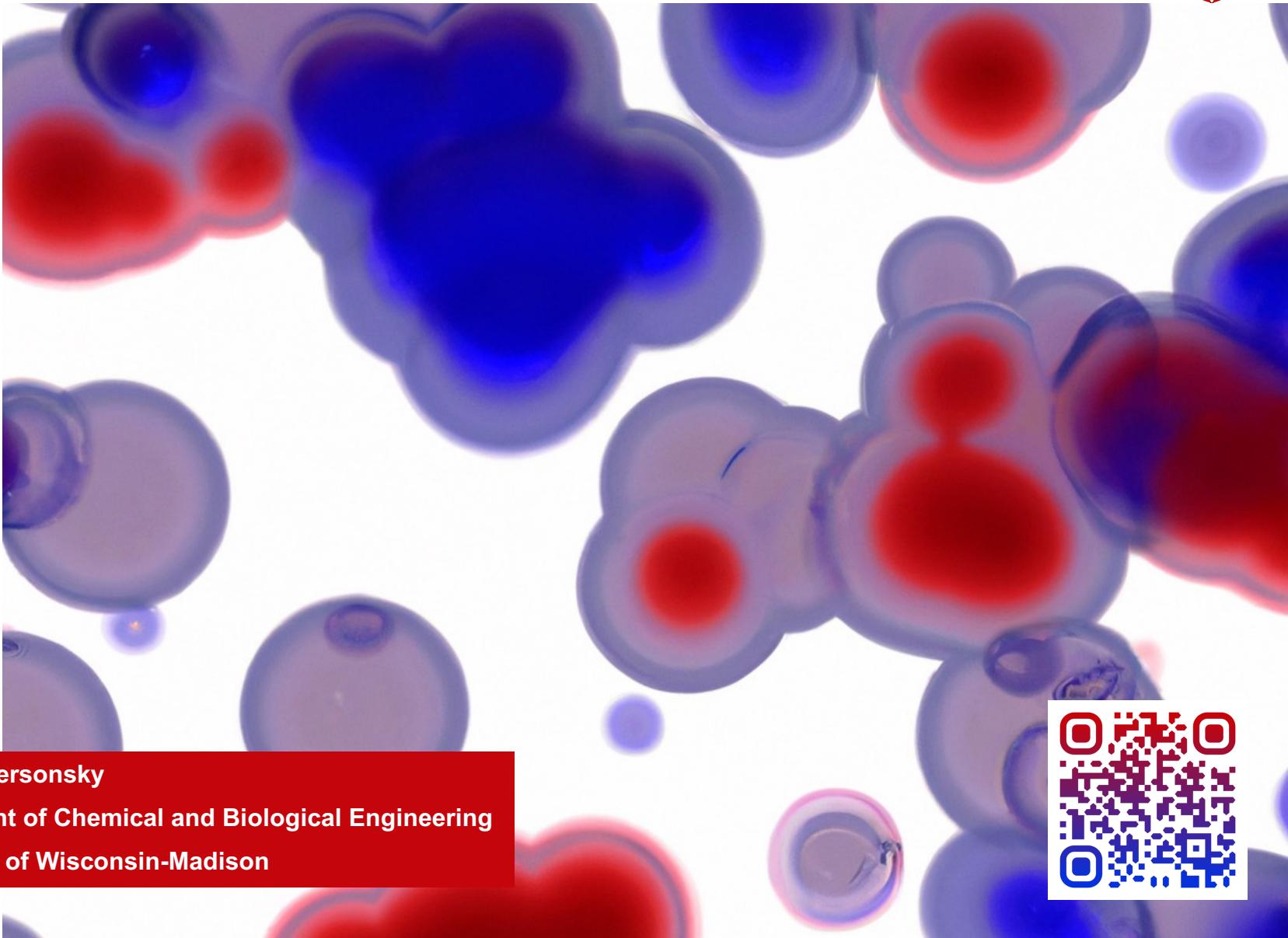
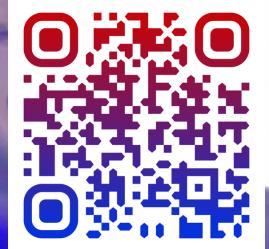


# Disentangling the impact of packing in colloidal and molecular self-assembly



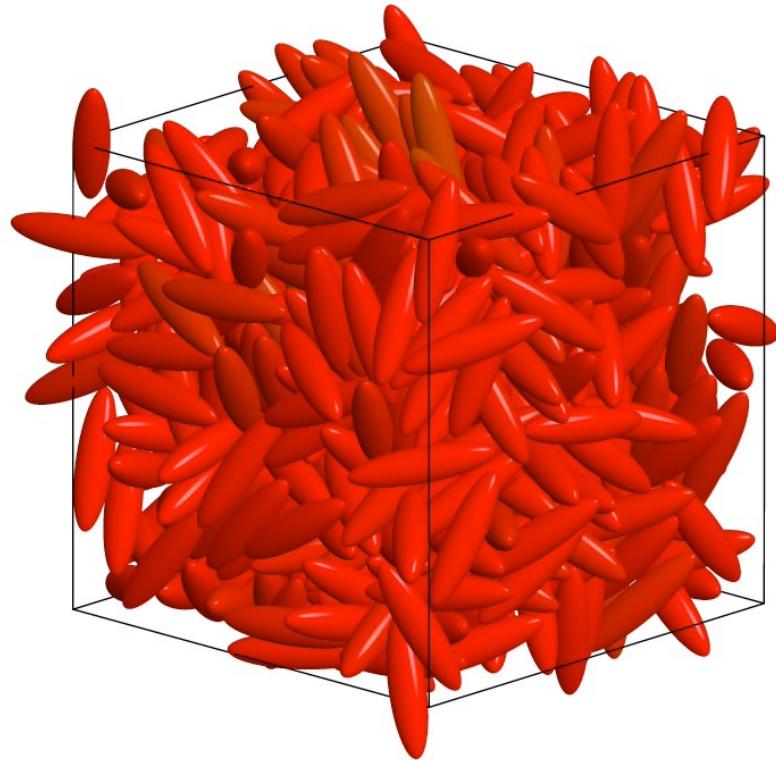
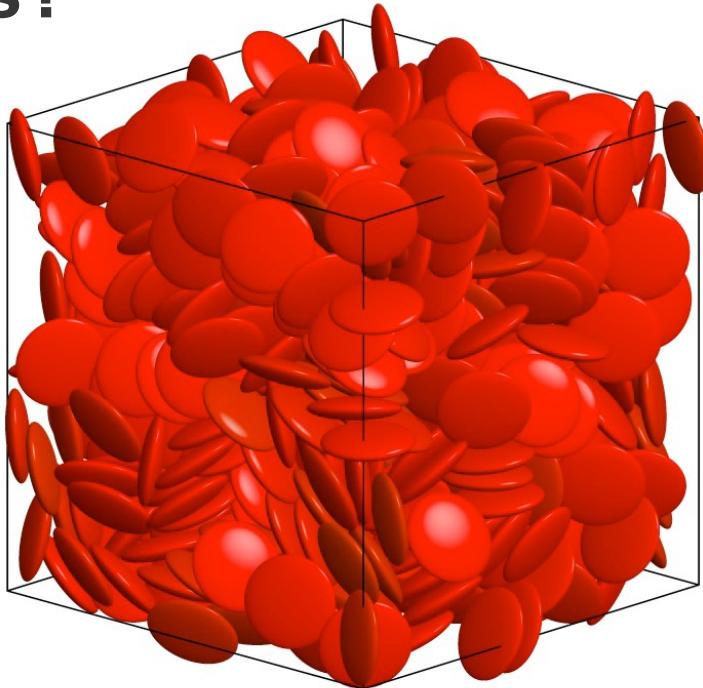
Rose K. Cersonsky

Department of Chemical and Biological Engineering  
University of Wisconsin-Madison





# What is the role of “packing” in the self- organization of particles?

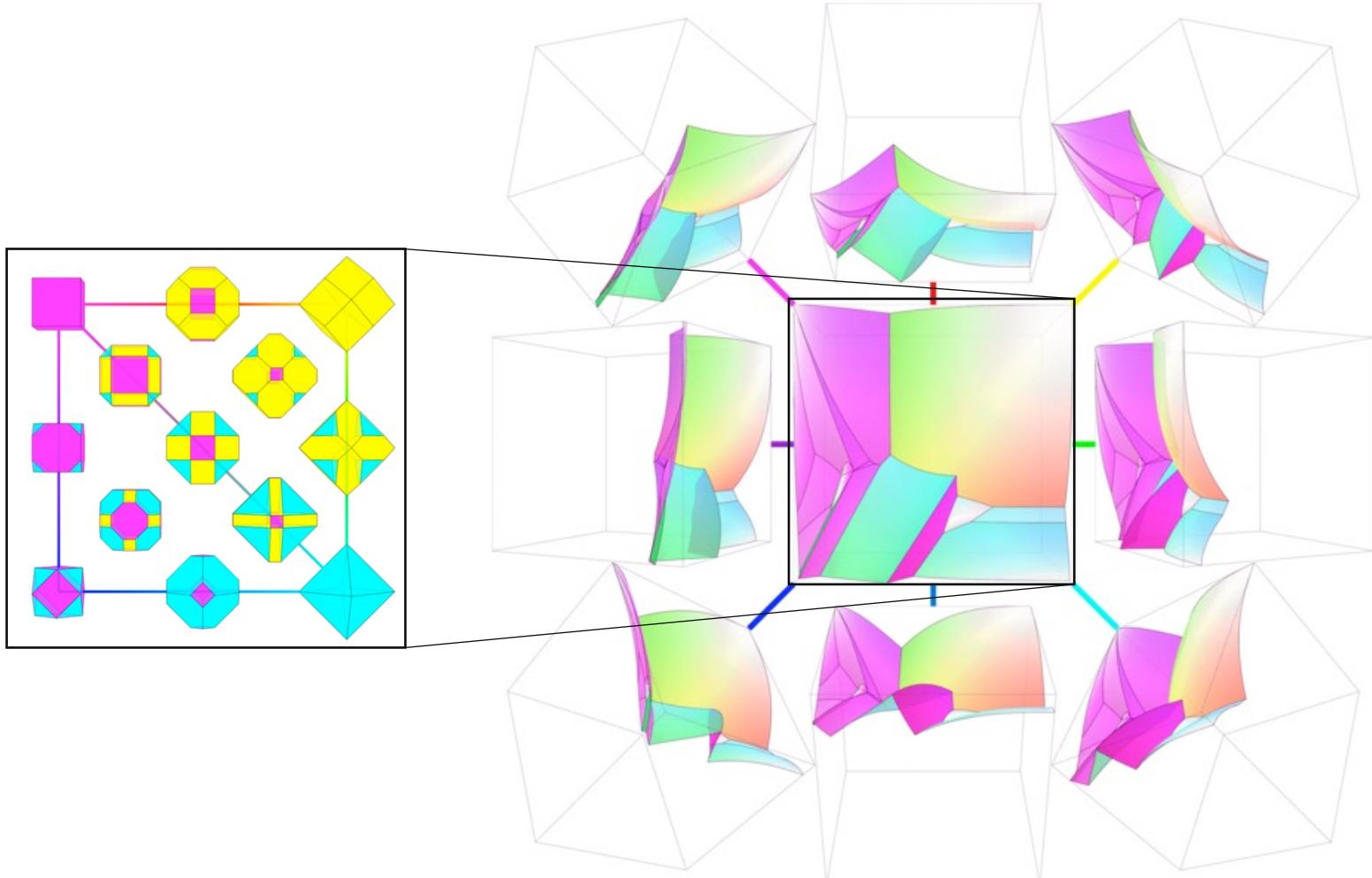


Here, hard particles are being simulated in a box keeping the number of particles ( $N$ ) and volume of the box and particles ( $V$ ) are kept constant. Red indicates a fluid-like arrangement, green a locally-ordered one.





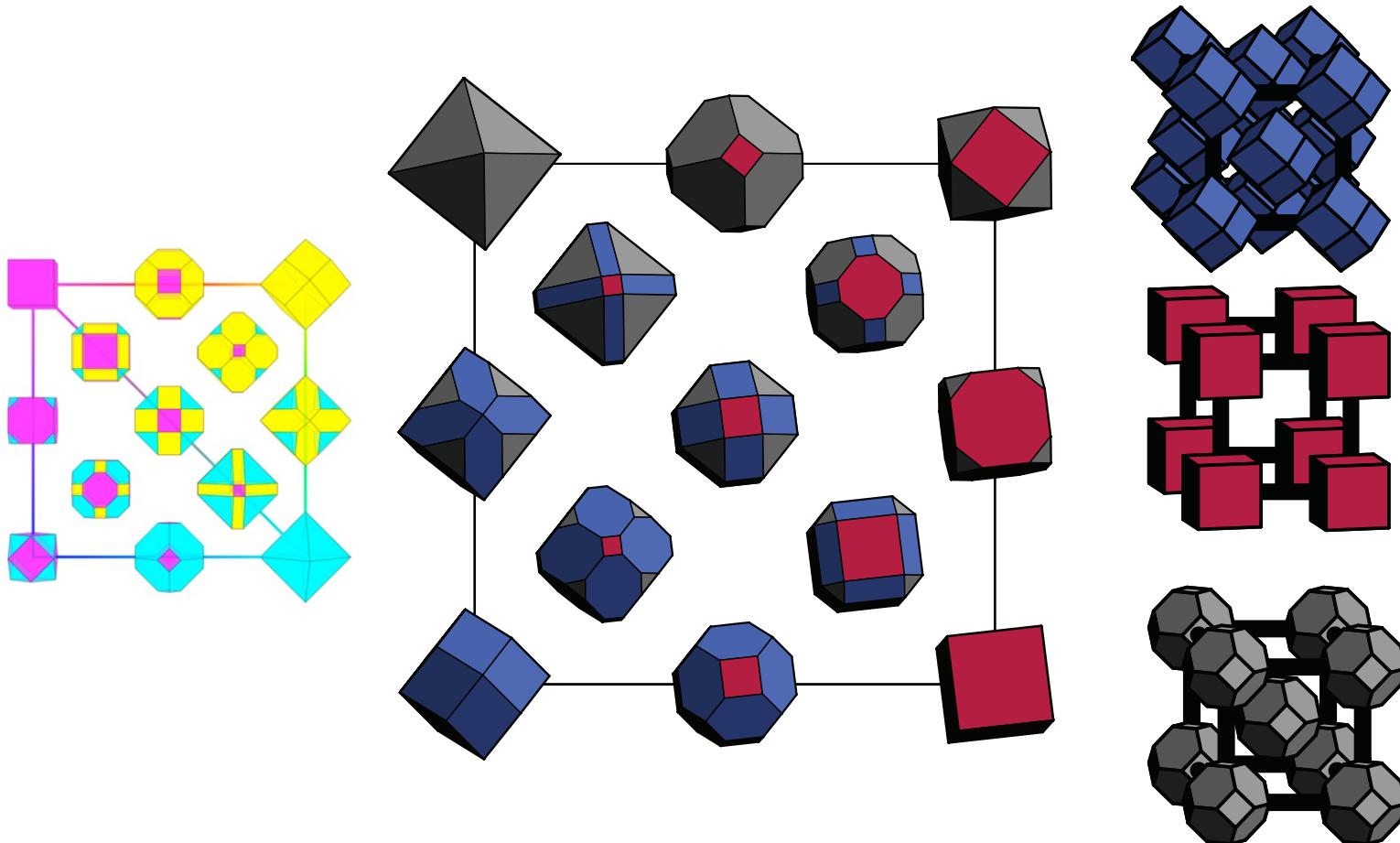
Here I will define packing as the arrangement of particles that minimizes free-volume; the best arrangement of particles under infinite pressure.





# What particles are most likely to be influenced by their “packings” in their self-organization processes?

Hard particles that both assemble and pack the same crystallographic structures





Given this construction, we can develop a thermodynamic ansatz for identifying the onset of packing behavior in these statistical ensembles.

$$F = \text{Free Energy} - \text{Temperature} \cdot \text{Entropy} - N \sum_i \mu_i \alpha_i$$

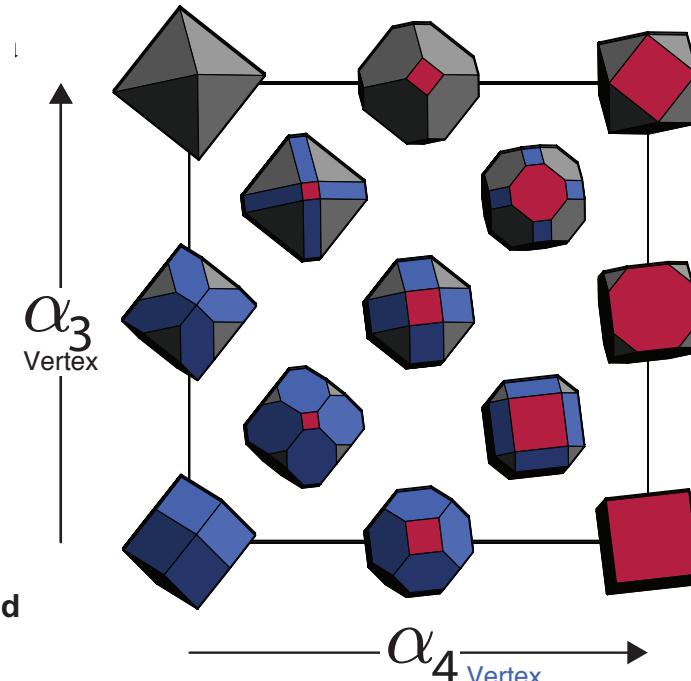
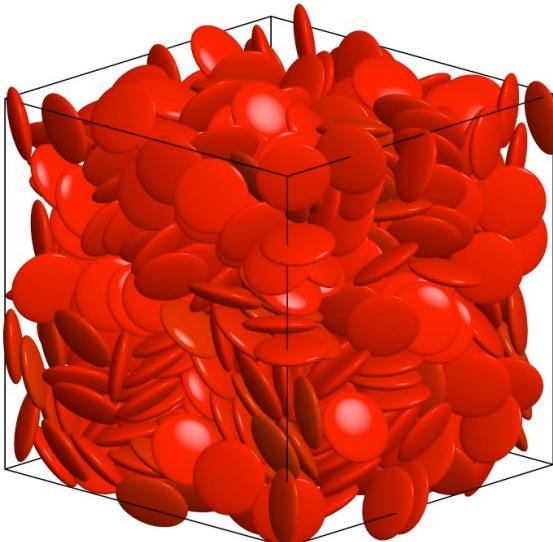
Internal Energy

Entropy

Temperature

“alchemical variable”

“alchemical potential”





**Given this construction, we can develop a thermodynamic ansatz for identifying the onset of packing behavior in these statistical ensembles.**

$$F = \text{Free Energy} - T S - N \sum_i \mu_i \alpha_i$$

Entropy      Temperature      “alchemical variable”  
 “alchemical potential”

Here I have changed our Maxwell relation to be in terms of volume density  $\phi$ , where  $\phi = N \frac{V_{\text{particle}}}{V_{\text{cell}}}$

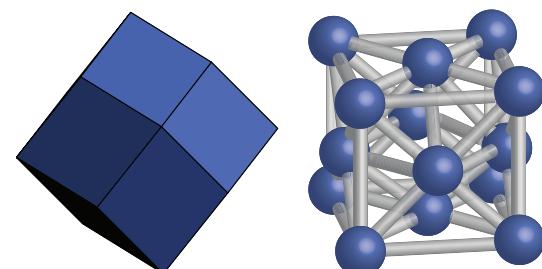
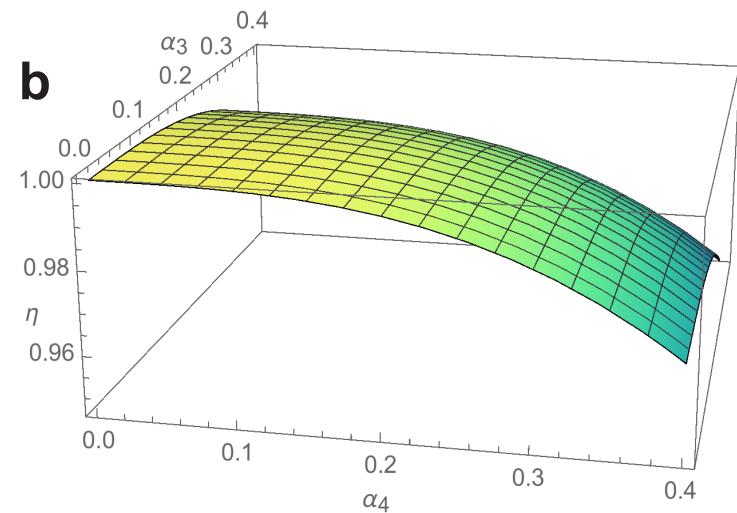
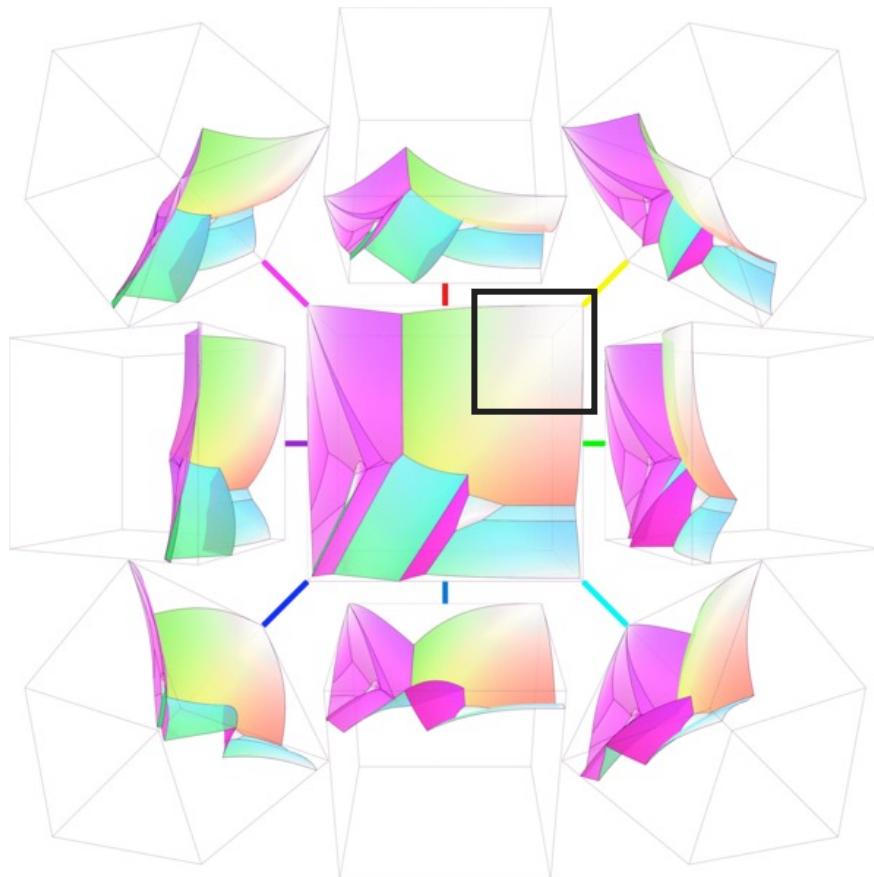
$$\left(\frac{\partial \mu_i}{\partial P}\right)_{N,T,\alpha_j \neq i} = \frac{1}{\phi^2} \left(\frac{\partial \phi}{\partial \alpha}\right)_{N,P,T,\alpha_j \neq i}$$





This relation tells us:  $\left(\frac{\partial \mu_i}{\partial P}\right)_{N,T,\alpha_j \neq i} = \frac{1}{\phi^2} \left(\frac{\partial \phi}{\partial \alpha}\right)_{N,P,T,\alpha_j \neq i}$

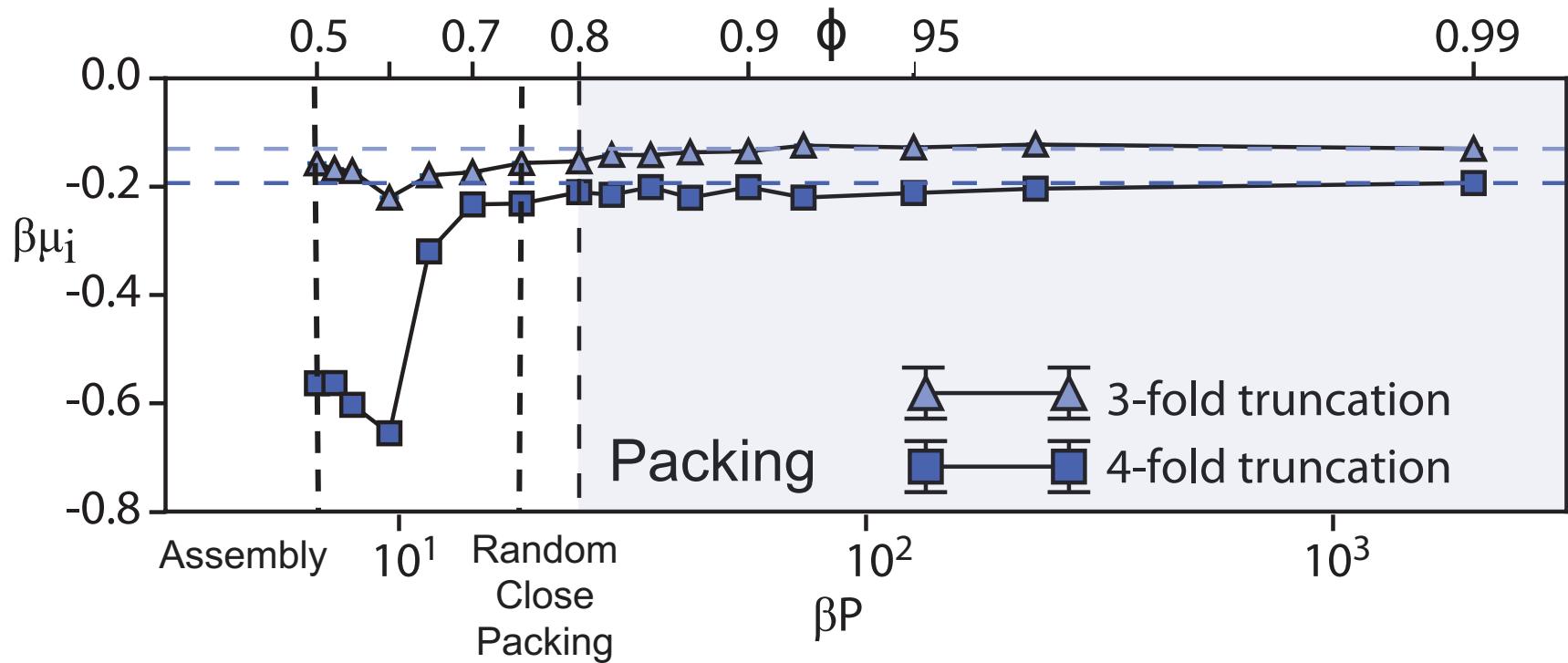
the thermodynamics of our system is consistent with packing when the derivative of the alchemical potential with respect to pressure equals that of our packing surface with respect to shape.





Looking at alchemical potential as a function of pressure, we can identify the onset of “packing” behavior.

$$\mu_i = \frac{1}{N} \left( \frac{\partial F}{\partial \alpha_i} \right)_{N, \eta, T, \alpha_j \neq i}$$



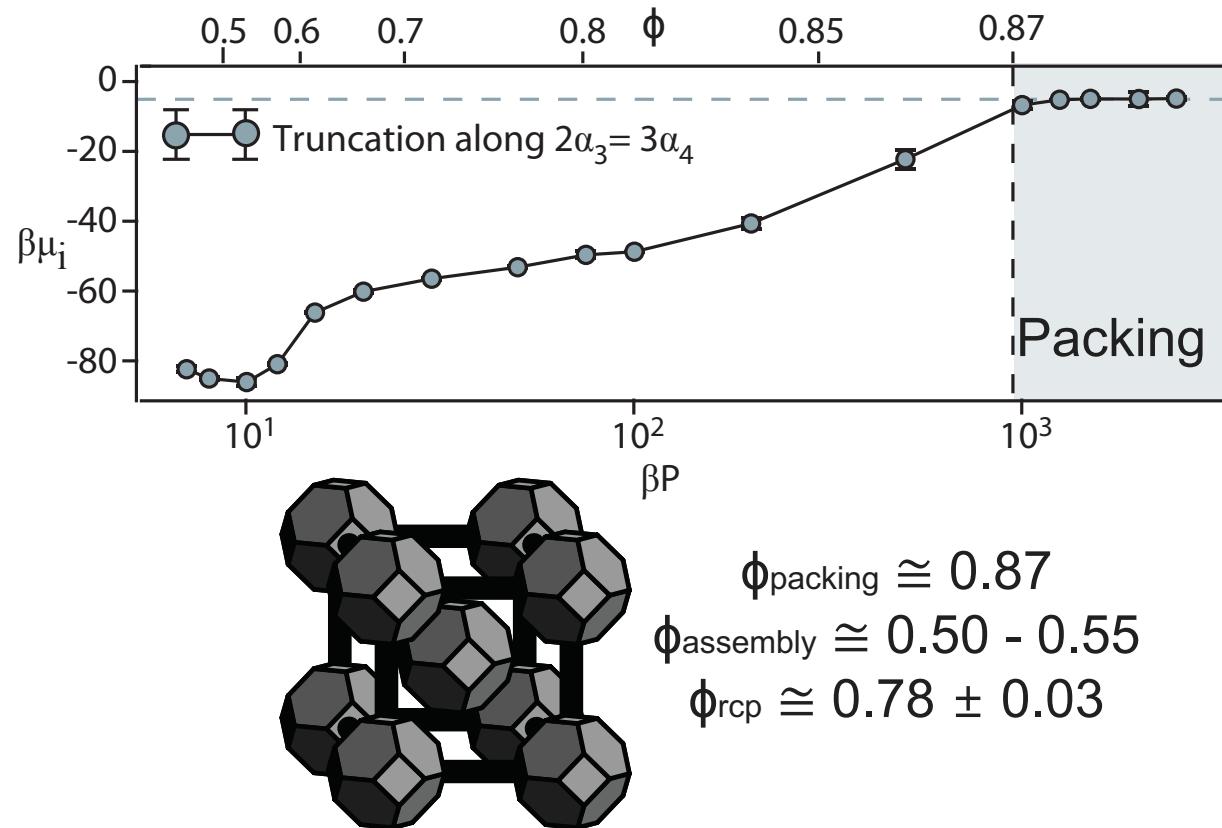
Here, NPT simulations were run until equilibrium to determine the Bennett acceptance ratio of “shape” moves. Using this acceptance ratio, we could compute the slope of the free energy surface with respect to shape. RCP densities were computed using Phys Rev Lett 99:155501.

$\phi_{\text{packing}} \cong 0.80$   
 $\phi_{\text{assembly}} \cong 0.50$   
 $\phi_{\text{rcp}} \cong 0.76 \pm 0.03$



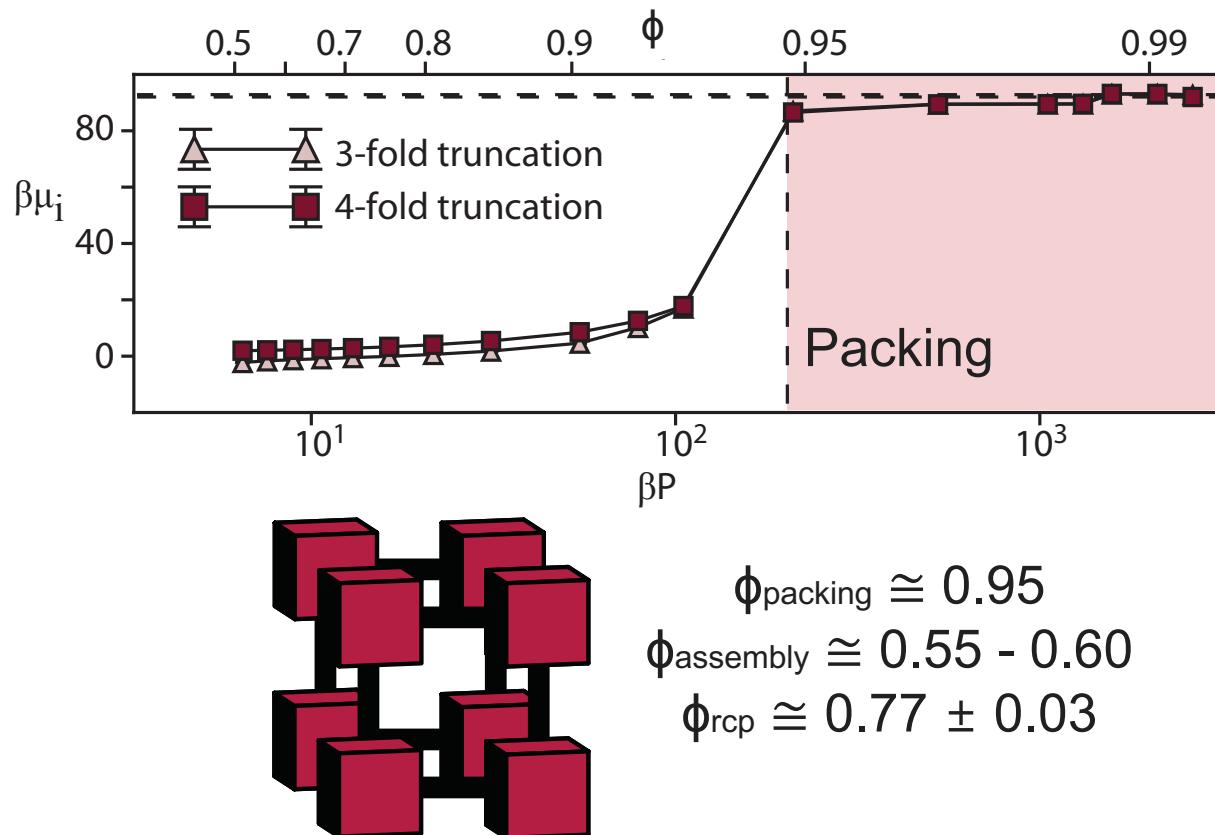


Thus, we can see that our system is not exhibiting the signature of *thermodynamic packing* when self-assembly occurs.





Thus, we can see that our system is not exhibiting the signature of *thermodynamic packing* when self-assembly occurs.



$$\phi_{\text{packing}} \cong 0.95$$

$$\phi_{\text{assembly}} \cong 0.55 - 0.60$$

$$\phi_{\text{rcp}} \cong 0.77 \pm 0.03$$





**Given that packing is not the driving force in  
colloidal self-assembly, then how can we  
understand the complex interactions that lead  
to crystallization?**

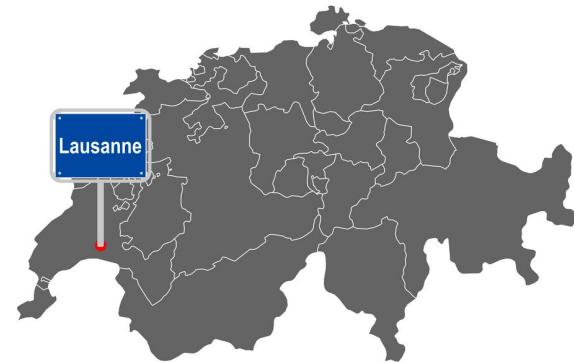




Here, I'll switch to another area where packing is often invoked: molecular crystallization.



2016, University of Michigan

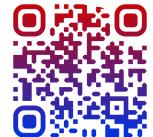
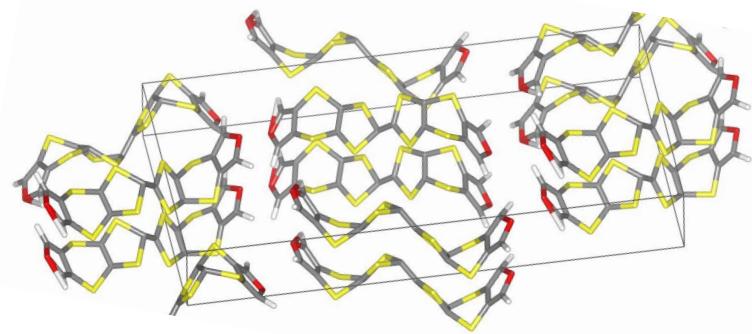
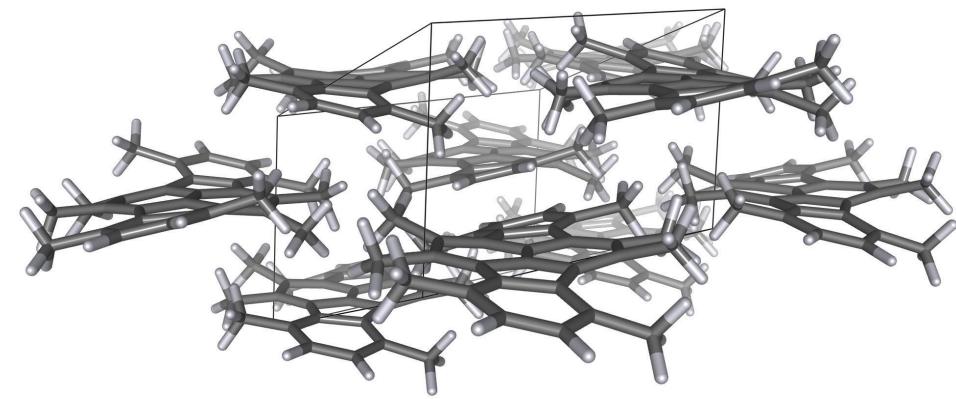
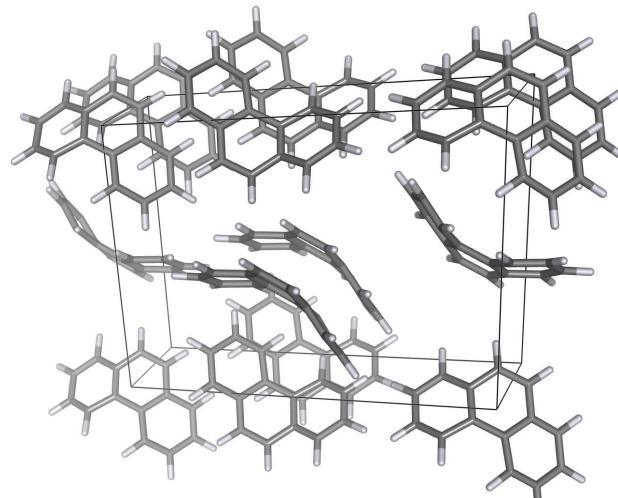
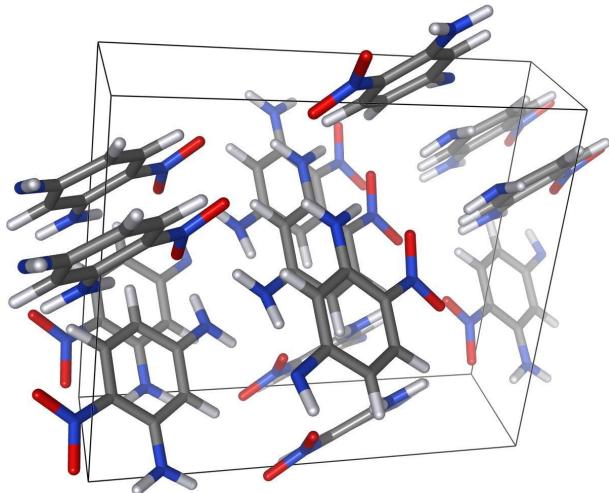


2022, École Polytechnique Fédérale de Lausanne (EPFL)





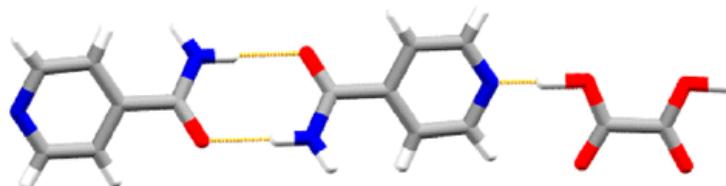
**Given a set of molecular chemistries, which forces are most responsible for binding the crystal together in solid form?**



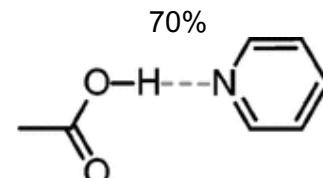


# To understand this question, it's useful to understand the typical methods for understanding *synthon* and *tecton* hierarchies.

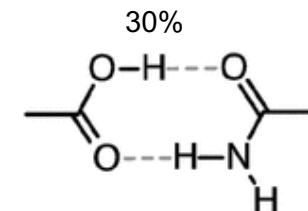
- In experiment, molecular packings are often rationalized by supramolecular synthons – libraries of intermolecular interactions common to molecular crystals.
- The hierarchies of these synthons is typically determined by mixing molecules with the synthons of interest and observing the interactions of the product.



(isonicotinamide)-(oxalic acid) cocrystal  
CSD ref. code: ULAWAF



70%  
carboxylic-pyridine,  
amide/amide



30%  
carboxylic-amide

- Computationally, this is done by querying CSD and determining the prevalence of each synthon interaction.





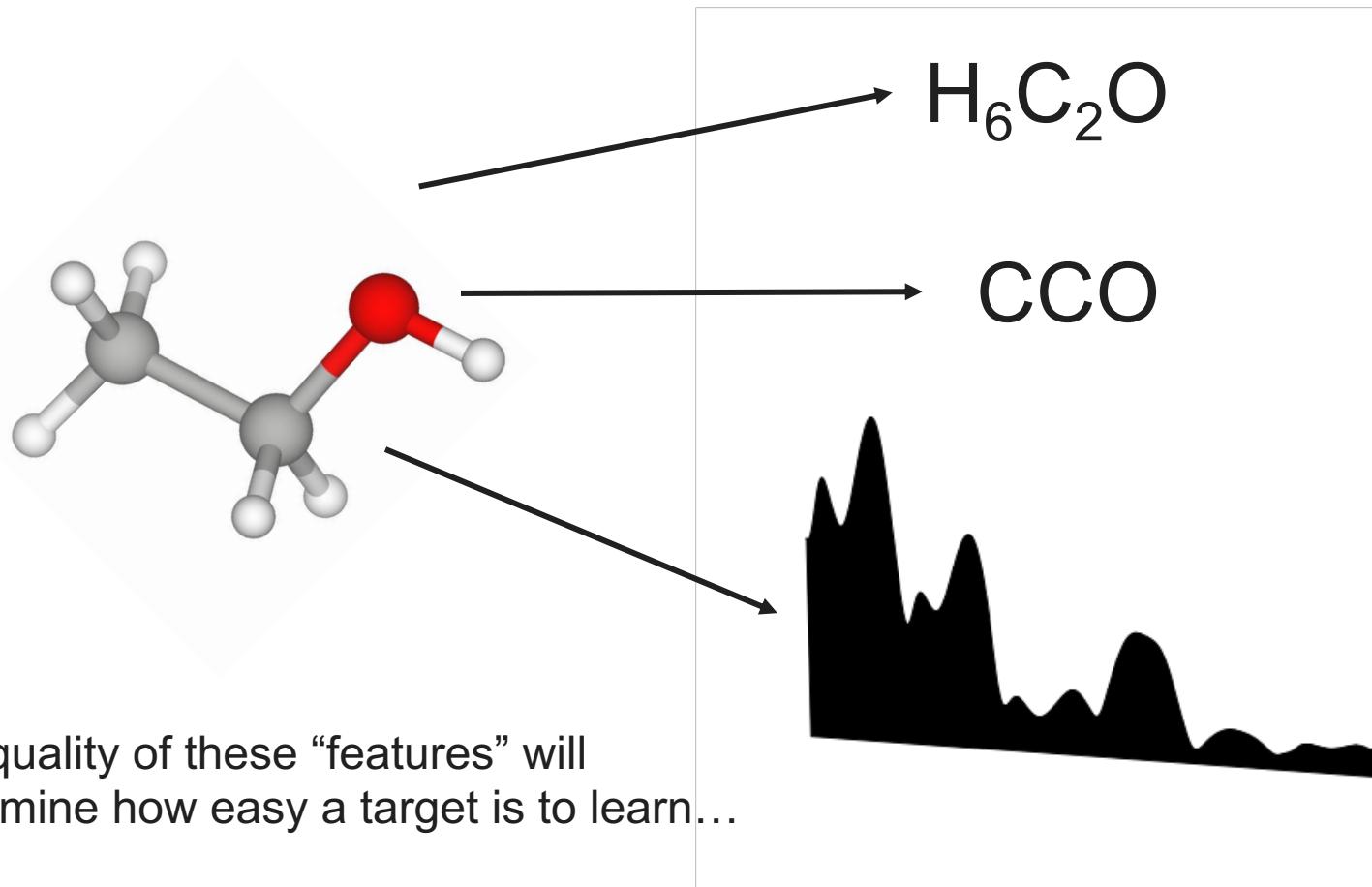
**Computing the relative energetics of these interactions is non-trivial, particularly when you must do so from first-principles to obtain accurate energetics.**

*Here is where we can leverage atom-centered approaches in machine learning.*





All machine learning models require that we translate our chemistry into “features” that can be used to learn off of.



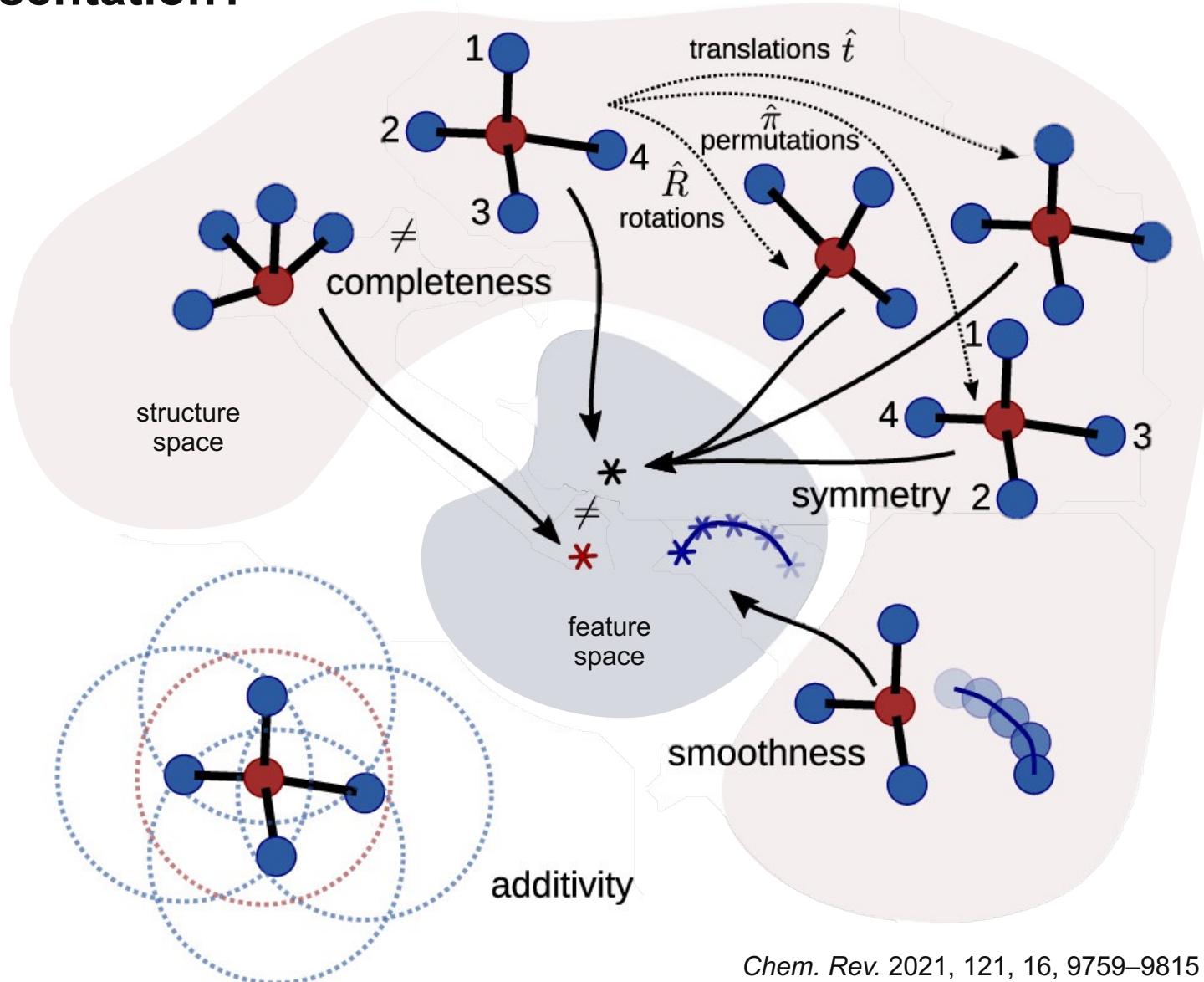
The quality of these “features” will determine how easy a target is to learn...

and how much you “learn” in the process!



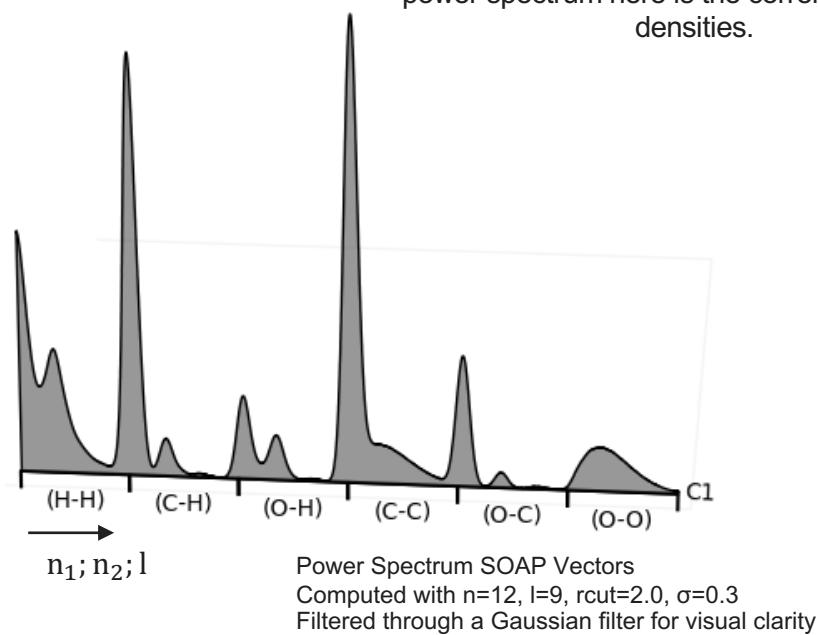
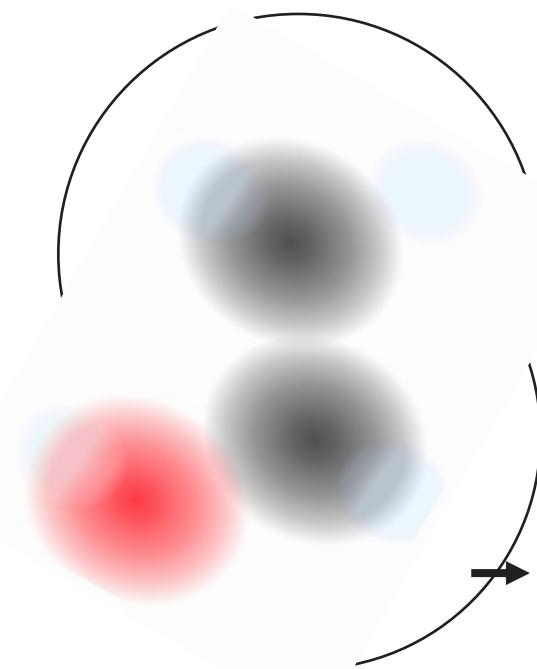


# In thermodynamic contexts, what do we want from a representation?





A collection of atoms can be represented by the combination of the atomic fingerprints.

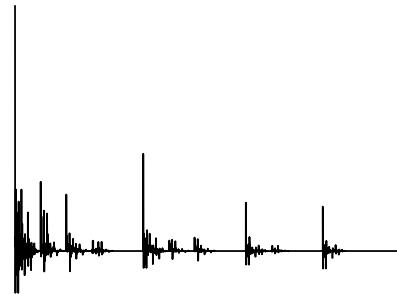
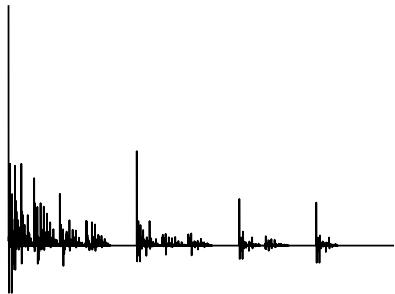


Each atom is represented as a Gaussian density expanded over radial bases (indexed by  $n$ ) and spherical harmonics (indexed by  $l$ ). The power spectrum here is the correlation of these densities.





| Representation       | Regression Result ( $R^2$ ) |
|----------------------|-----------------------------|
| Crystal Environments | 0.86                        |



the crystalline environments

- the gas-phase environments

= the "remnant" environments

**SOAP hyperparameters:** n=6, l=4, interaction cutoff (7.0 Å) with radial scaling, gaussian width (0.3 Å)  
**Regression hyperparameters:** 2823 train/ 558 test, sklearn.RidgeCV(cv=5, alphas=np.logspace(-12,-3,20))





**By using an additive descriptor, we can estimate the contribution of each motif to the cohesive energy.**



the “remnant” fingerprint  
averaged over the crystal



weights

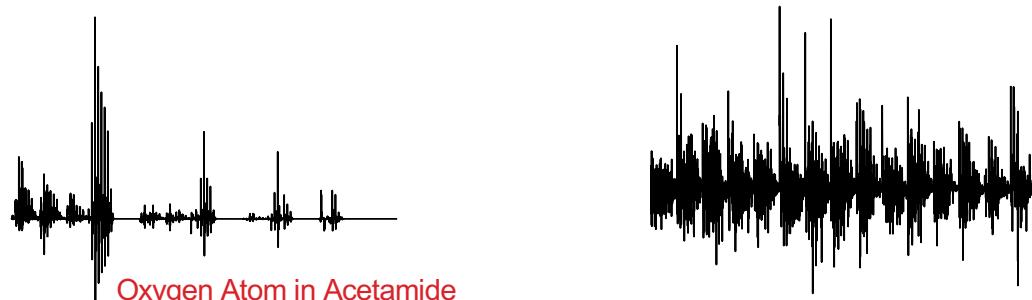


binding energy  
of the crystal





By using an additive descriptor, we can estimate the contribution of each motif to the cohesive energy.



the “remnant” fingerprint  
of each atom



weights

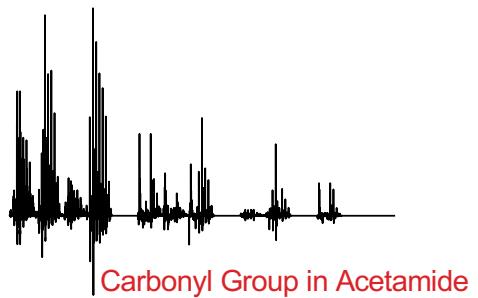


binding energy  
**contribution**  
**of the atom**





By using an additive descriptor, we can estimate the contribution of each motif to the cohesive energy.



Carbonyl Group in Acetamide

the “remnant” fingerprint  
**averaged over a**  
**collection of atoms**



weights

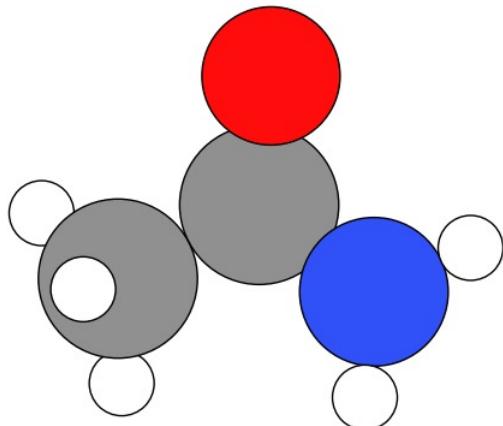


binding energy  
**contribution**  
**of the collection**

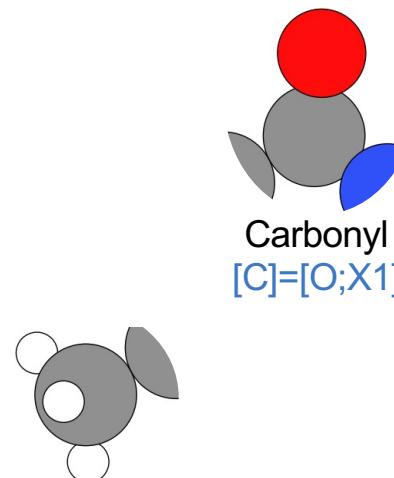




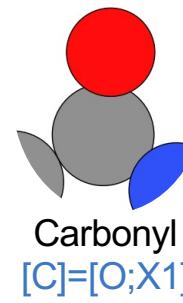
We used SMARTS string to automate labeling 48 popular molecular subgroups, resulting in approximately 70,000 motifs.



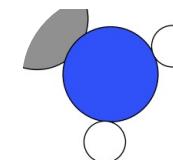
Canonical SMILES:  
 $\text{CC}(=\text{O})\text{N}$



Methyl  
 $[\text{C}; \text{H}3; \text{X}4]$



Carbonyl  
 $[\text{C}]=[\text{O}; \text{X}1]$

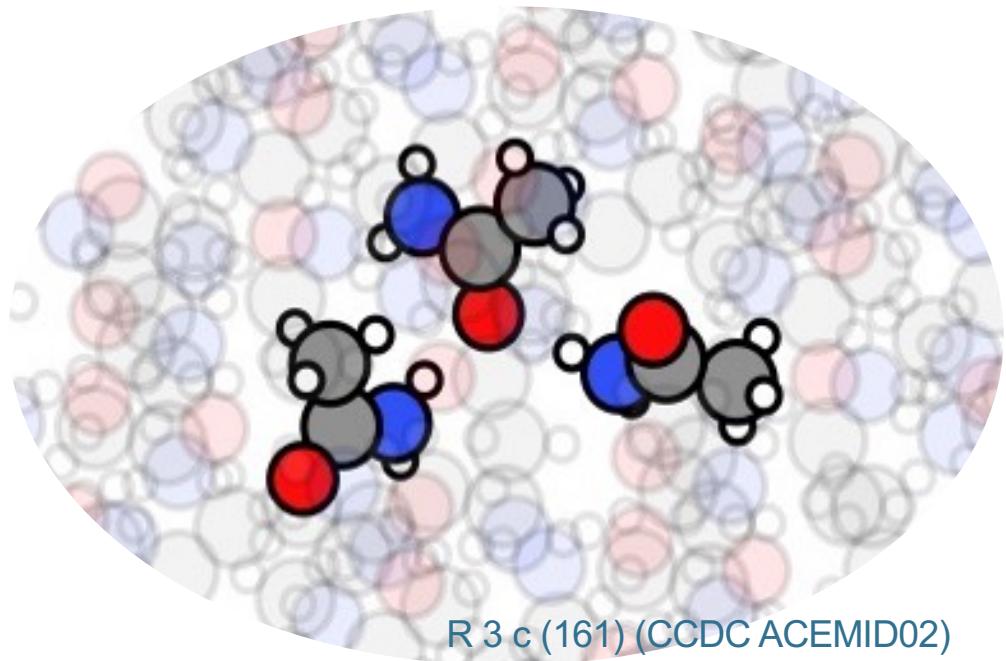
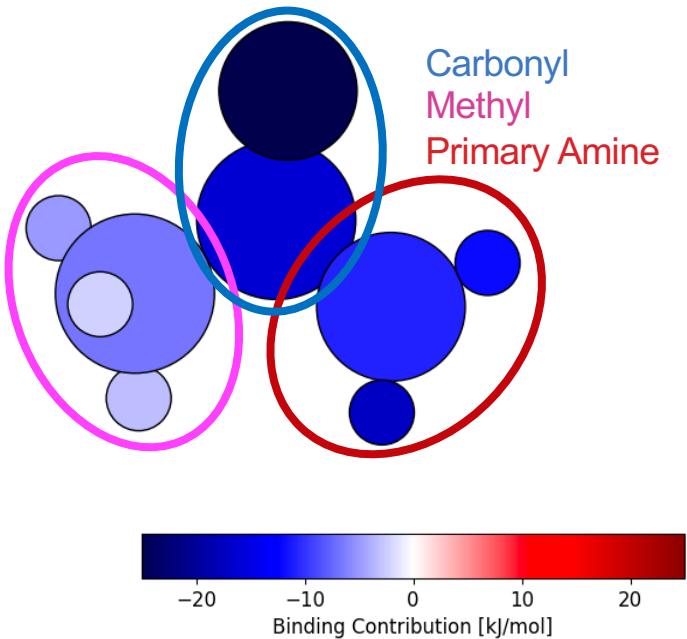


Primary Amine  
 $[\text{N}; \text{X}3; \text{H}2]$



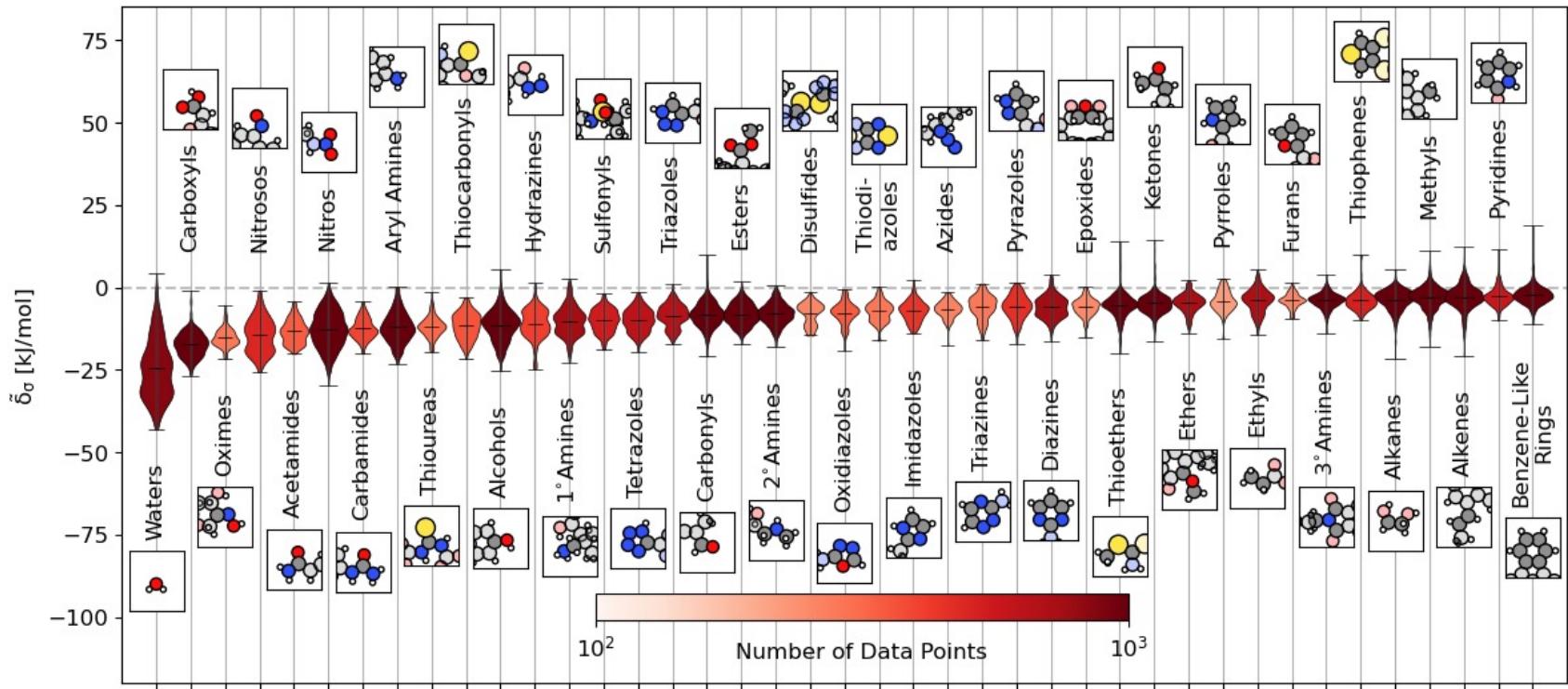


With these categorizations, we can see the contribution of each sub-group.



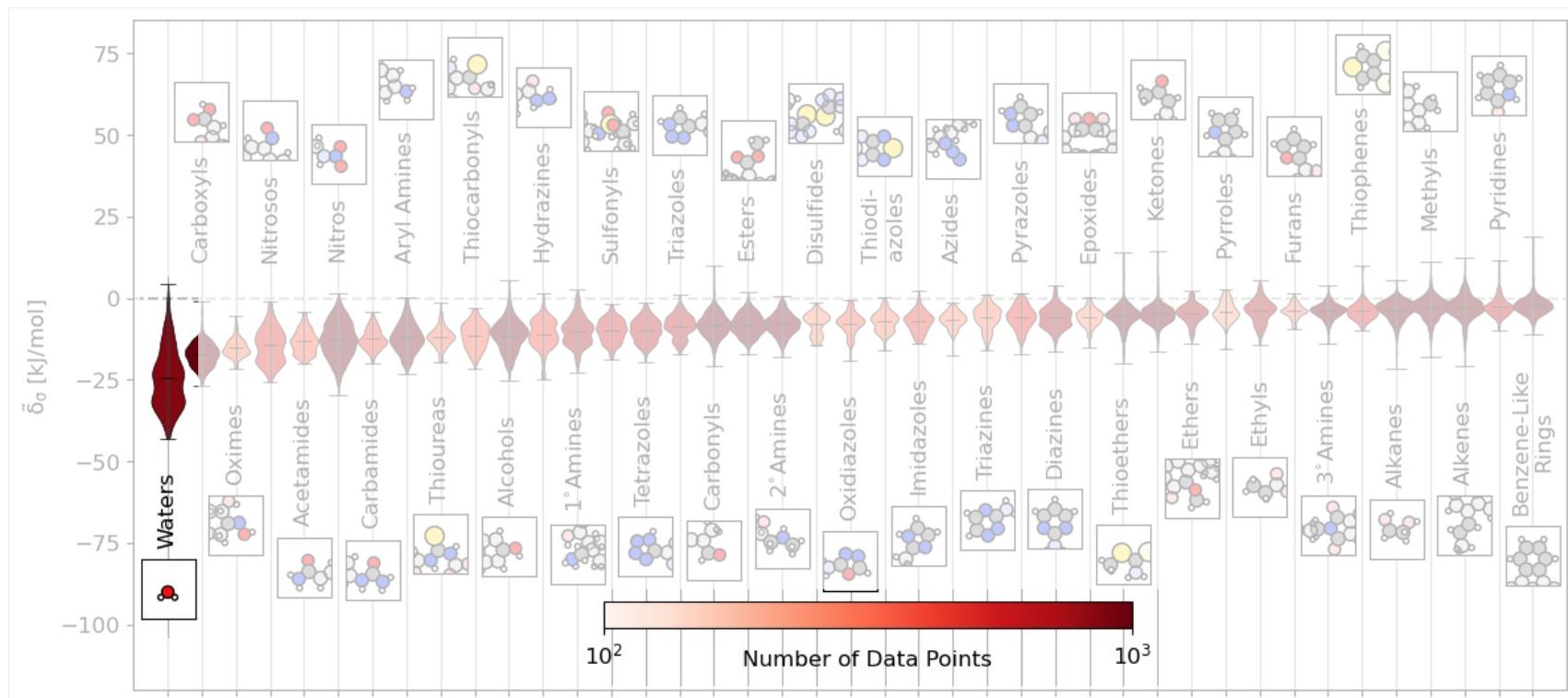


# Each different subgroup results in a range of contributions...



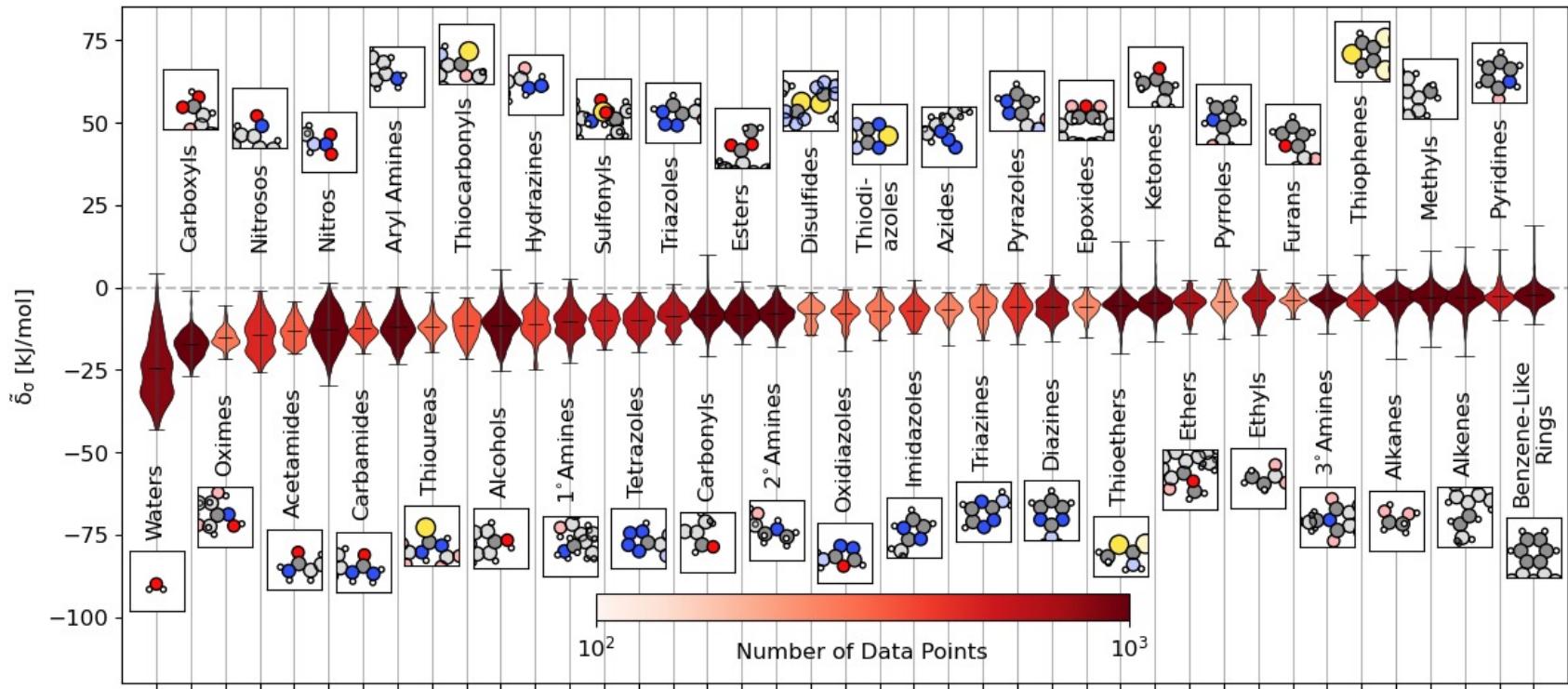


# Each different subgroup results in a range of contributions...





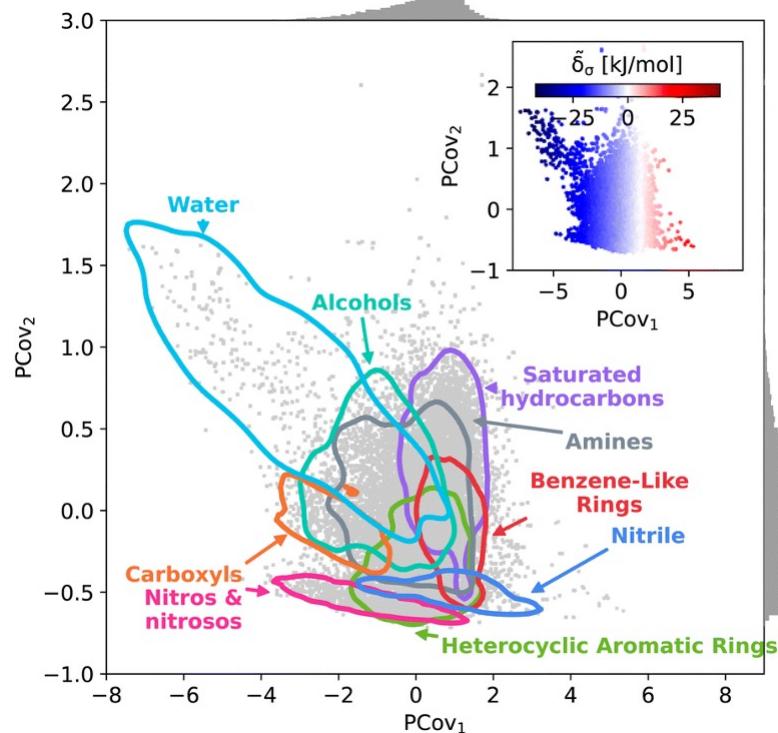
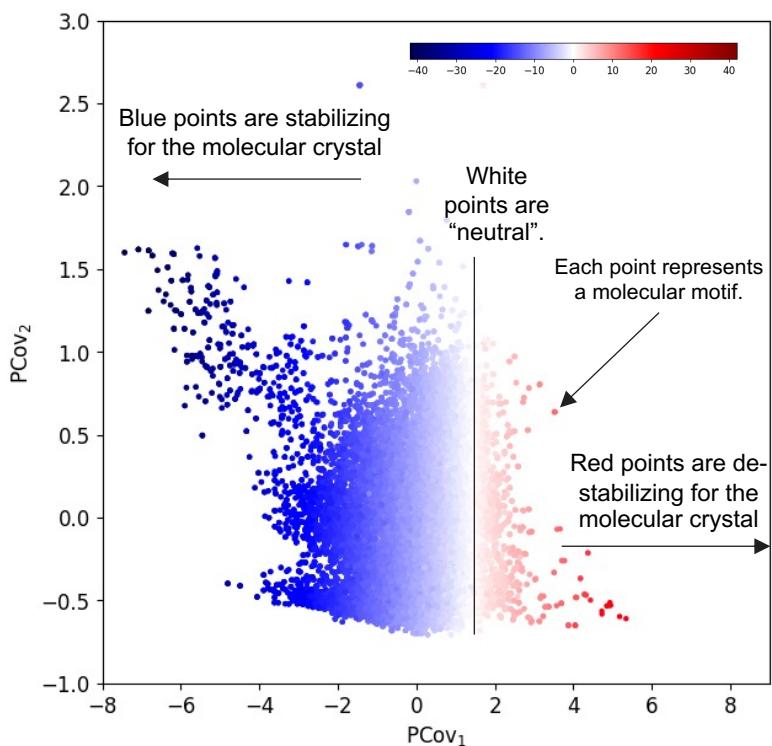
# Each different subgroup results in a range of contributions ... but how do we interpret these results?





A mapping technique to make structure-property maps.

# We use a PCovR mapping to understand the similarities of intermolecular interactions.



B. A. Helfrecht, RKC, G. Fraux, and M. Ceriotti. 2020 Mach. Learn.: Sci. Technol. 1 045021

S. de Jong, H.A.L. Kiers, Chemom. intell. lab. syst. 14 (1992) 155-164.

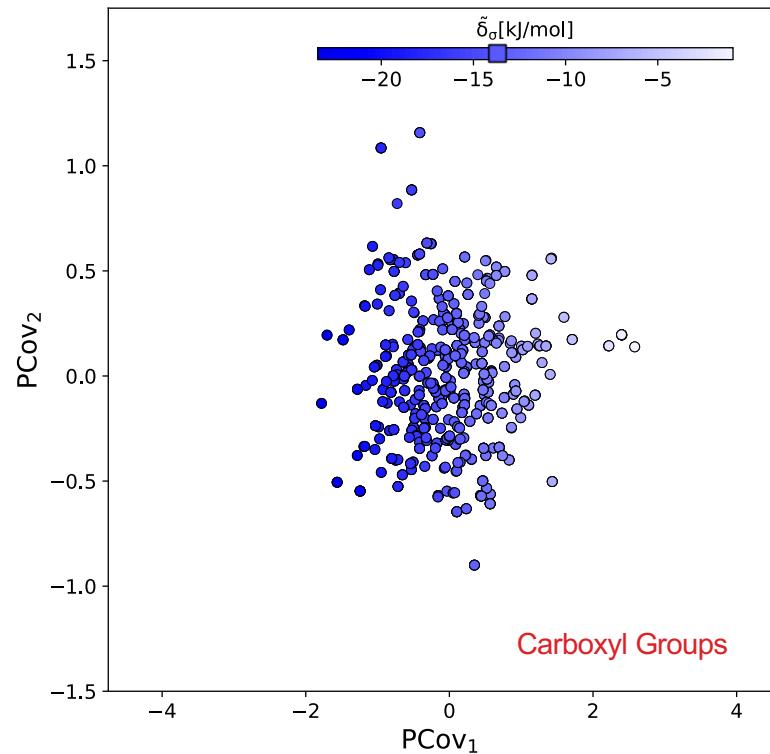
[scikit-matter.readthedocs.io](https://scikit-matter.readthedocs.io)





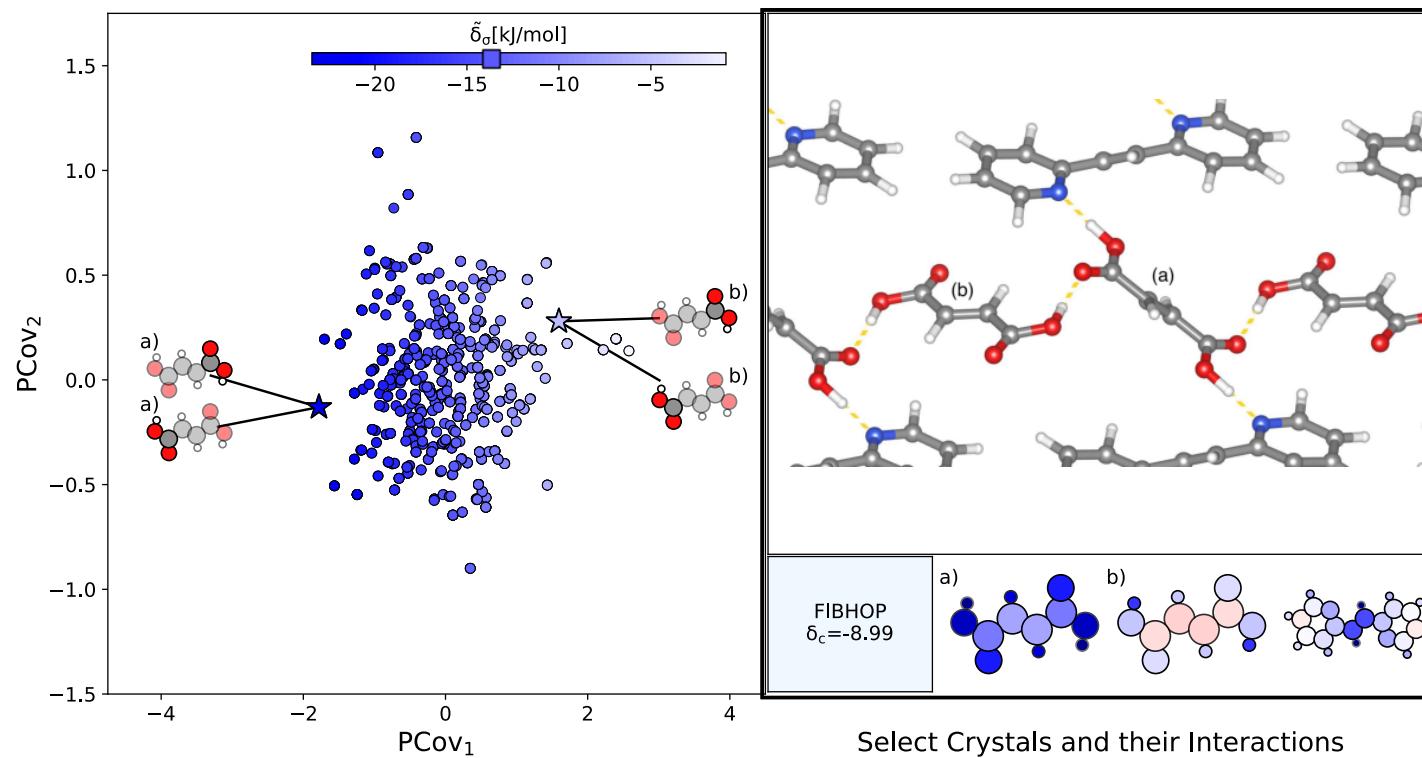
# We also can use a PCovR mapping to understand the interactions possible for a single functional group.

Here the map is showing the structure-property relationship for carboxyl groups – what defines a good environment for carboxyls?



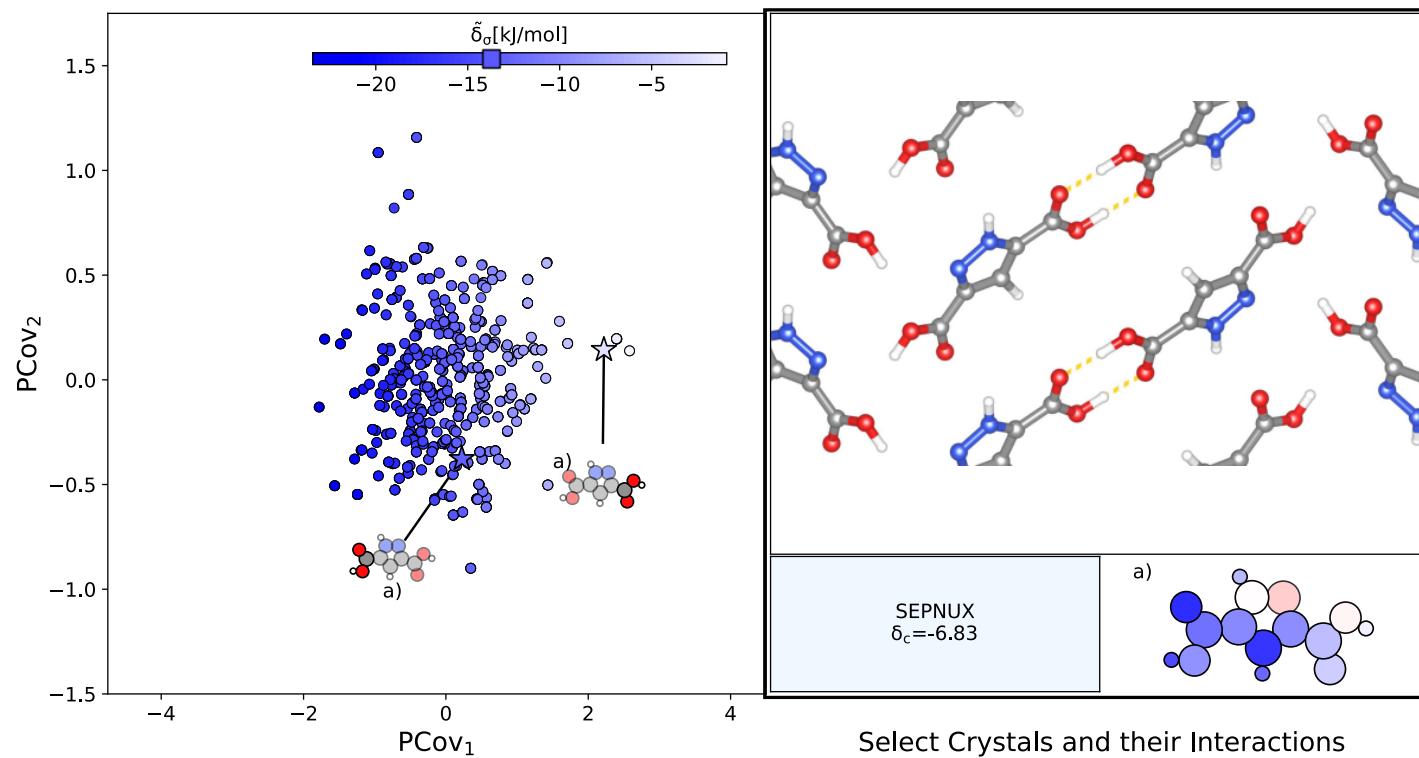


We also can use a PCovR mapping to understand the interactions possible for a single functional group.



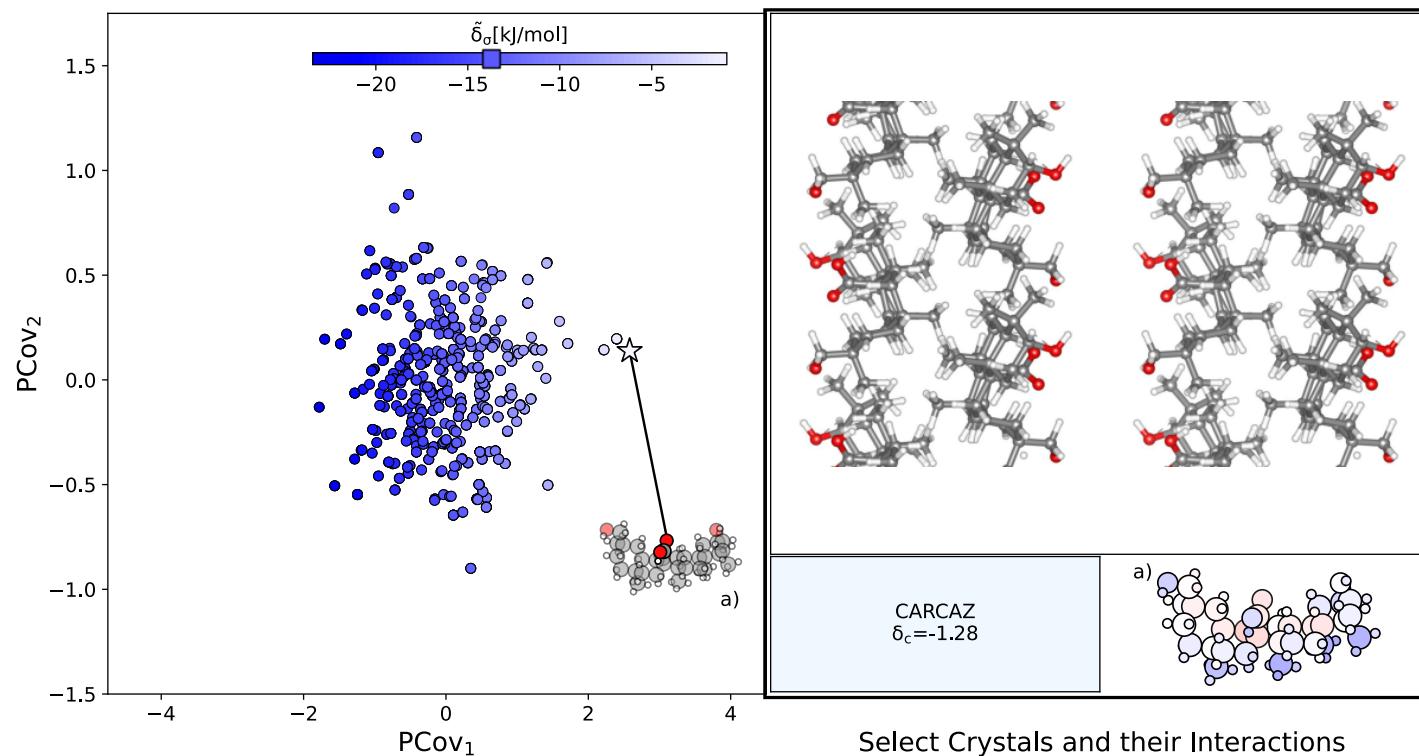


We also can use a PCovR mapping to understand the interactions possible for a single functional group.



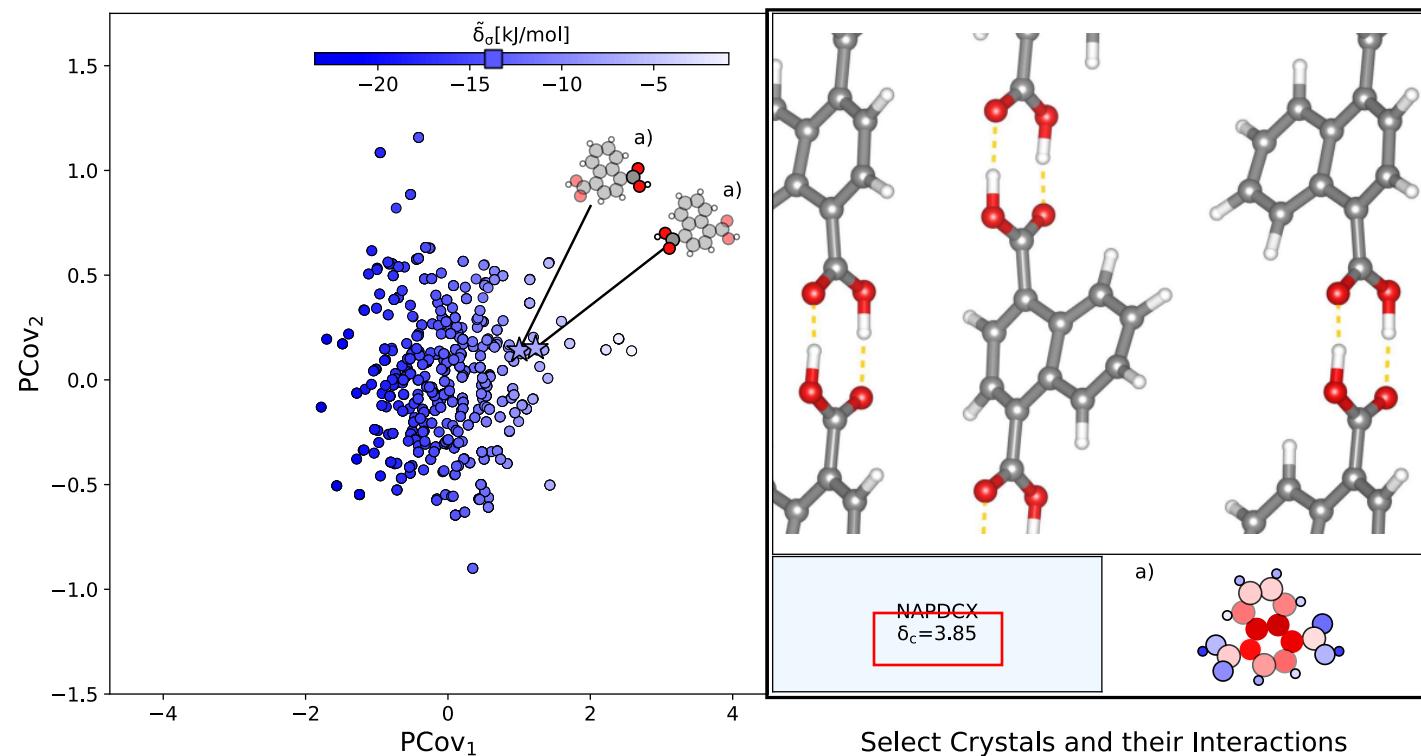


We also can use a PCovR mapping to understand the interactions possible for a single functional group.





We also can use a PCovR mapping to understand the interactions possible for a single functional group.



# Disentangling the impact of packing in colloidal and molecular self-assembly



- "Relevance of packing to colloidal self-assembly." **RKC** et al., *Proceedings of the National Academy of Sciences* 115.7 (2018): 1439-1444.
- *Mapping techniques for structure-property mappings*: B. A. Helfrecht, **RKC**, et al. 2020 *Mach. Learn.: Sci. Technol.* 1 045021.
- Interactive visualization software for unsupervised learning. G. Fraux, **RKC**, et al. 2020 *JOSS* 5(51), 2117.
- *Unsupervised Learning for Quantum Chemistry*: **RKC**, S. De. 2022, Elsevier.
- **RKC**, et al., 2023 *Chem. Sci.* **14**, 1272–1285.
- **RKC**, et al, *Lattice energies and relaxed geometries for 2'707 organic molecular crystals and their 3'242 molecular components.*, Materials Cloud **2023.5** (2023), doi: [10.24435/materialscloud:71-21](https://doi.org/10.24435/materialscloud:71-21). Visualization at: <https://molmotifs.matcloud.xyz/>
- scikit-learn compatible software for chemistry and materials: A. Goscinski, ..., **RKC**, 2023 Open Research Europe, 3(81).

If current trends do not change, fields such as chemical engineering and materials science will not reach gender parity any time soon. Why is this?

What can we do?

Pártay, L.B., Teich, E.G. & Cersonsky, R.K.C *Not yet defect-free: the current landscape for women in computational materials research.* *npj Comput Mater* 9, 98 (2023).  
<https://doi.org/10.1038/s41524-023-01054-z>

My group is hiring PhDs and postdocs!

In the Cersonsky Lab, our ultimate goal is to build a unified machine learning feature space and methodology for studying the thermodynamical behavior of multiscale and hierarchical materials.

[rose.cersonsky@wisc.edu](mailto:rose.cersonsky@wisc.edu)

