MSRI
Mathematical Sciences Research Institute

17 Gauss Way Berkeley, CA 94720-5070 p: 510.642.0143 f: 510.642.8609 www.msri.org

NOTETAKER CHECKLIST FORM

(Complete one for each talk.)

Name: Malgorzata Marciniak **Email/Phone:** mmarciniak@lagcc.cuny.edu 5734620411

Speaker's Name: Maciej Haranczyk

Talk Title: Facilitating Discovery of Nanoporous Materials for Energy Applications with Porosity Analysis Tools

Date: 10 /03 /2018 **Time:** 11 : ___ am / pm (circle one)

Please summarize the lecture in 5 or fewer sentences:

The presentation offers a view at the structure of nanoporous materials from the perspective of how mathematics, in particular computational geometry and topology, can be in analysis of the shape and structure of materials. We want to combine different aspects and different tools to enable this discovery. i.e., discuss new models for porosity analysis, use molecular simulations to predict absorption and create algorithms for discovery (sampling techniques, machine learning or optimization technique) of finding the best structures for particular applications.

CHECK LIST

(This is **NOT** optional, we will **not pay** for **incomplete** forms)

- ☑ Introduce yourself to the speaker prior to the talk. Tell them that you will be the note taker, and that you will need to make copies of their notes and materials, if any.
- ☑ Obtain ALL presentation materials from speaker. This can be done before the talk is to begin or after the talk; please make arrangements with the speaker as to when you can do this. You may scan and send materials as a .pdf to yourself using the scanner on the 3rd floor.
  - **Computer Presentations**: Obtain a copy of their presentation
  - **Overhead**: Obtain a copy or use the originals and scan them
  - **Blackboard**: Take blackboard notes in black or blue **PEN**. We will **NOT** accept notes in pencil or in colored ink other than black or blue.
  - **Handouts**: Obtain copies of and scan all handouts
- ☑ For each talk, all materials must be saved in a single .pdf and named according to the naming convention on the "Materials Received" check list. To do this, compile all materials for a specific talk into one stack with this completed sheet on top and insert face up into the tray on the top of the scanner. Proceed to scan and email the file to yourself. Do this for the materials from each talk.
- ☑ When you have emailed all files to yourself, please save and re-name each file according to the naming convention listed below the talk title on the "Materials Received" check list. (*YYYY.MM.DD.TIME.SpeakerLastName*)
- ☑ Email the re-named files to notes@msri.org with the workshop name and your name in the subject line.

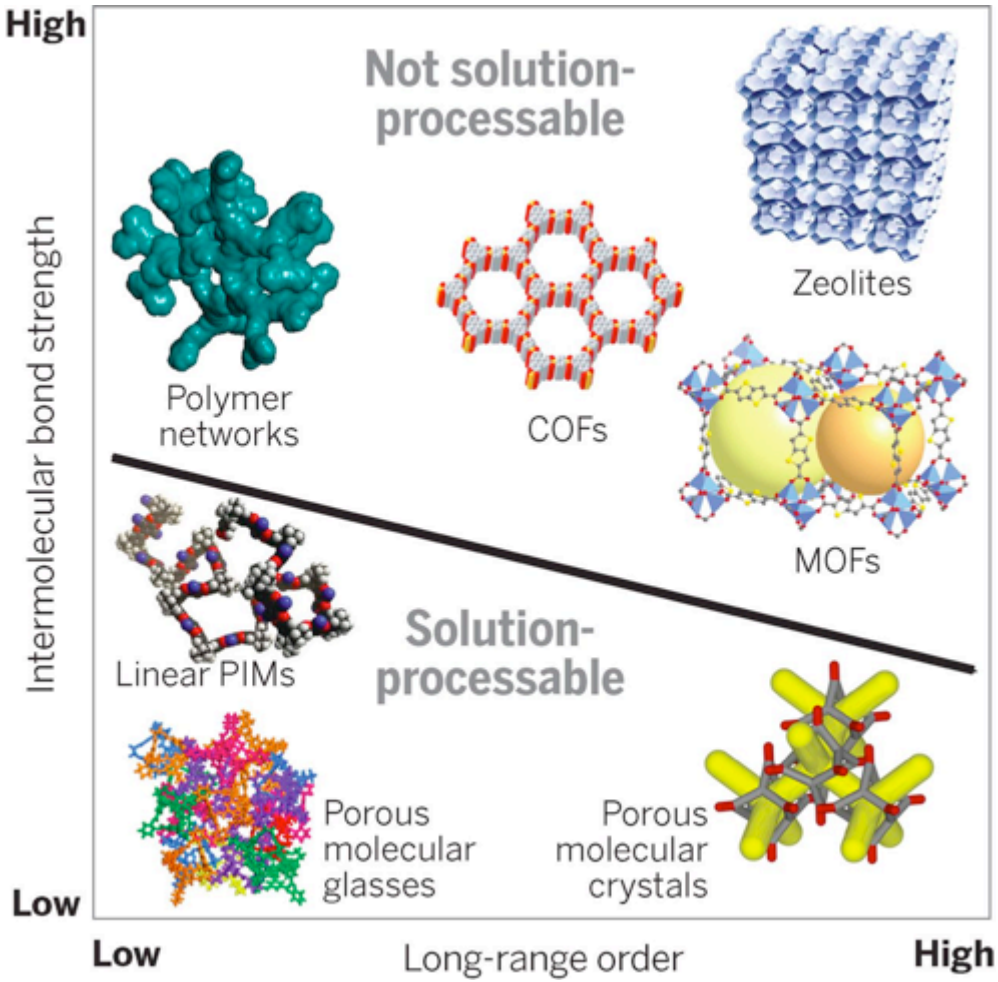
Facilitating Discovery of Nanoporous Materials for Energy Applications with Structure and Porosity Analysis Tools

Maciej Haranczyk

Cory Simon, Matt Witman, Nils Zimmermann, Jose Perez, Richard Martin, Berend Smit, Ismael Garcia, Marco Bernabei

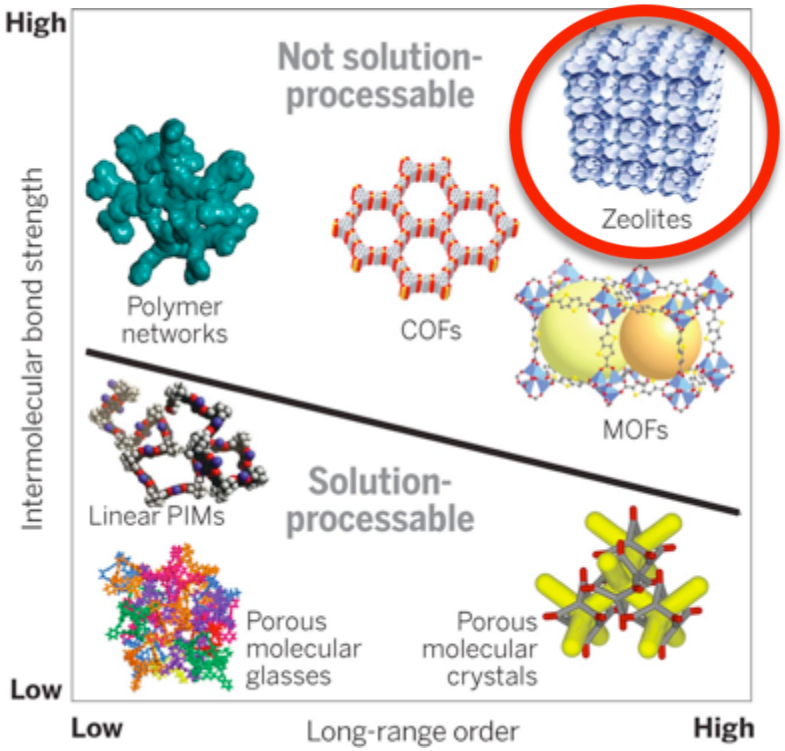
Berkeley, CA, October 2018

Nanoporous materials



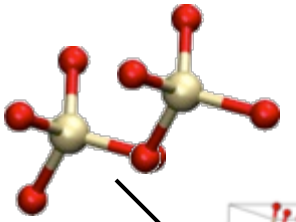
Slater&Cooper, Science 2015

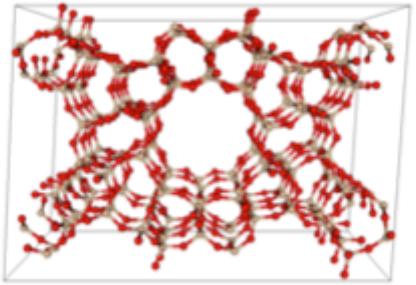
Diversity and modularity of nanoporous materials



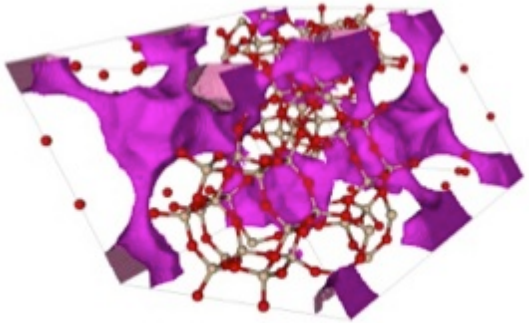
Zeolites

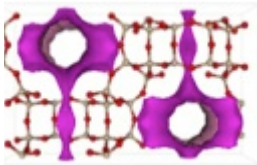
$(SiO_2)x$ building unit of zeolites

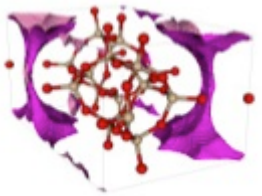




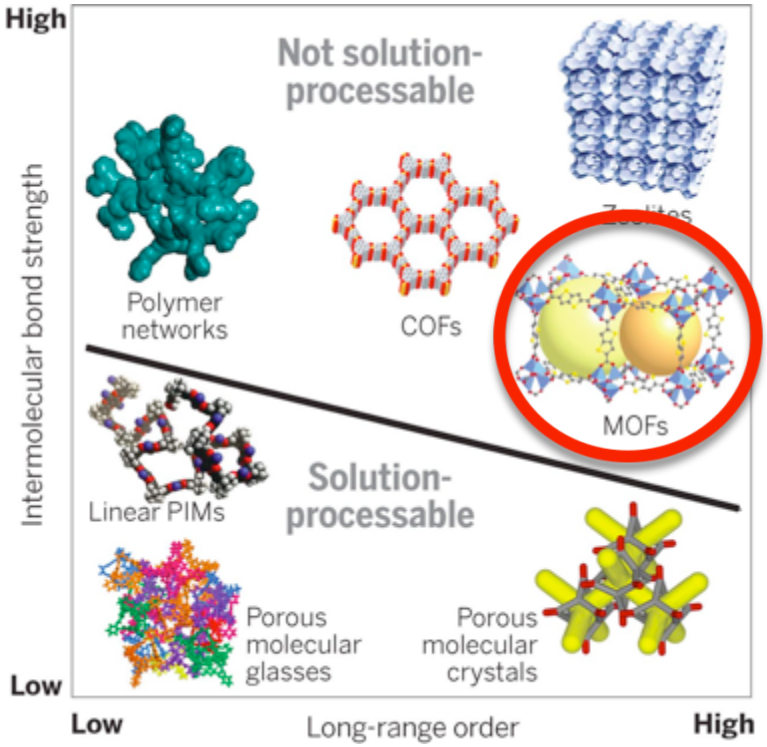
Diversity of zeolites (mainly topology, chemistry): ~1.5k made, ~3M predicted





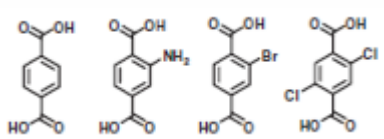


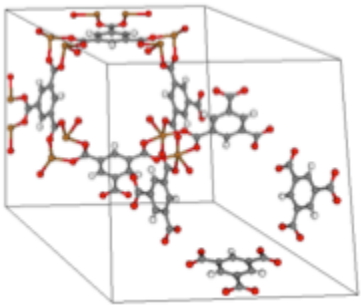
Diversity and modularity of nanoporous materials

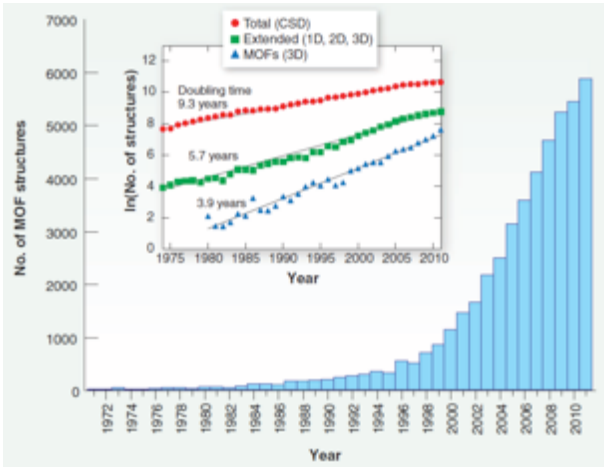


Metal Organic Frameworks (MOFs)

Metal (oxide) clusters +



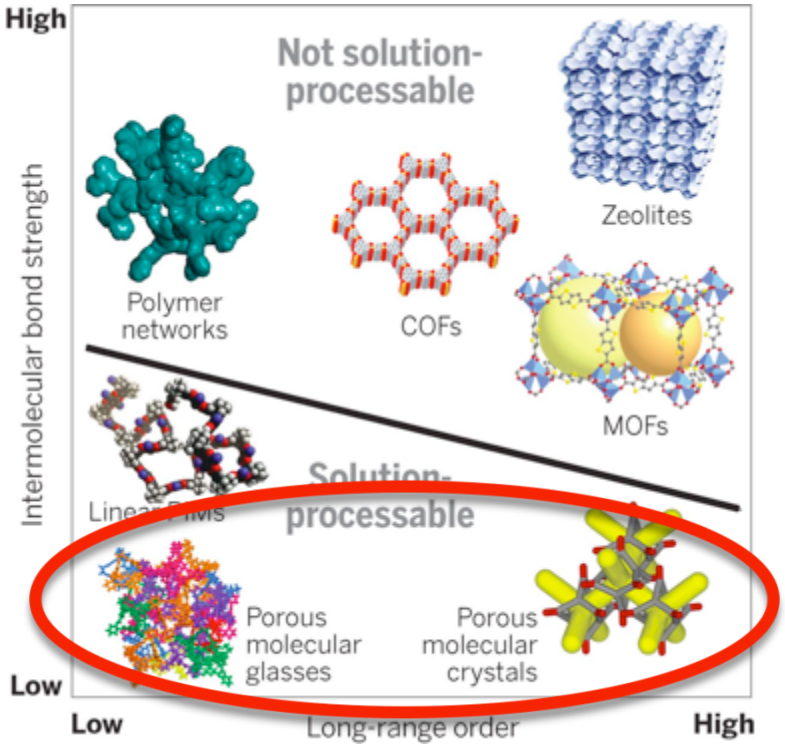




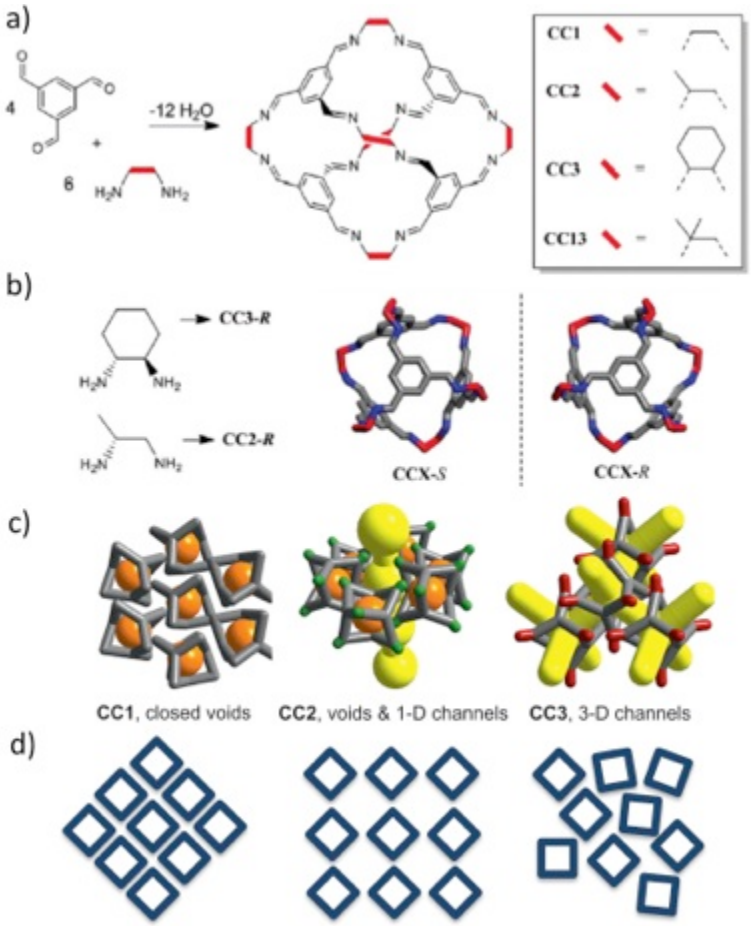
Furukawa et al. *Science* **2013**, 341, 974

Diversity of MOFs (topology, chemistry):
~10k made,
~0.5M predicted

Diversity and modularity of nanoporous materials



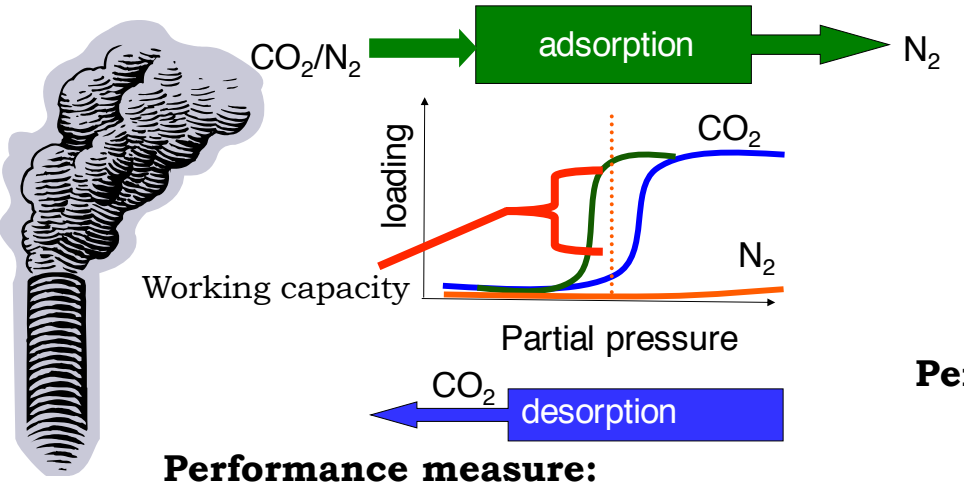
Diversity of PMC: ~50? made, 100 predicted ?



Application Context and Targets

CO_2 capture

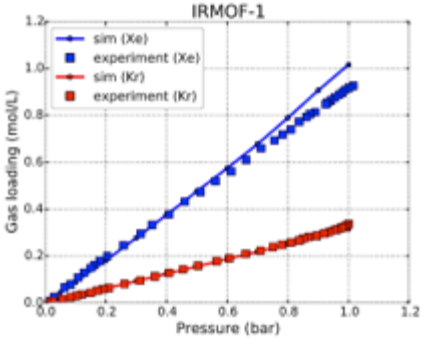
Flue gas: 1atm, 14% CO_2, 40°C



Performance measure:
Parasitic Energy= (energy cost) / (amount of CO_2 captured)

Xe/Kr Separations

Xe/Kr mixture: 20/80, 1atm total, 298K



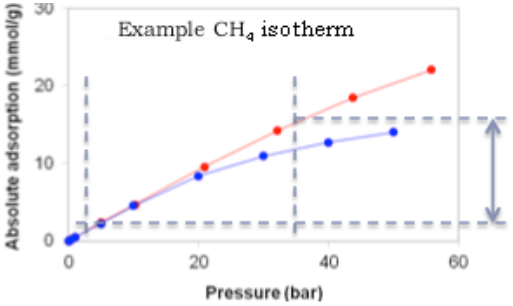
Performance measure: Selectivity + Material uptake

Other Targets

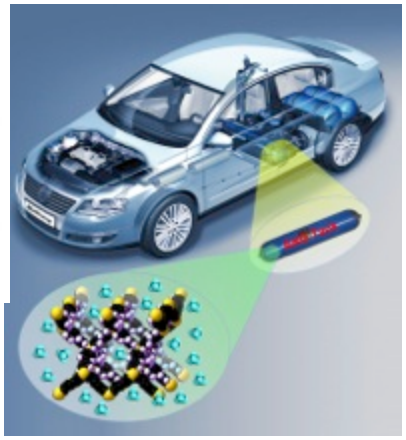
"Performance" measure (examples):

- General crystal design:
  - Large surface area (gravimetric/volumetric)
  - Pore size

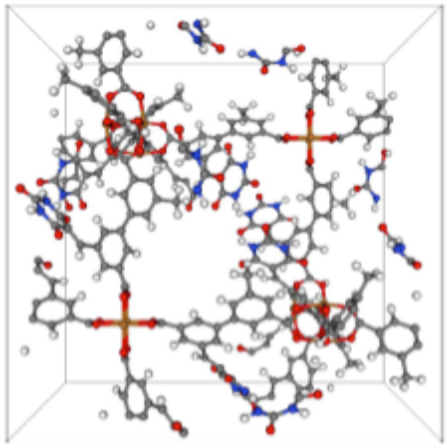
Methane storage



Performance measure:
Working capacity = Uptake@65bar-Uptake@5.8bar

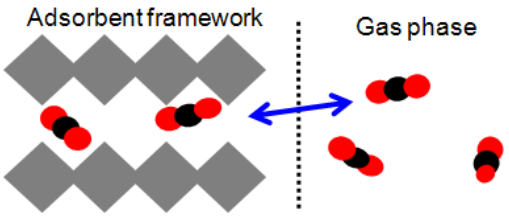


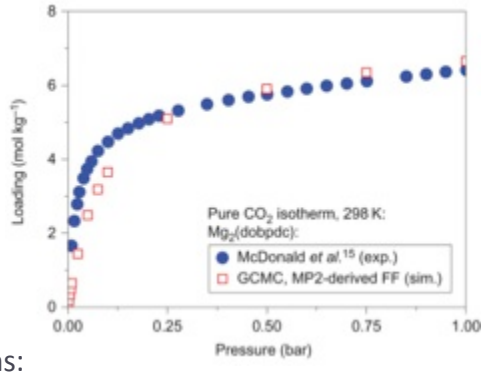
Property and Performance Prediction



Crystal structure (random example (MOF))

Molecular simulations can accurately model adsorption of guest molecules





Adsorption isotherms are computed using grand canonical Monte Carlo. Typical assumptions:

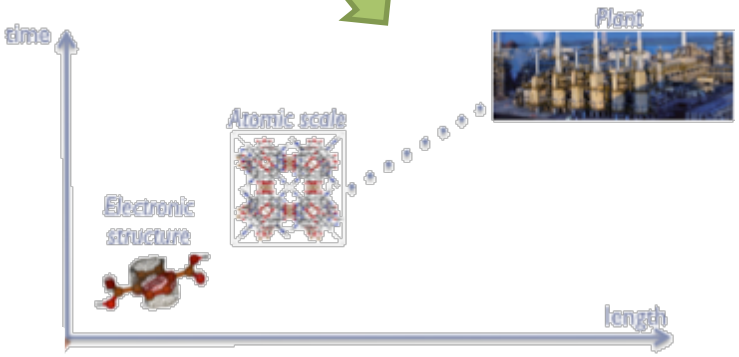
- Crystal structure is rigid,
- Classical Force-fields (Lennard-Jones + Coulomb terms)

Henry coefficient: $K_H = \beta \langle \exp(-\beta U_{ins}) \rangle$

Heat of adsorption: $H_A = \dfrac{\langle U_{ins} \exp(-\beta U_{ins}) \rangle}{\langle \exp(-\beta U_{ins}) \rangle}$

$\beta = 1/(kBT)$

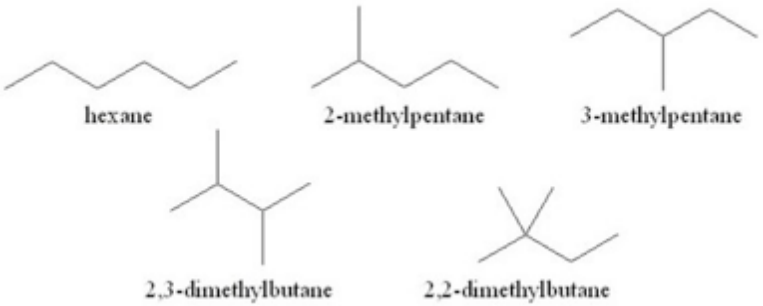
U_{ins} = test molecule insertion energy



Various scales need to be included to accurately predict performance, e.g. electronic structure methods are used to tune classical force-fields, and various engineering approaches are used to estimate material's performance

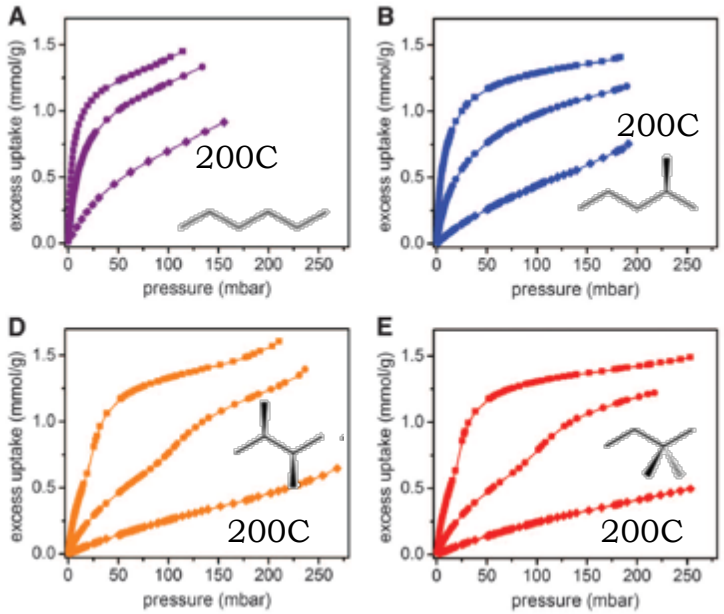
Hot Topics: Shape and Structure of Materials

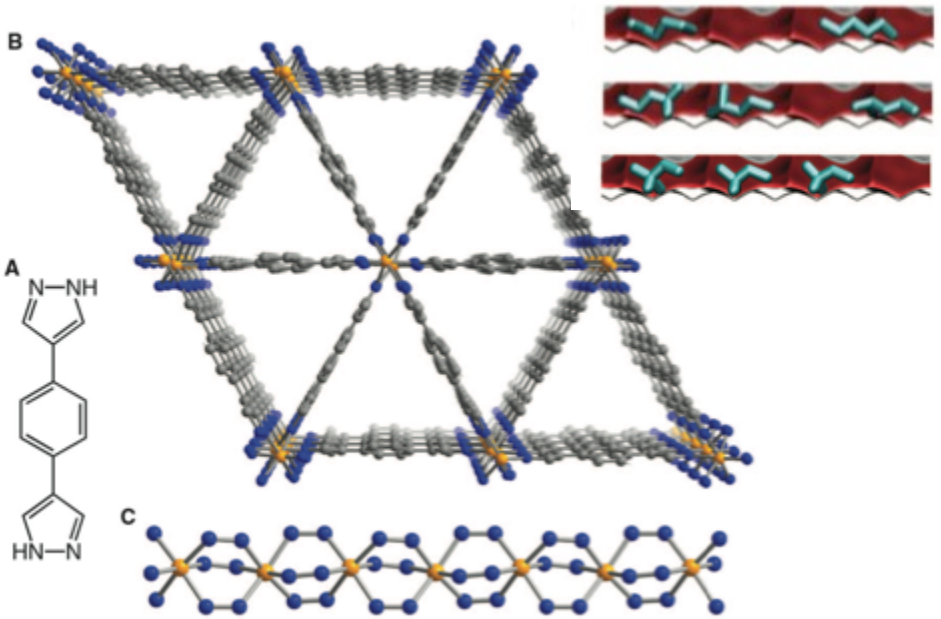
MOFs for Hexane Isomers Separations



Hexane isomers separation is important for petrochemical industry.

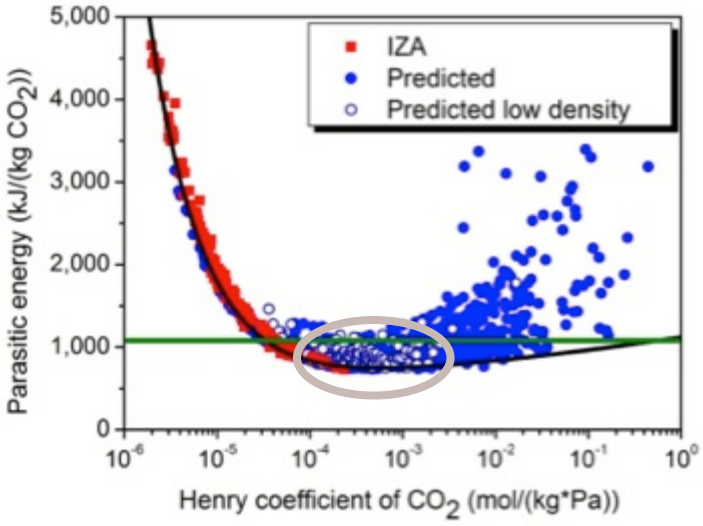
$Fe_2(BDP)_3$ MOF has been demonstrated to have high selectivity towards non-branched isomers.





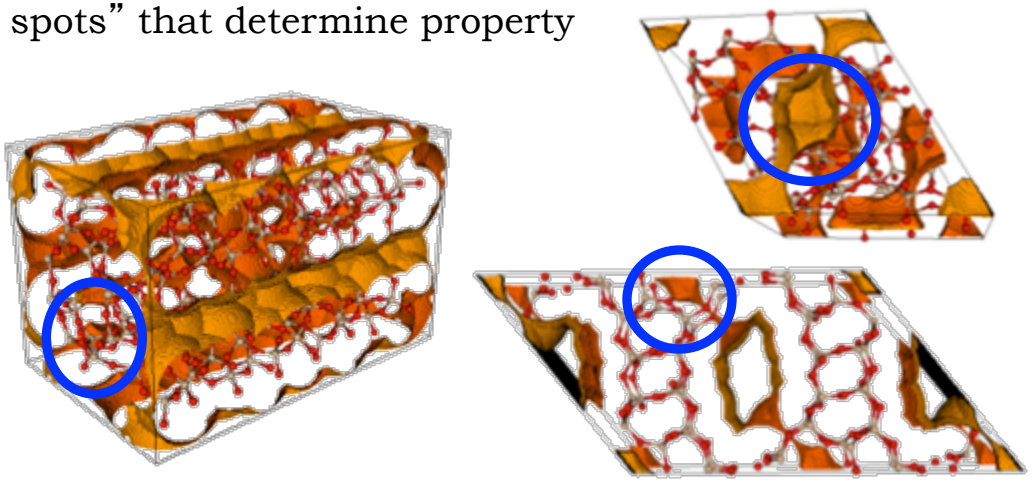
Zoey R. Herm et al. Science **340**, 960, (2013);

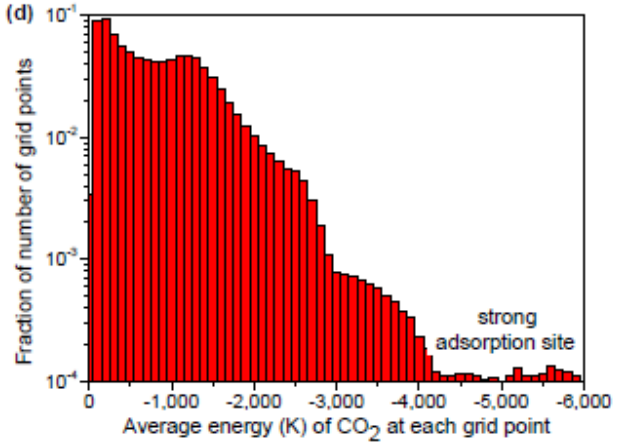
Best Performing Zeolites in CO_2 Capture Application



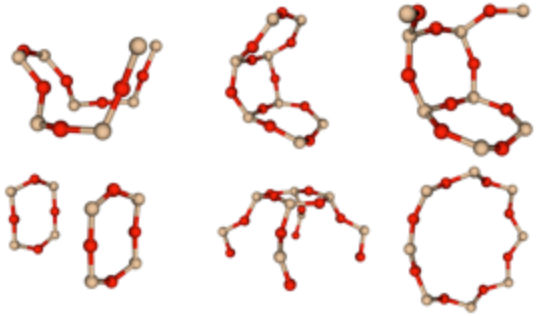
L.-C. Lin et al, Nature Materials, 2012

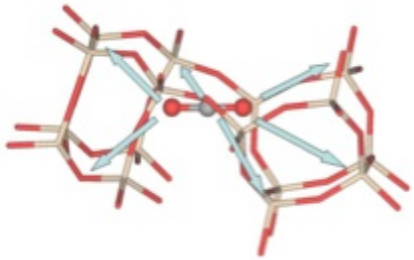
The most diverse of the zeolites with best performance exhibit common features – "sweet spots" that determine property





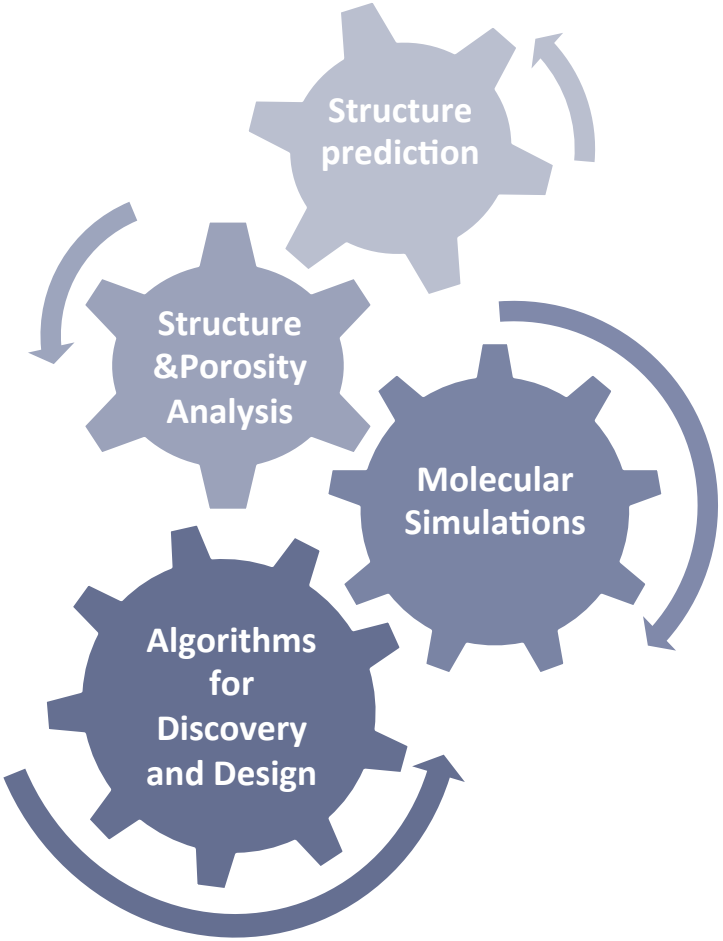
Anatomy of a binding site: two fragments arranged in a way to maximize interactions





Martin, et al , ChemPhysChem, 2012

Discovery Methodology Overview

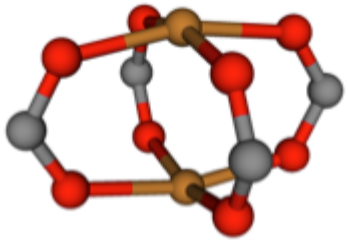


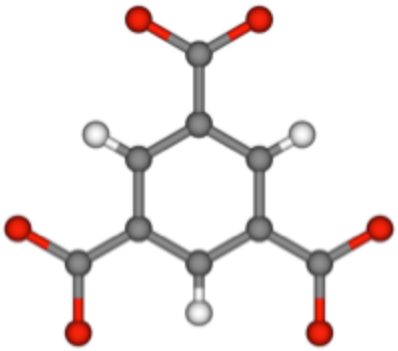
Towards Crystal Structure Prediction and Structure Enumeration

Towards Structure Prediction for Advanced Porous Materials

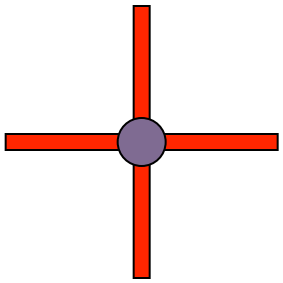
New classes of advanced porous materials such as **MOFs** permit unlimited structural variation through control of *topology and functionality*

Metal cations + *Organic bridging molecules* = *Framework*



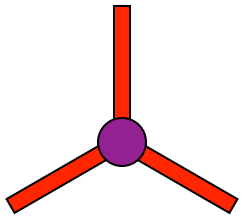


?



Square

+



Triangle

=

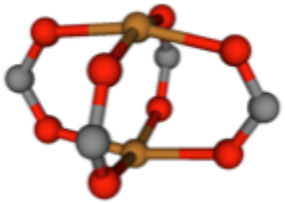
How can we reliably predict the structures of these advanced porous materials?

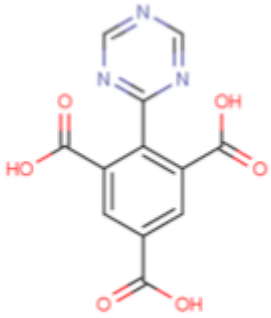
Topology-based Structure Prediction

RCSR

Reticular Chemistry Structure Resource

Example: predicting the structure of a MOF comprising square and triangular components





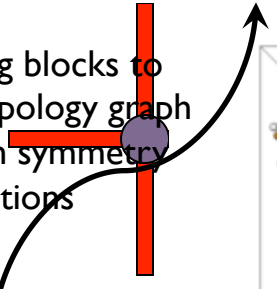
Building blocks:
4-connected metal (square) and
3-connected linker (triangle)

Ockwig *et al.*, *Acc. Chem. Res.* **2005**, *38*, 176
O' Keeffe *et al.*, *Acc. Chem. Res.* **2008**, *41*, 1782

Catalogue of possible topologies
rcsr.anu.edu.au

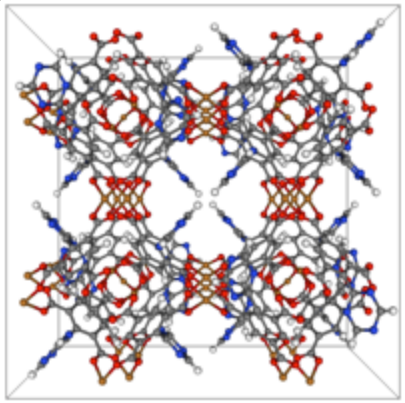
High-symmetry candidate:
tbo topology or "net"

Map building blocks to vertices of topology graph and perform symmetry operations

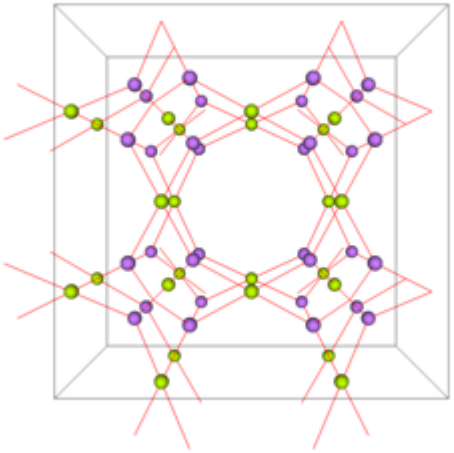


Square

Alternative, or functionalized linkers can be easily accommodated



Final model exhibits specified topology

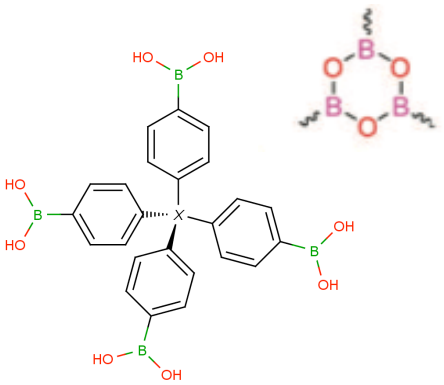


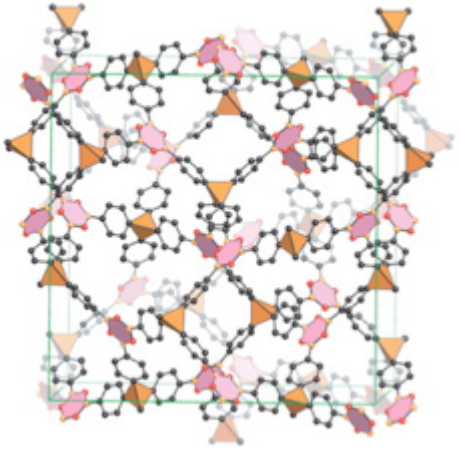
Martin, R.L.; Haranczyk, M., *Cryst. Growth Des.* 2014

Structure Prediction

Example: Quality of predicted structure

COF-102

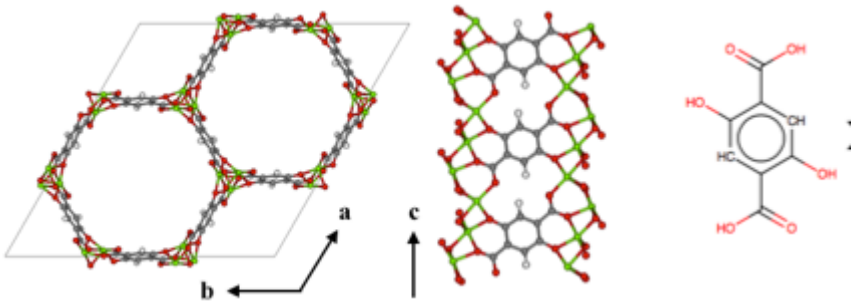




	Experiment	Model (unrelaxed)	PM6-relaxed model
Density (g/cm^3)	0.42	0.40 (-5.44%)	0.43 (2.40%)
Included sphere diameter (Å)	9.04	9.56 (5.79%)	8.52 (-5.73%)
Free sphere diameter (Å)	7.99	7.67 (-4.06%)	7.84 (-1.93%)
Accessible surface area (m^2/g)	5189.68	5730.52 (10.42%)	5425.27 (4.54%)
Accessible volume (cm^3/g)	1.00	1.07 (6.75%)	0.93 (-6.67%)

Tailored Structure Prediction Effort

MOF-74 and its analogues

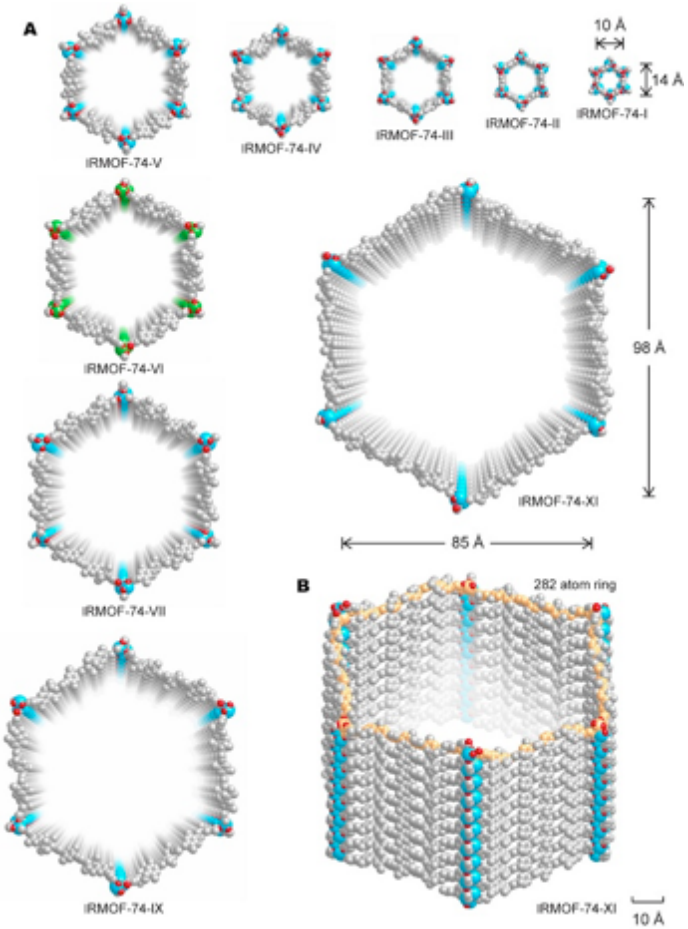


Examples of various metals

M_2(dobpdc)
M = Mg, Mn, Fe, Co, Ni, Zn

Gygi et al Chem. Mater. 2016

Examples of various linkers:

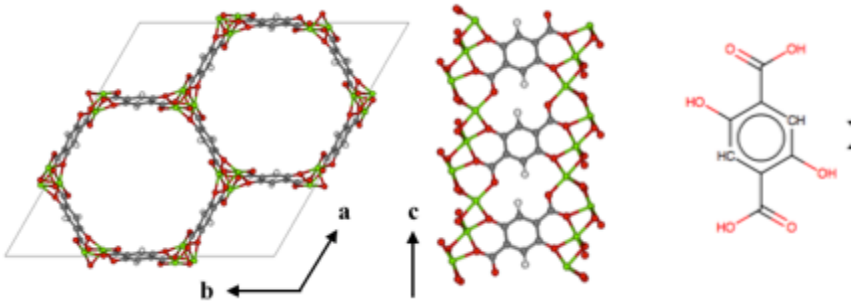


Deng et al. Science 2012

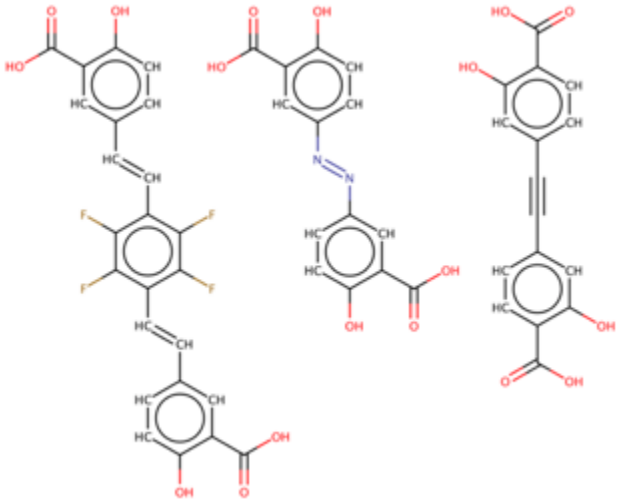
Maciej Haranczyk (maciej.haranczyk@imdea.org)

Structure Prediction of New Materials

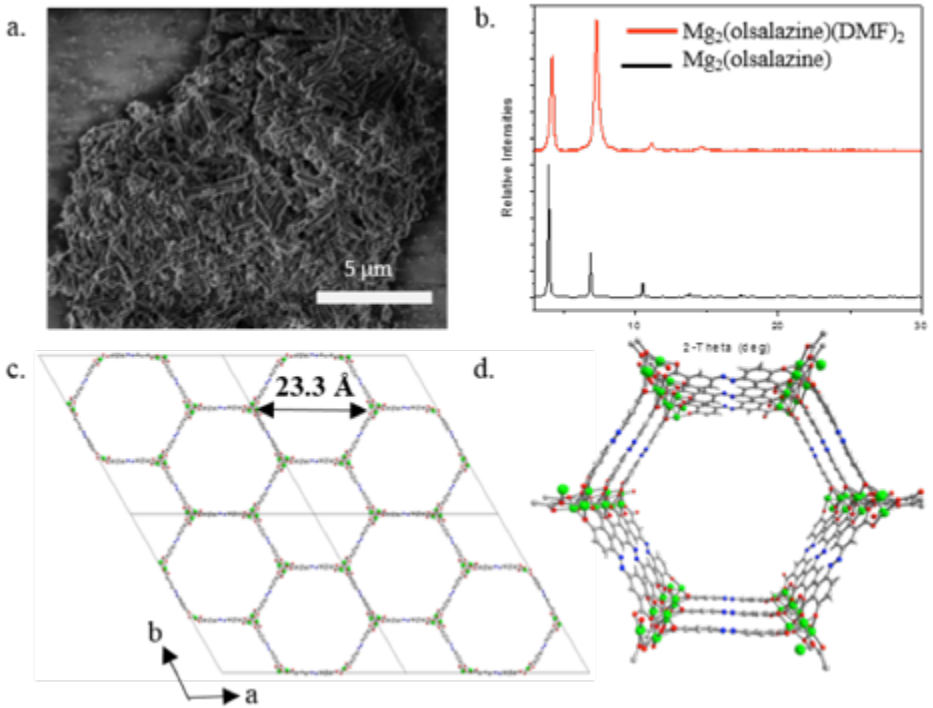
Ab initio structure prediction of MOF-74 analogues



Examples of compatible linkers



MOF-74 analogue made:
Mg_2(olsalazine)



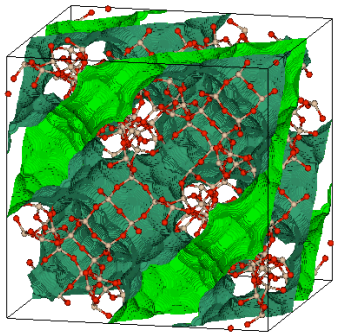
M. Witman *et al Chemical Science 2016*

Maciej Haranczyk (maciej.haranczyk@imdea.org)

High-throughput Structure Analysis and Structural Descriptors for Porous Materials

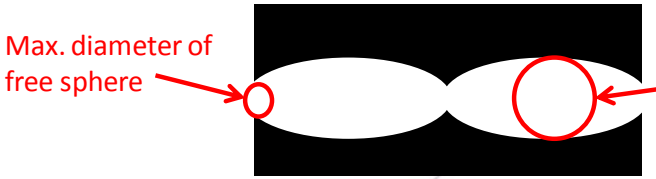
High-Throughput Analysis of Porous Materials and Their Voids

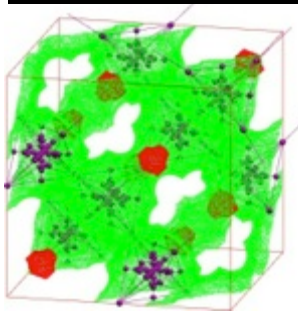
Material
= positions of atoms + unit cell



?

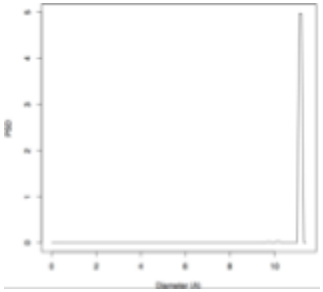
Void space geometry and topology

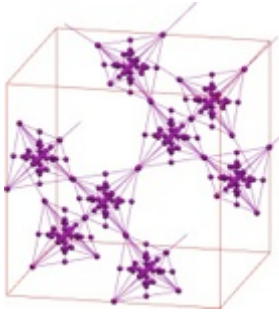




Monte Carlo sampling of accessible surface area and volume

Pore size distribution

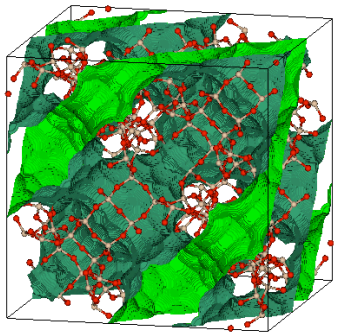


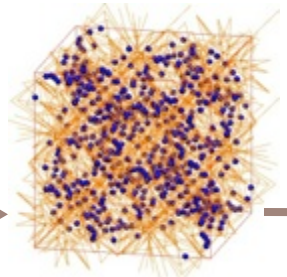


Network representing accessible void space is used to compare materials

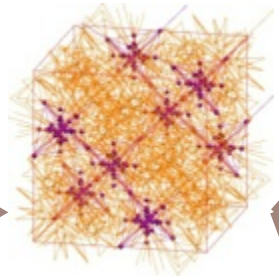
High-Throughput Analysis of Porous Materials and Their Voids

Material = positions of atoms + unit cell



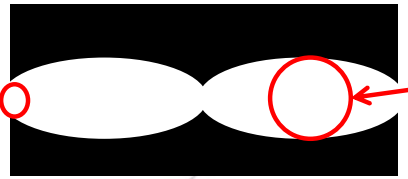


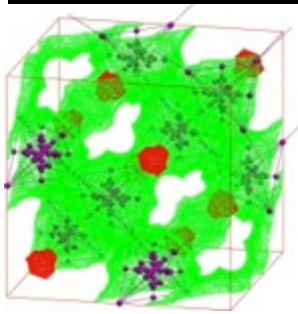
Voronoi network (a periodic graph)



Probe-accessible part of the network in pink

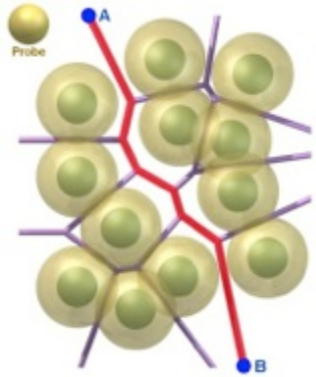
Void space geometry and topology



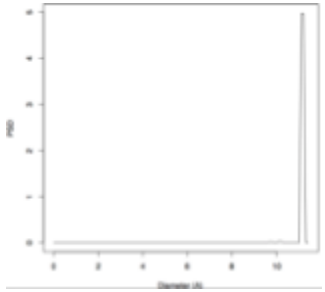


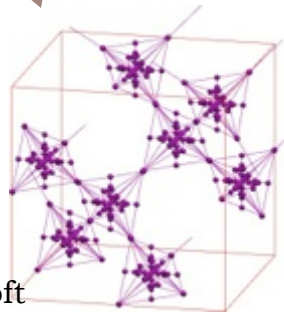
Monte Carlo sampling of accessible surface area and volume

Voronoi decomposition in 2D



Pore size distribution



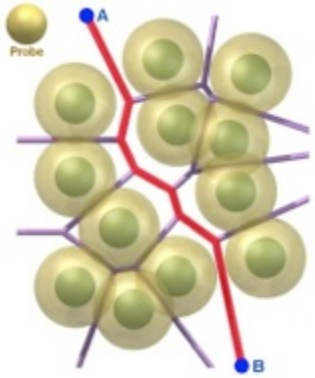


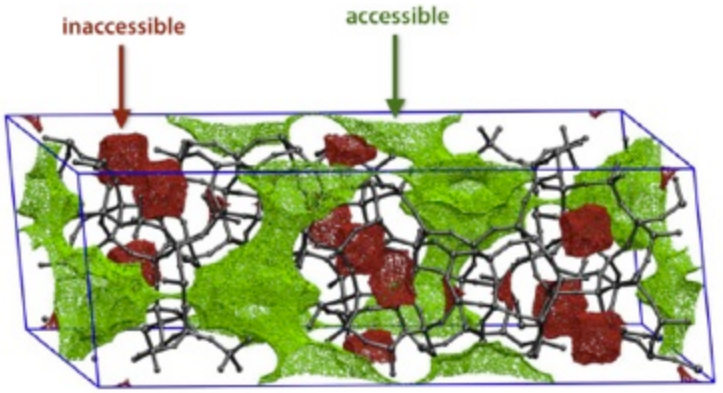
Network representing accessible void space is used to compare materials

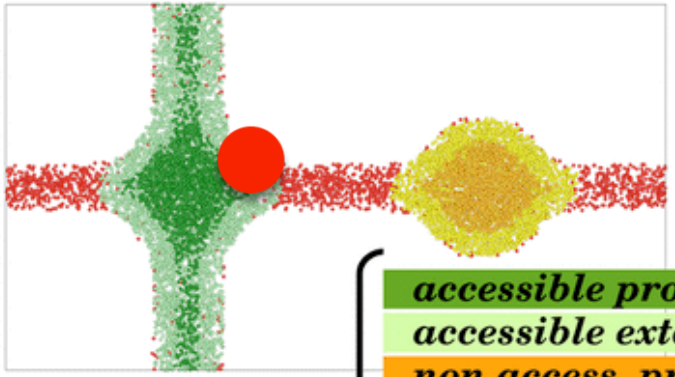
Willems et al. *Mesoporous&Microporous Materials*, 2012
Implemented in Zeo++: www.zeoplusplus.org; employs Voro++ by C. Rycroft

High-Throughput Analysis of Porous Materials and Their Voids

Voronoi decomposition in 2D

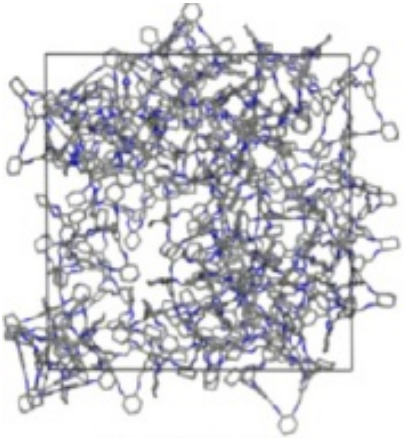




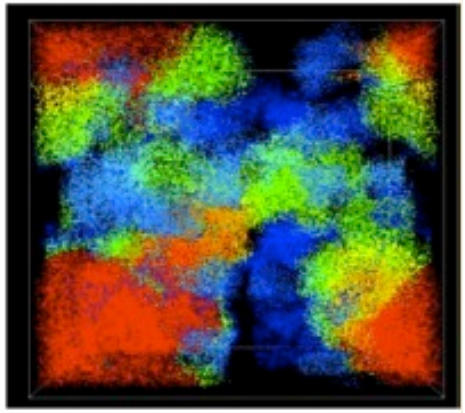


pore volume characterization

- *accessible probe center*
- *accessible extended*
- *non access. probe center*
- *non access. extended*
- *narrow*

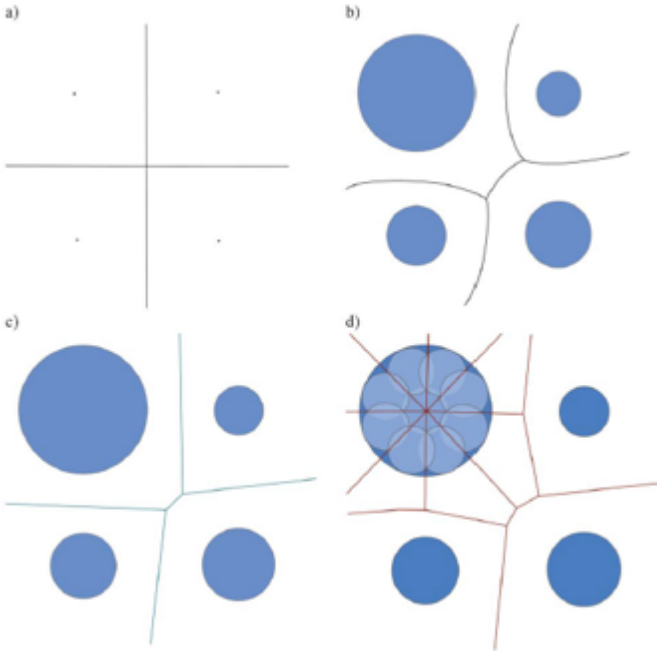


Amorphous **CC3**



Ongari et al. Langmuir, 2017; Pinheiro et al J. Mol. Graph. Model 2013
Implemented in Zeo++: www.zeoplusplus.org

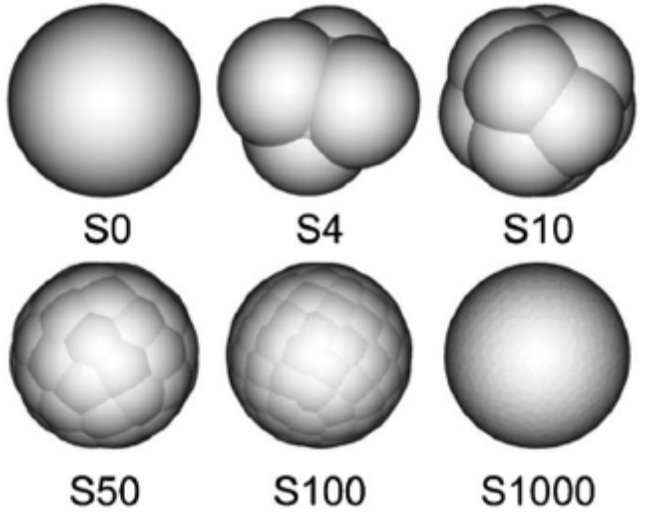
High-Throughput Analysis of Porous Materials and Their Voids

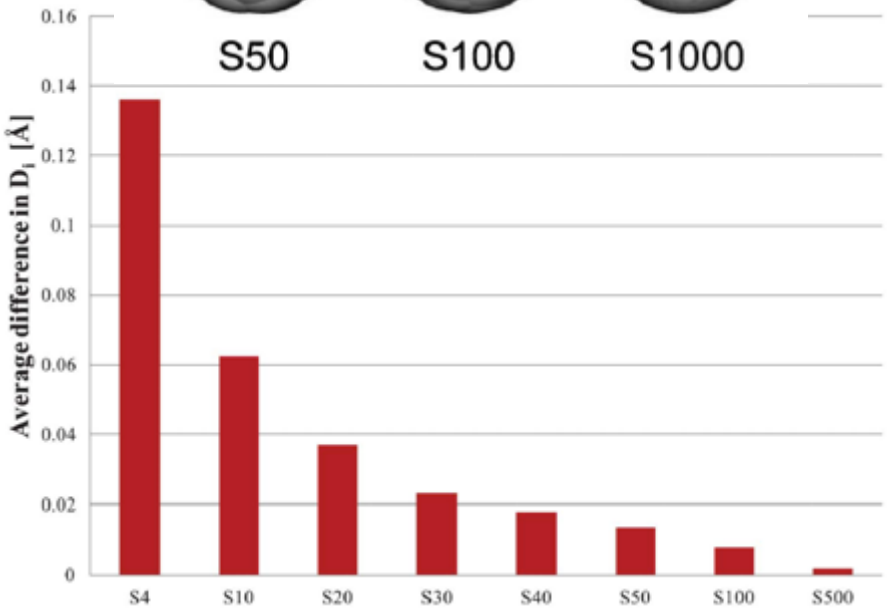


$$d(\mathbf{x},\mathbf{x}_i) < d(\mathbf{x},\mathbf{x}_j)$$

$$d(\mathbf{x},\mathbf{x}_i) - r_i < d(\mathbf{x},\mathbf{x}_j) - r_j$$

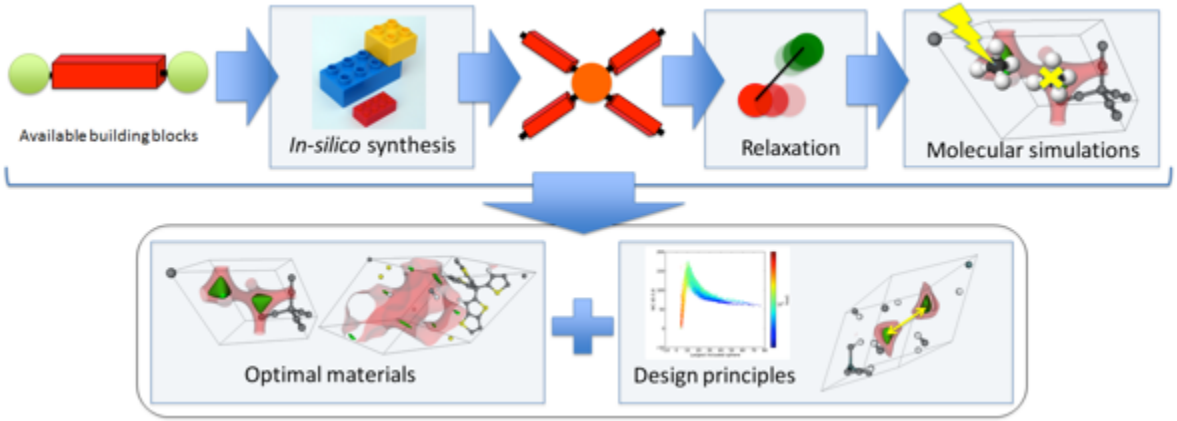
$$d(\mathbf{x},\mathbf{x}_i)^2 - r_i^2 < d(\mathbf{x},\mathbf{x}_j)^2 - r_j^2$$





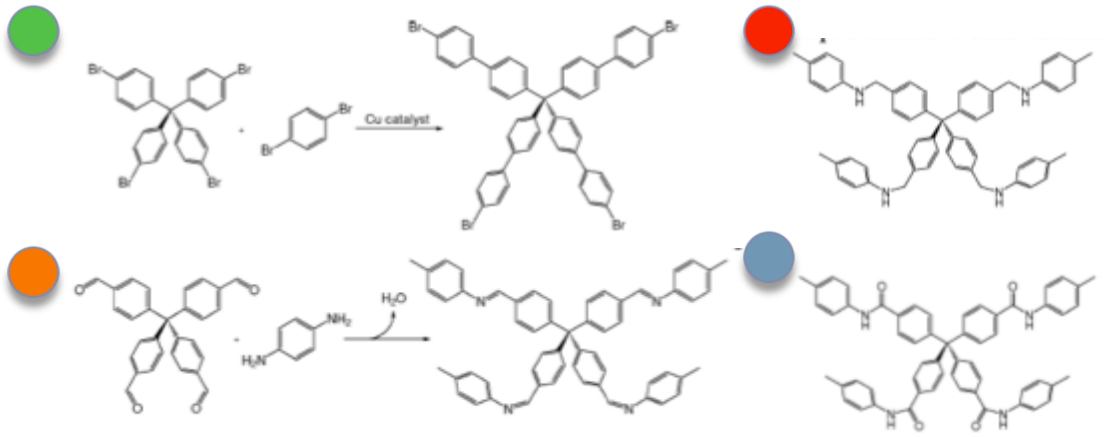
Pinheiro et al. CrysEngComm, 2013
Implemented in Zeo++: www.zeoplusplus.org

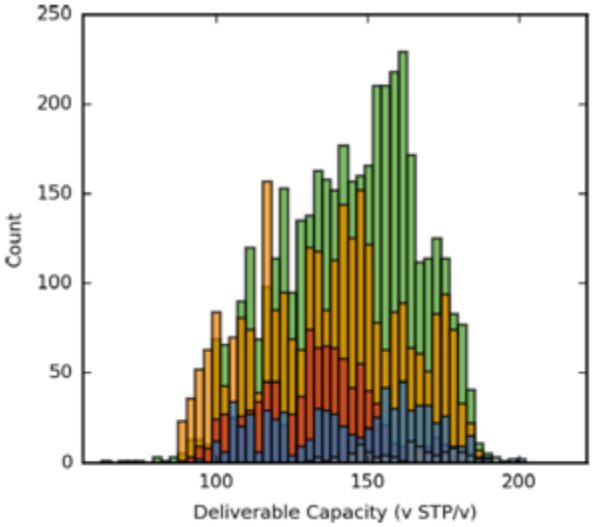
Brute-force Search for Gas Storage Materials

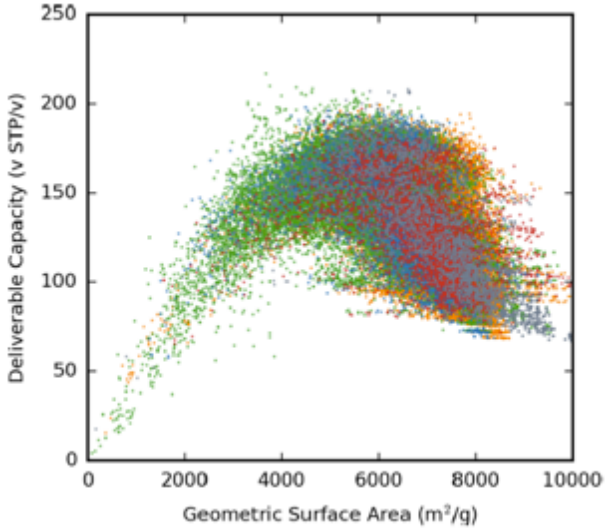


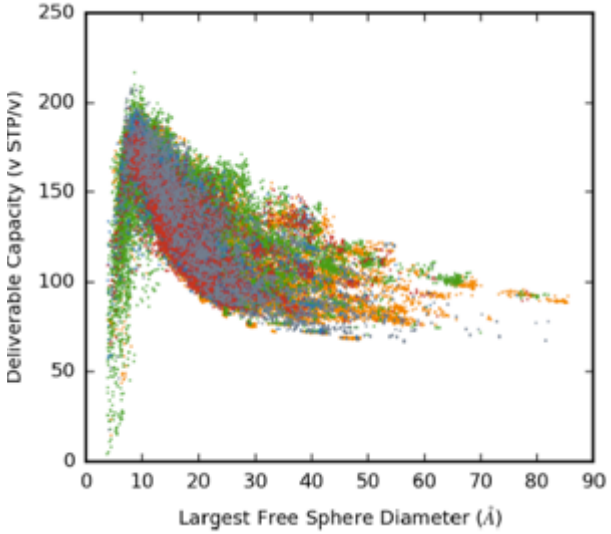
Covalent Organic Frameworks for Methane Storage

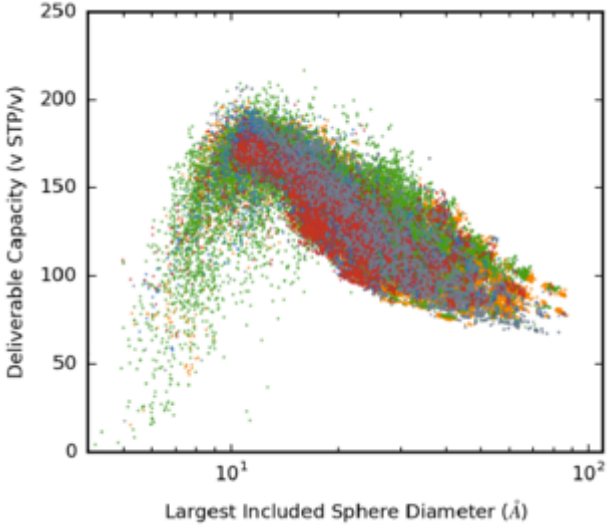
69840 COFs based on 666 linkers and 4 synthetic routes









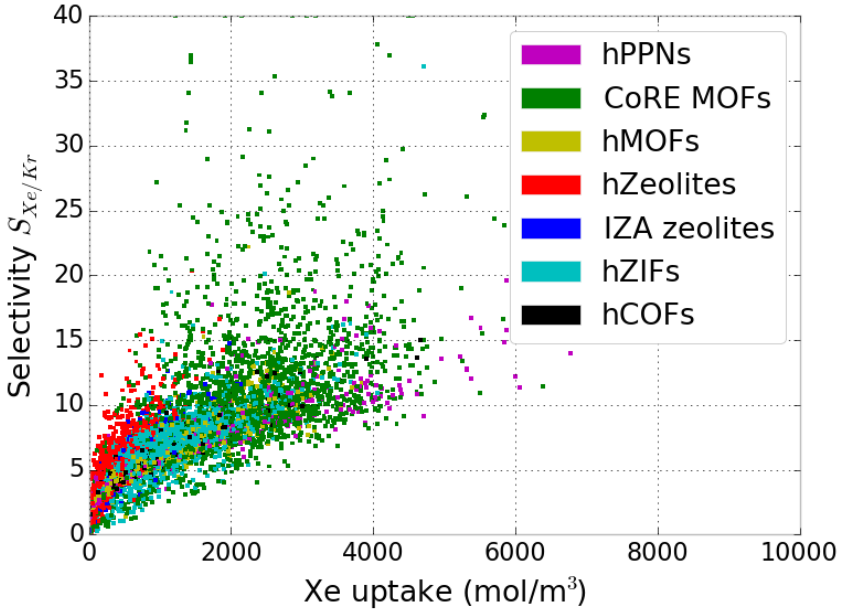


Mercado *et al. Chem. Mat.* 2018

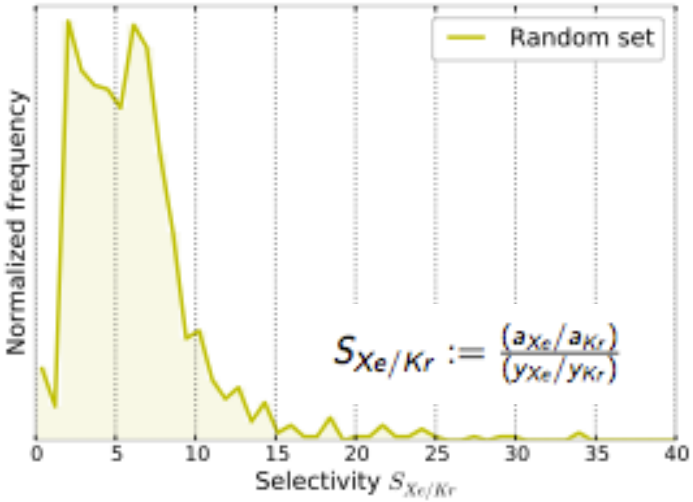
Accelerated Discovery with Machine Learning

Discovery of Materials for Xe/Kr Separations

Nanoporous Materials Genome = ca. 650k materials
Our diverse sample = ca. 15k materials



Majority of structures have poor performance…



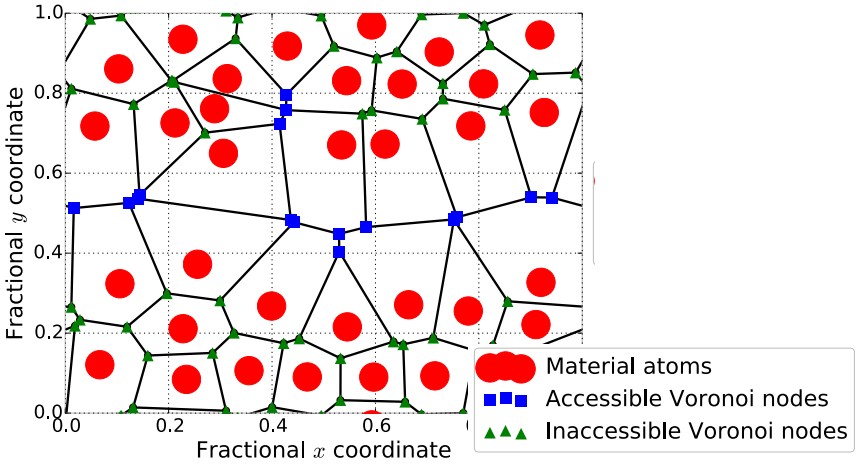
BUT with our discovery approaches we can fish them out.

C. Simon, R. Mercado, S. Schnell, B. Smit, M. Haranczyk. What Are the Best Materials To Separate a Xenon/Krypton Mixture? Chemistry of Materials, 2015, 27 (12), pp 4459-4475.

Discovery of Materials for Xe/Kr Separations

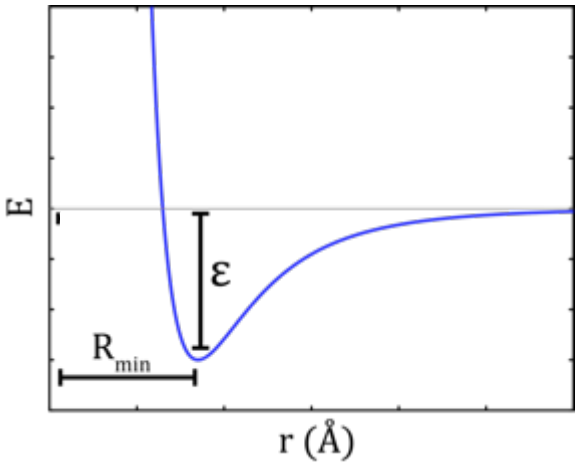
Structural descriptors (feature vector)

Descriptor	Symbol
Void fraction [unitless]	ϵ_v
Crystal density [kg/m^3]	ρ
Largest free sphere diameter [Å]	D_f
Largest included sphere diameter [Å]	D_i
Accessible surface area [m^2/cm^3]	a
Surface density [kg/m^2]	ρ_s



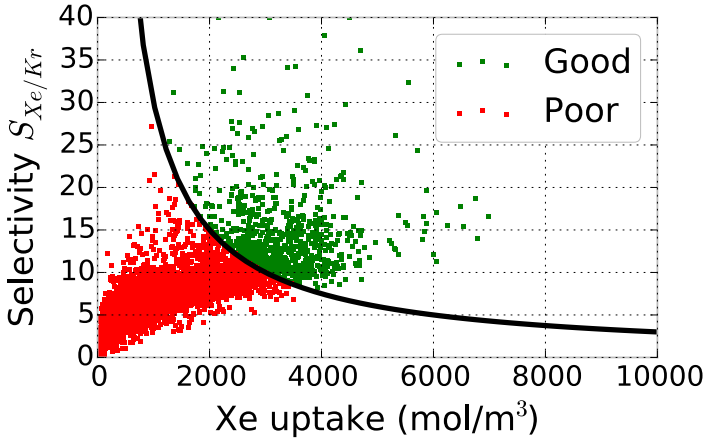
Lennard-Jones potential

$$\mathcal{V}(r) = \varepsilon\left[\left(\frac{R_{min}}{r}\right)^{12} - 2\left(\frac{R_{min}}{r}\right)^{6}\right]$$

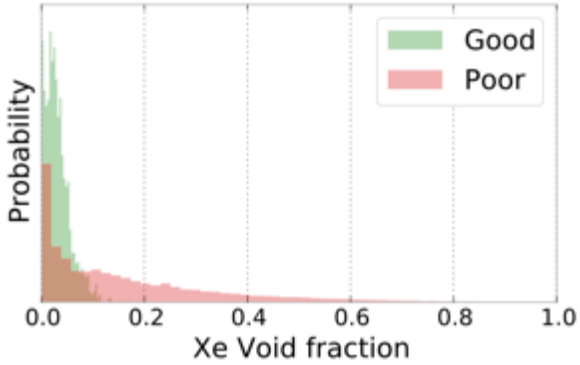


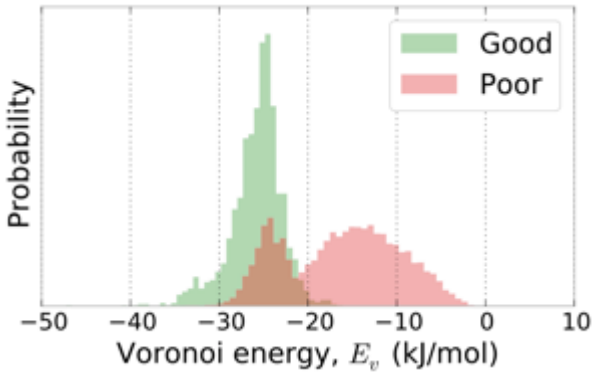
Discovery of Materials for Xe/Kr Separations (II)

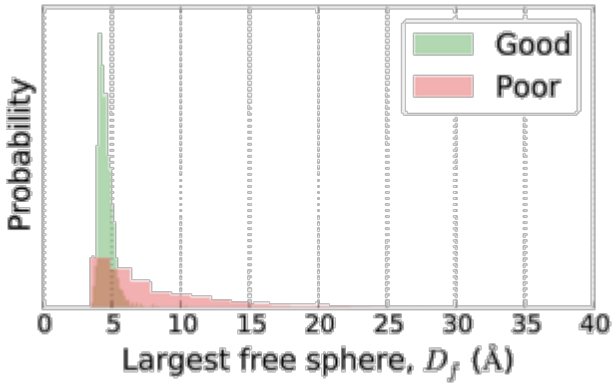
Sample of diverse materials become the training set



No single feature (descriptor) determines properties of interest

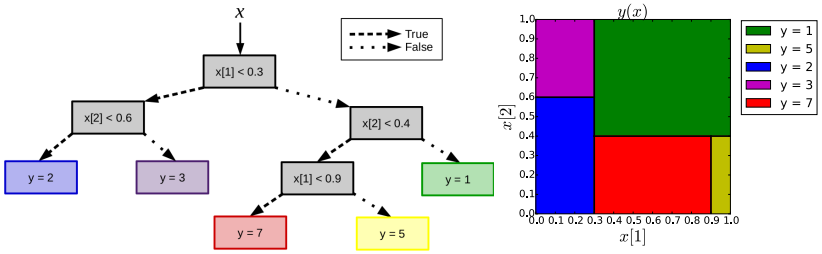


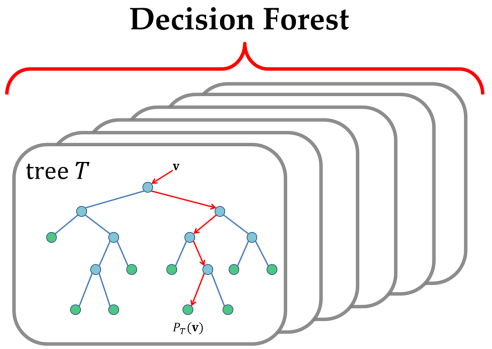




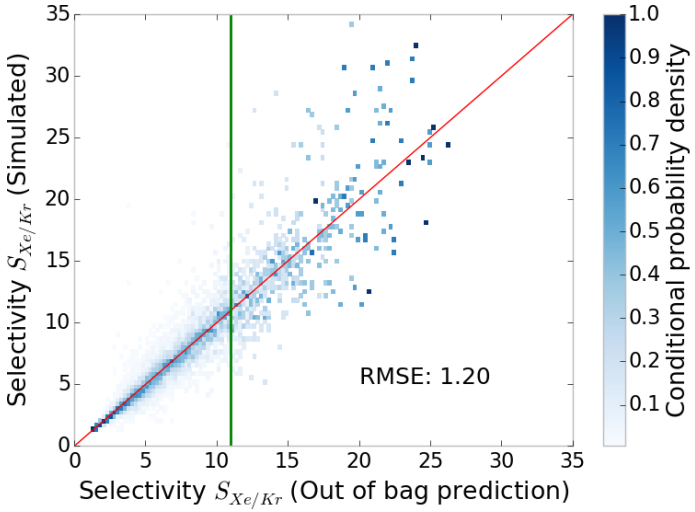
Discovery of Materials for Xe/Kr Separations (III)

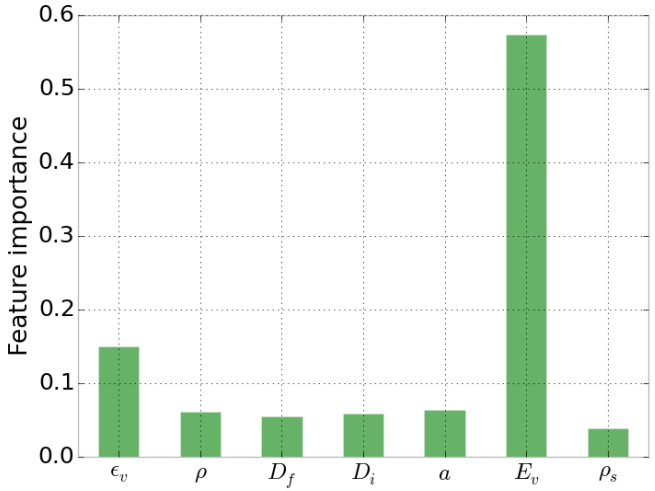
Statistical machine learning is used to correlated feature vectors with performance data.
Here, we use a decision forest





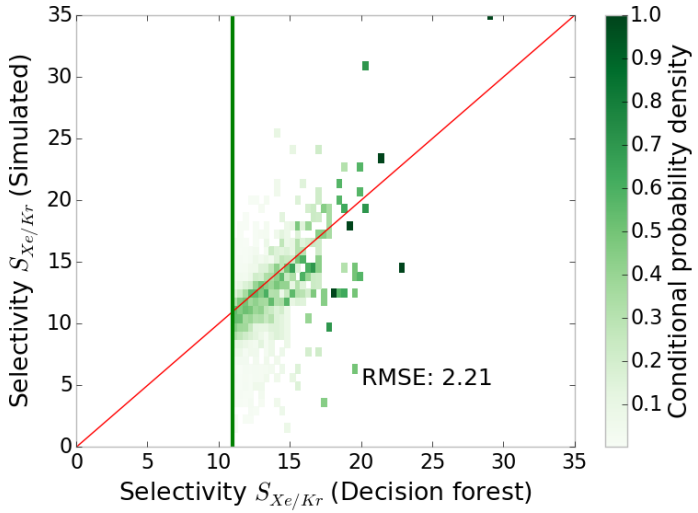
Validation of training

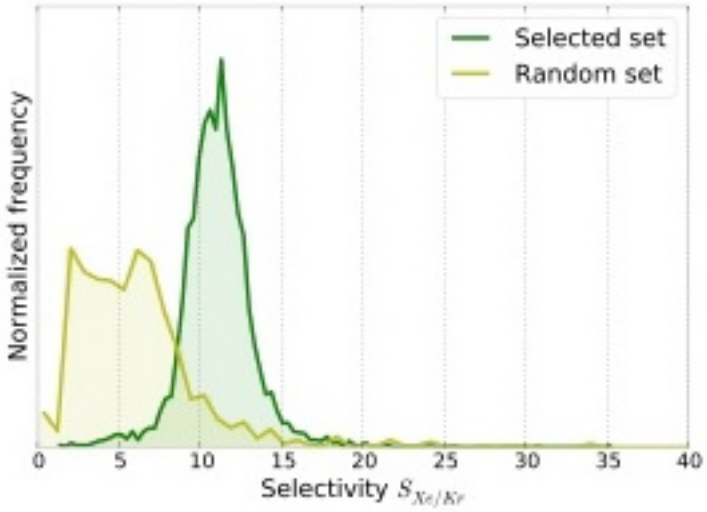




Discovery of Materials for Xe/Kr Separations (IV)

Performance prediction for all structures in the Nanoporous Materials Genome

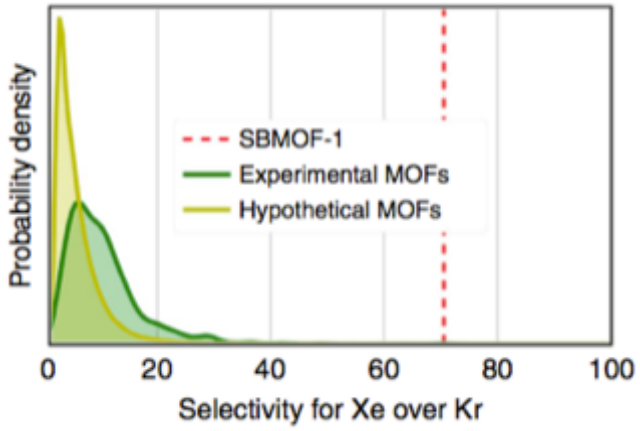


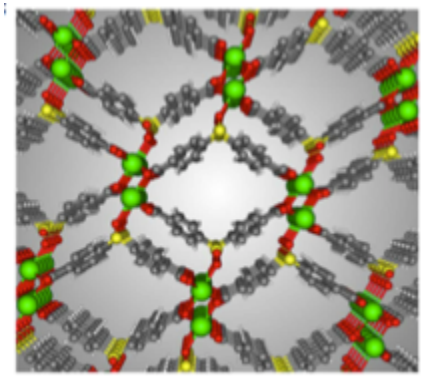


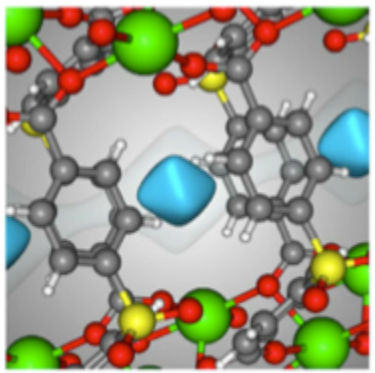
C. Simon, R. Mercado, S. Schnell, B. Smit, M. Haranczyk. What Are the Best Materials To Separate a Xenon/ Krypton Mixture? Chemistry of Materials, 2015, 27 (12), pp 4459-4475.

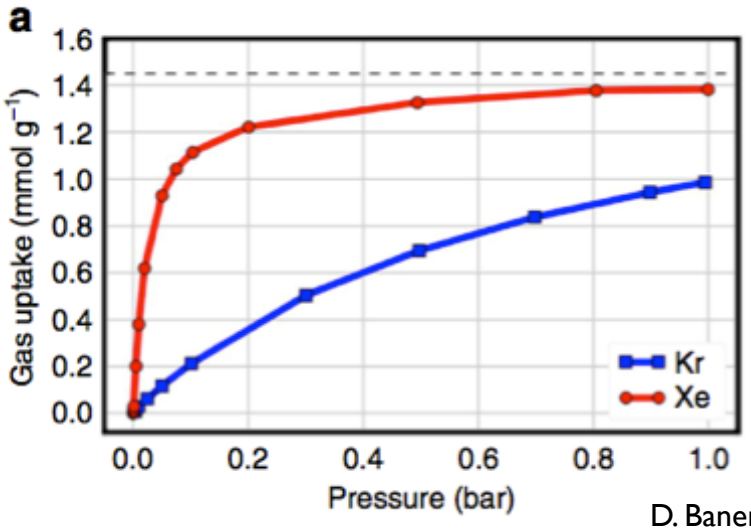
Maciej Haranczyk (mharanczyk@lbl.gov)
Computational Research Division

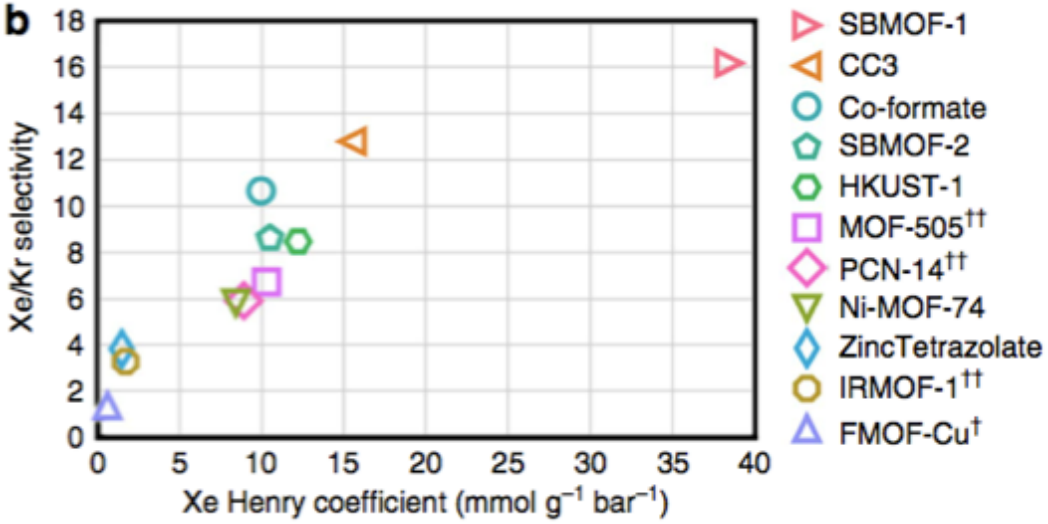
SBMOF-1 Discovery and Experimental Verification







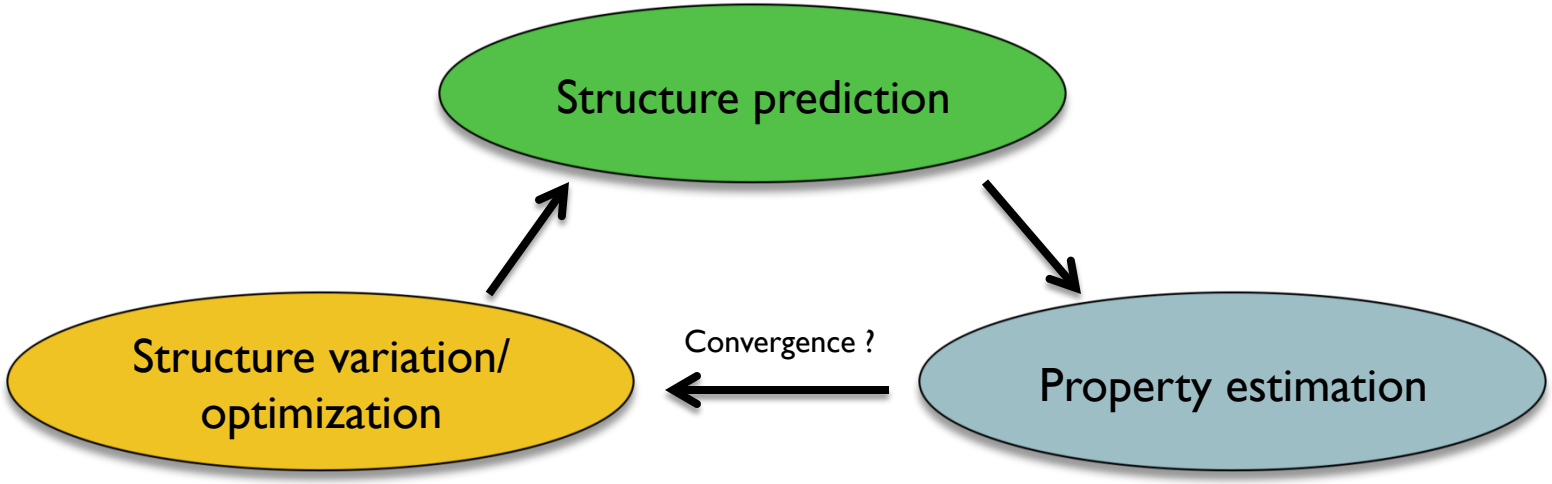




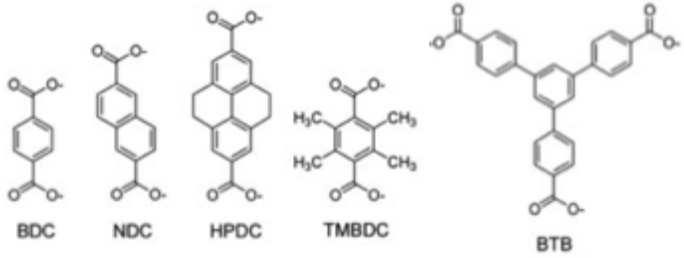
D. Banerjee, C.M. Simon, A.M. Plonka, R.K. Motkuri, J. Liu, X. Chen, B. Smit, J.B. Parise, M. Haranczyk, P.K. Thallapally - Metal-Organic Framework with Optimal Adsorption, Separation, and Selectivity towards Xenon – *Nature Communications 7 (2016) 11831*

Optimization-based Materials Design

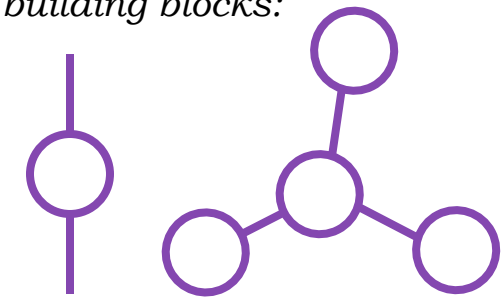
Optimization-based Material Design



Possible discrete chemical building blocks:

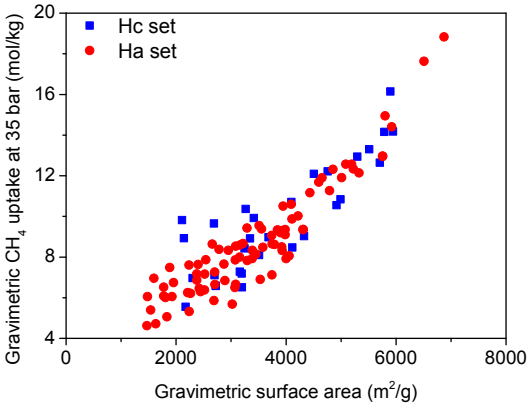


Possible "alchemical" building blocks:

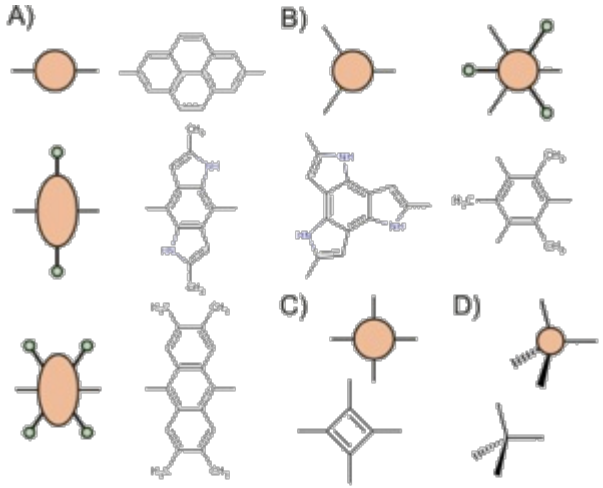


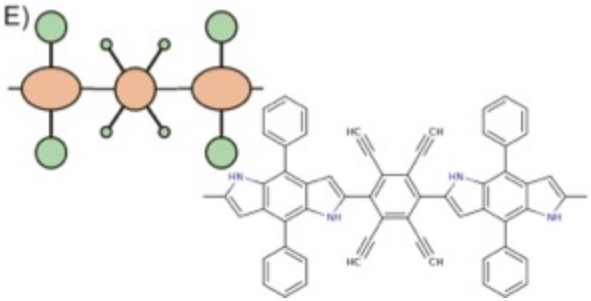
Exploring Frontiers of High Surface Area Materials: Gravimetric Surface Area (GSA)

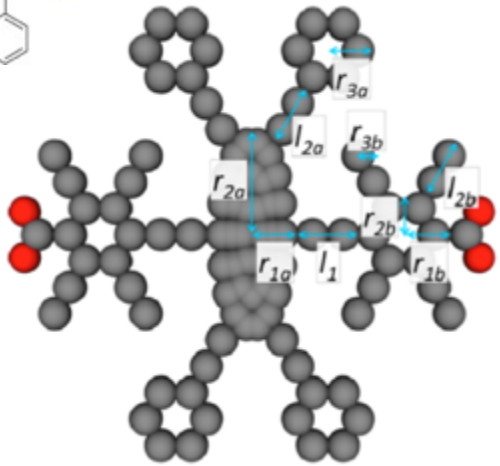
Motivation: GSA has been the most addressed property in the design of MOFs.



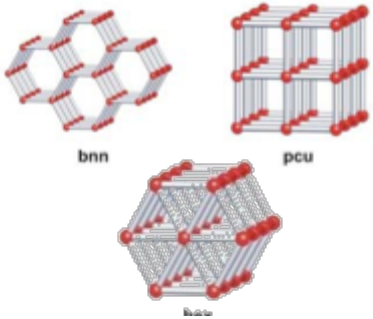
Example fragments

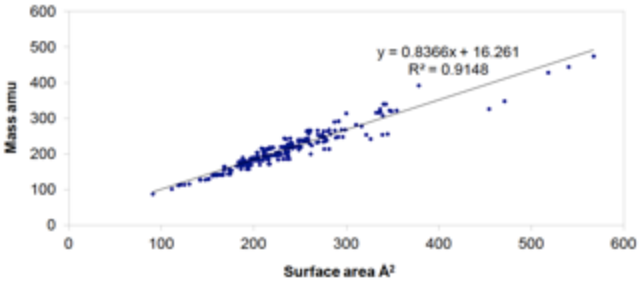




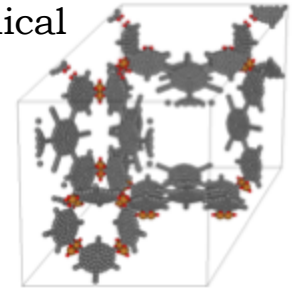


Possible topologies:





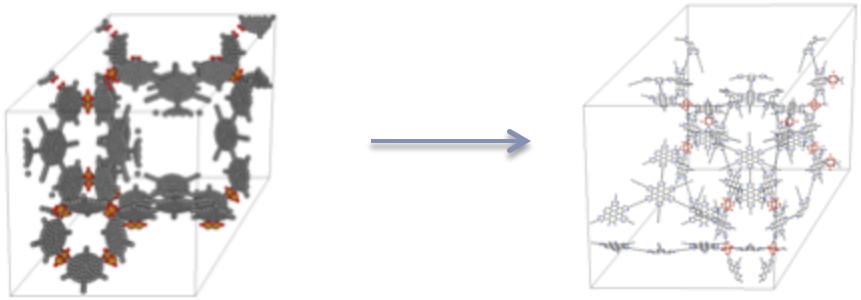
Alchemical MOF



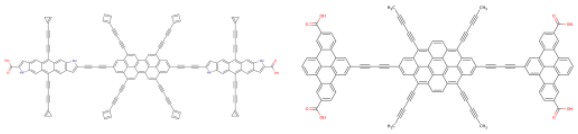
Exploring Frontiers of High Surface Area Materials: Gravimetric Surface Area (GSA)

Optimized Alchemical MOF

Real MOF

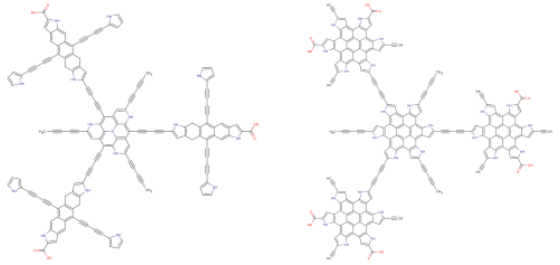


Optimal building blocks for each considered topology:



Best *pcu*: 9968.35 m^2g^{-1}

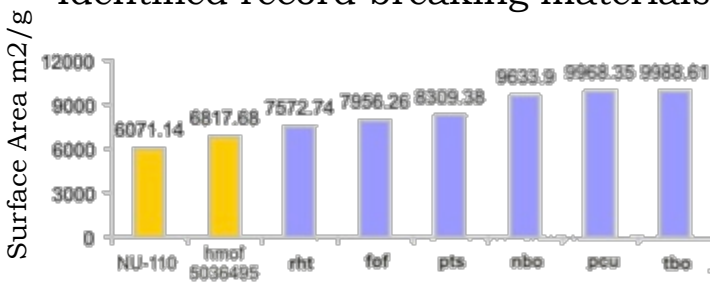
Best *nbo* (2-3-coordinated): 7956.26 m^2g^{-1}



Best *tbo*: 9988.61 m^2g^{-1}

Best *rht*: 7572.74 m^2g^{-1}

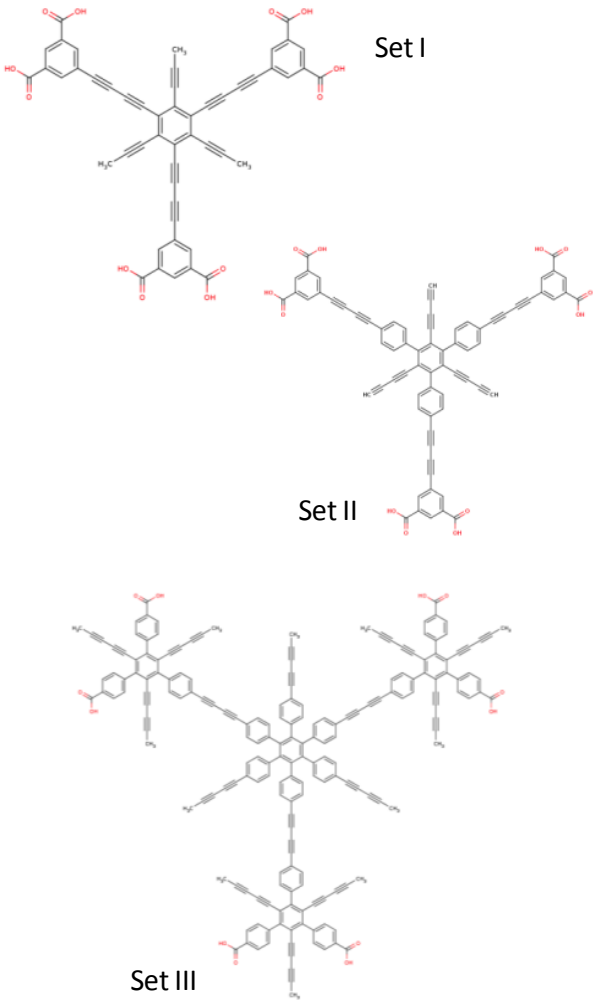
Identified record breaking materials



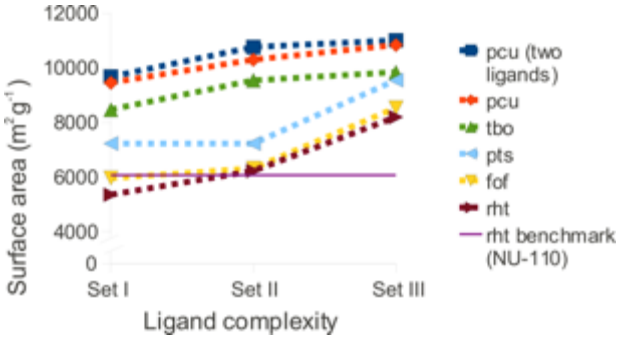
Martin, Haranczyk, Chemical Science, 2013

Exploring Frontiers of High Surface Area MOFs

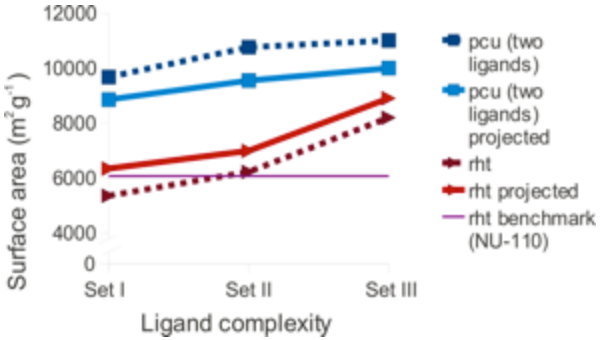
Chemical complexity – I-III



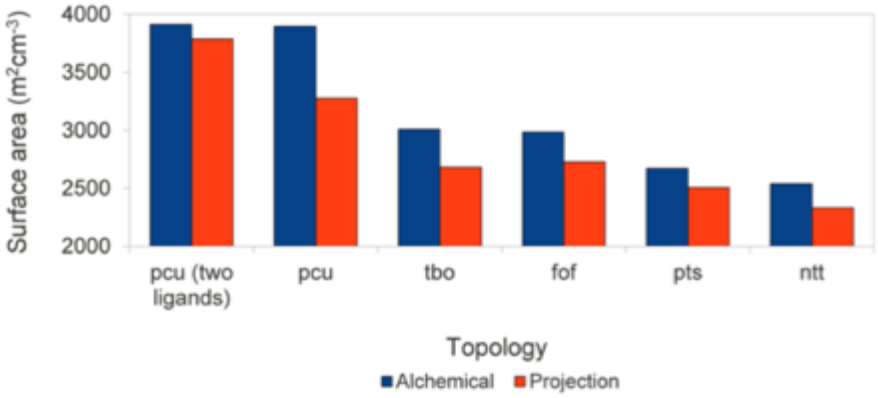
Alchemical linkers



Chemical linkers



Similar study for volumetric surface area (VSA)



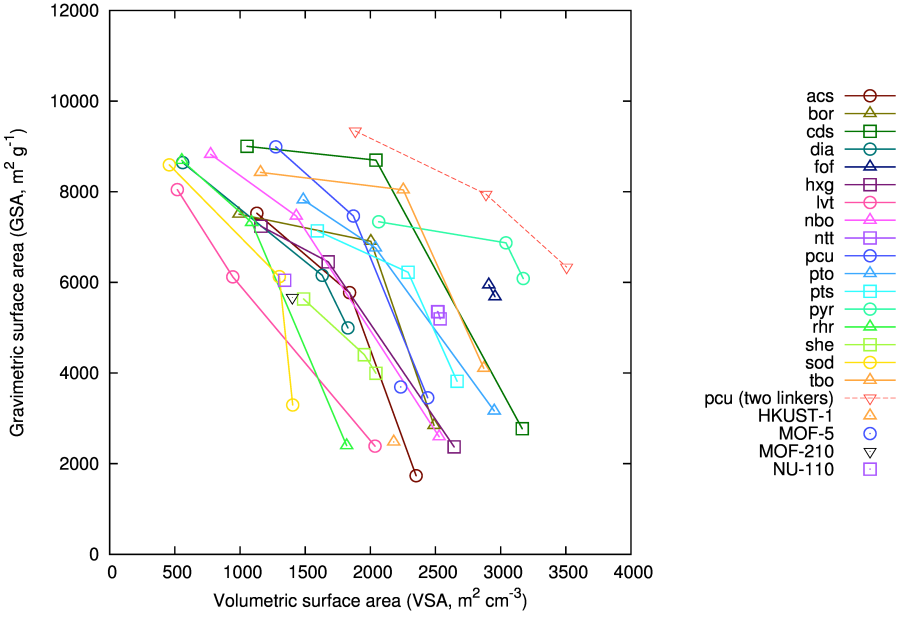
Martin, Haranczyk, J. Chem. Theory Comp. (JCTC) 2013

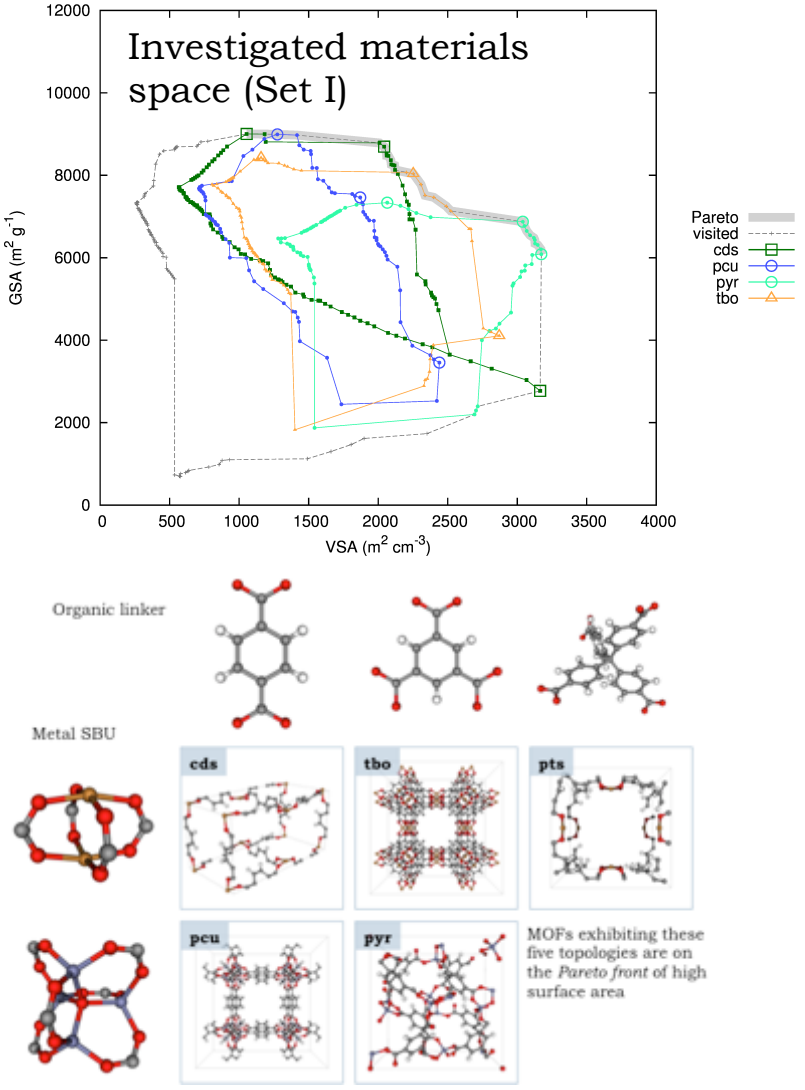
Towards Multiobjective Design

Multiobjective design can be achieved by combining objectives:

F = GSA * VSA

Comparison of topologies (Set I):





Martin, Haranczyk, Crystal Growth Design, 2013

Optimization in Real Chemical Space

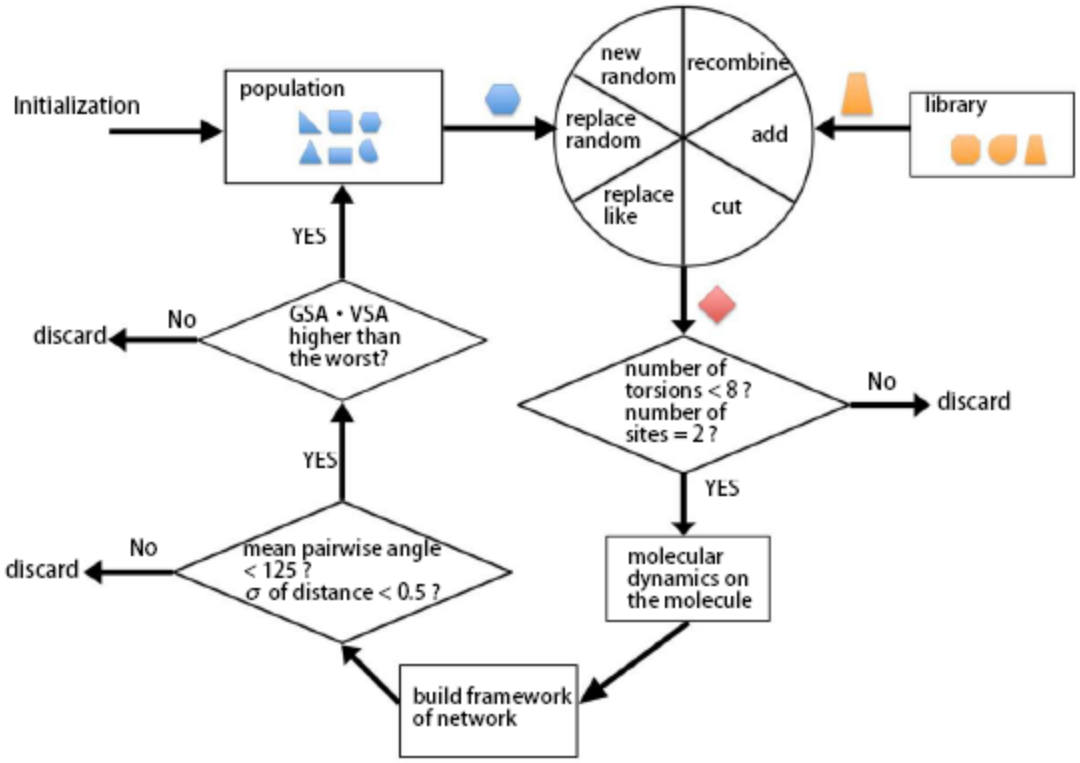
An alternative implementation of optimization-based design approach involves real molecular fragments and genetic algorithms .

Our chemical library consists of *84* reactions and over *40,000* compounds on which to perform them
(Aldrich catalogue(A); eMolecules.com(B))

At each step, *"mutations"* are performed on candidate molecules

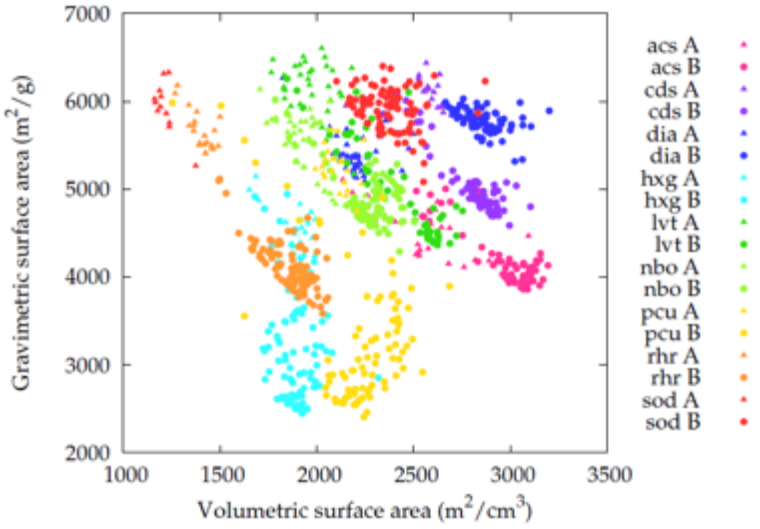
The *"fittest"* (e.g., highest resulting MOF surface area) progress and are the basis for subsequent molecules

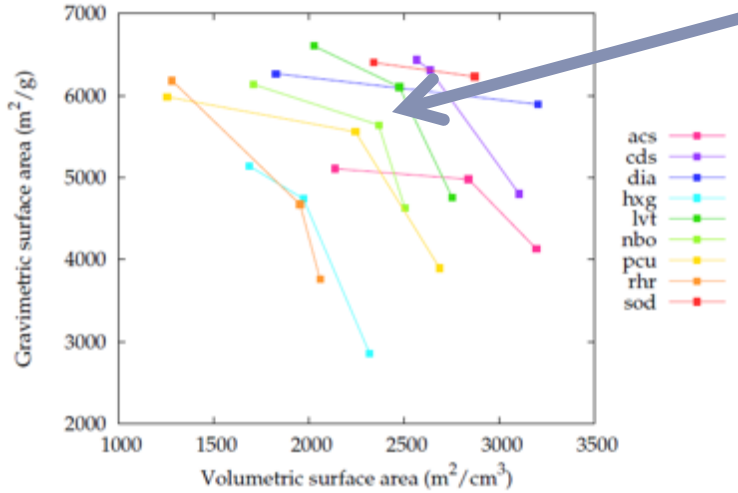
Genetic algorithm (GA) workflow



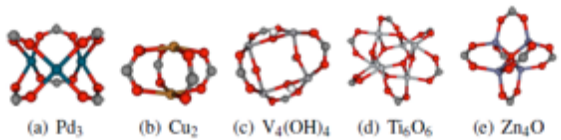
Bao et al. PCCP 2015; JPCC 2015

Optimization in Real Chemical Space

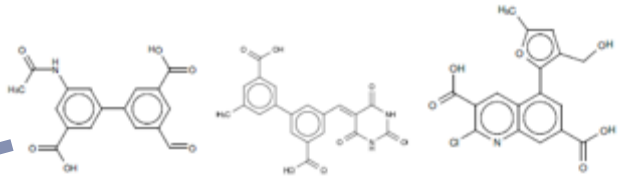




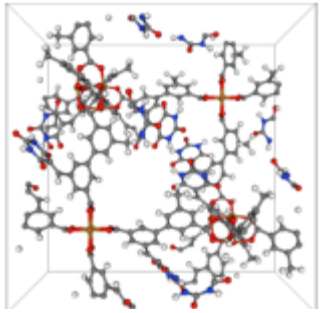
Net	SBU	Description
acs	Pd_3	trigonal prism
cds, lvt, nbo, rhr	Cu_2	square "paddlewheel"
dia, sod	$V_4(OH)_4$	tetrahedral
hxg	Ti_6O_6	hexagonal
pcu	Zn_4O	octahedral

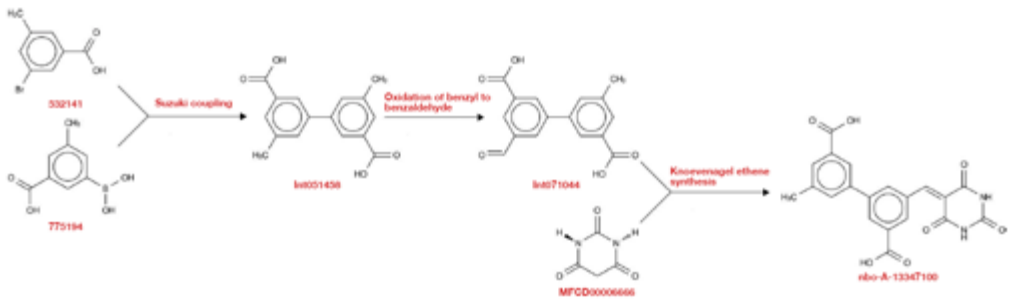


(a) Pd_3 (b) Cu_2 (c) $V_4(OH)_4$ (d) Ti_6O_6 (e) Zn_4O



(a) nbo-A-12901900 (b) nbo-A-13347100 (c) nbo-B-12769400



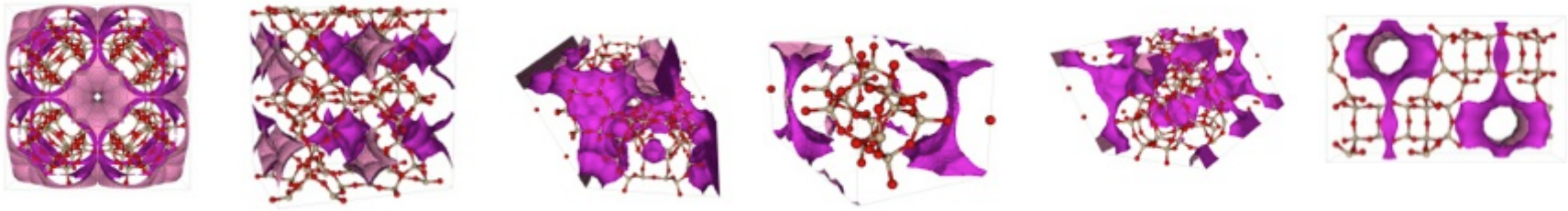


Bao et al. PCCP 2015; JPCC 2015

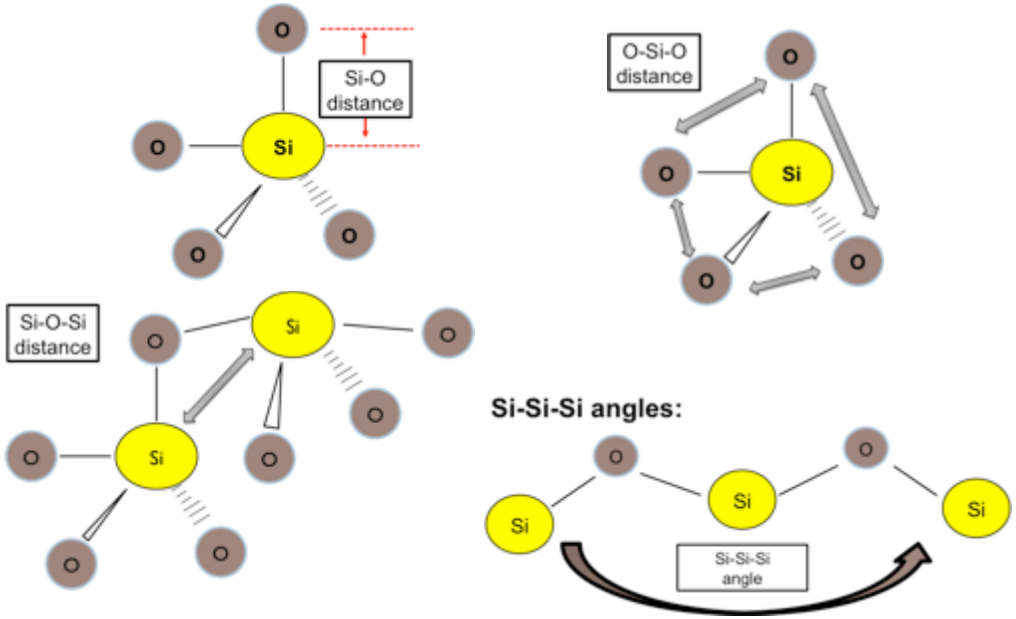
Synthesizability and Applicability Prediction with New Structure Descriptors

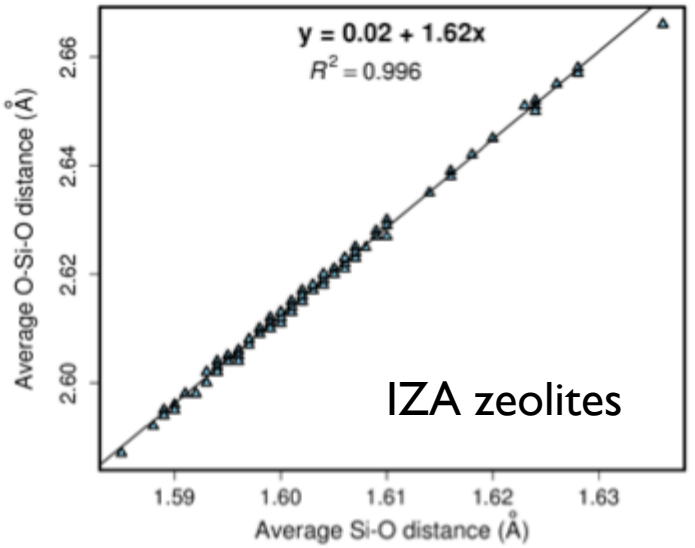
Towards Synthesizability Prediction: Zeolites

Hypothetical zeolite topologies ~3M vs 200-300 known ones



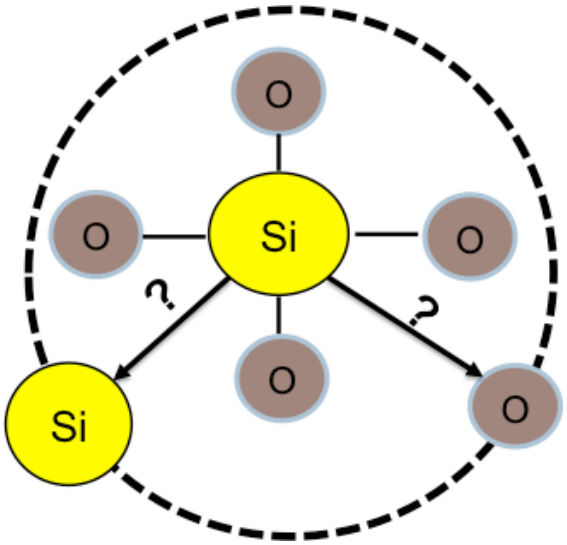
Li structural criteria based on known zeolites
(Li *et al.*, *Angew. Chem. Int. Ed.* **2013**, 10.1002/anie.201206340)





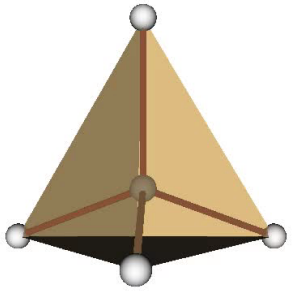
New Descriptors for Synthesizability Prediction

Distance to fifth neighbor:



Degree of tetrahedrality [0-1]:

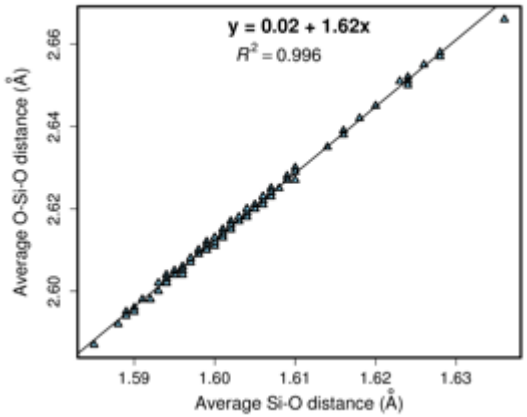
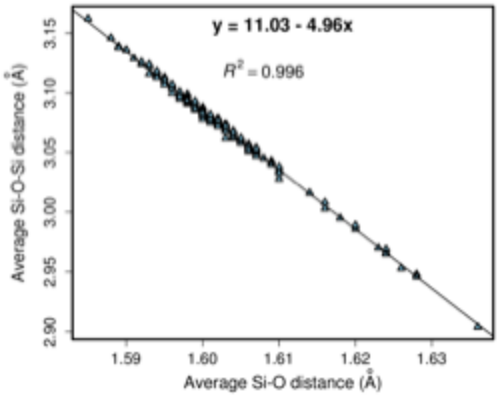
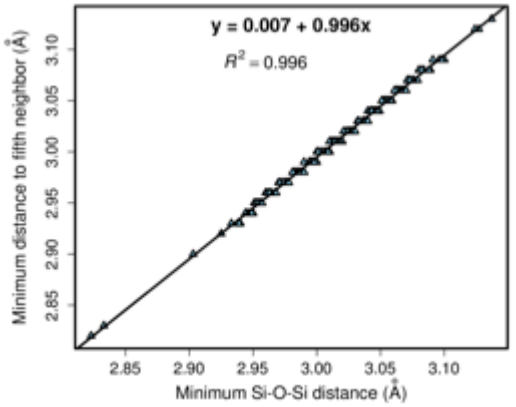
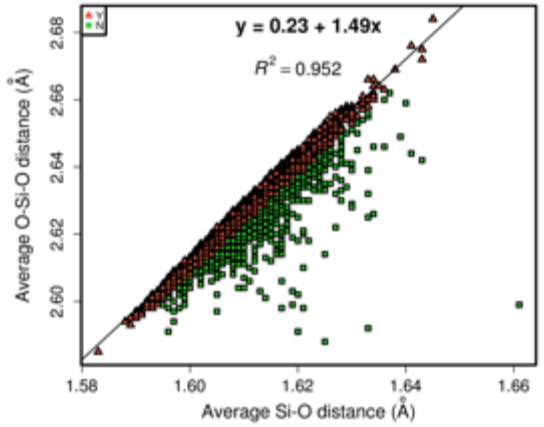
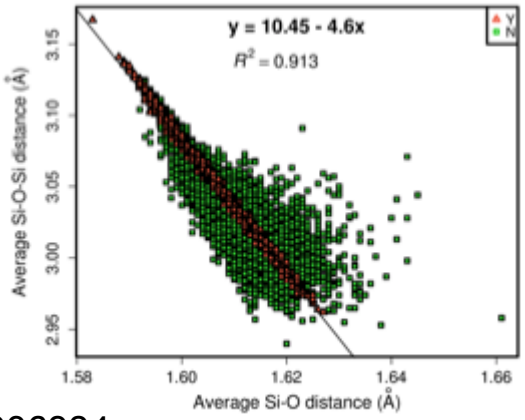
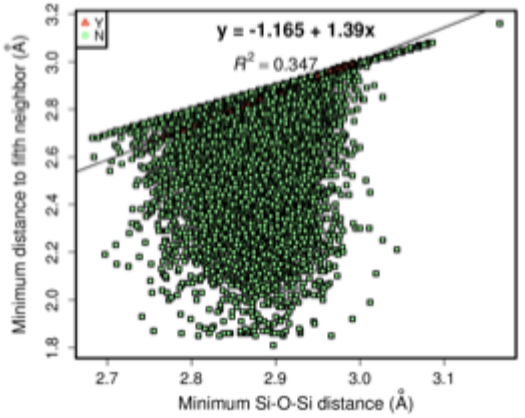
$q_{tet} = 1.00$



0 degrees

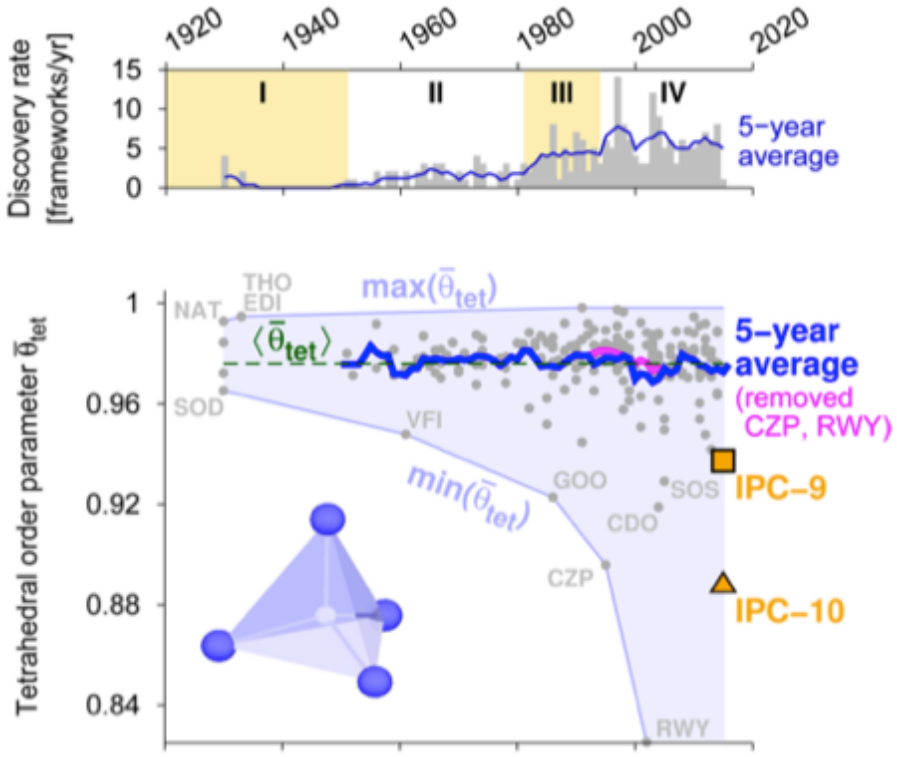
$$\theta_{tet} = \frac{1}{24}\sum_{j\neq k}^{4}\left\{\exp\left[\frac{-(\alpha_k - \alpha_{tet})^2}{2\Delta\alpha^2}\right]\sum_{m\neq j,k}^{4}\cos(3\varphi)\right.$$
$$\left.\exp\left[\frac{-(\alpha_m - \alpha_{tet})^2}{2\Delta\alpha^2}\right]\right\}$$

New Descriptors for Synthesizability Prediction and their Applications to Deems Database

	Average O-Si-O distance versus average Si-O distance	Average Si-O-Si distance versus average Si-O distance	Minimum distance to fifth neighbor versus minimum Si-O-Si distance
IZA zeolites			
Deem's zeolites			

Deem, *J. Phys. Chem. C* **2009**, 10.1021/jp906984z

Pushing It Further: Applicability Prediction



Tetrahedrality descriptor for industrially relevant zeolites

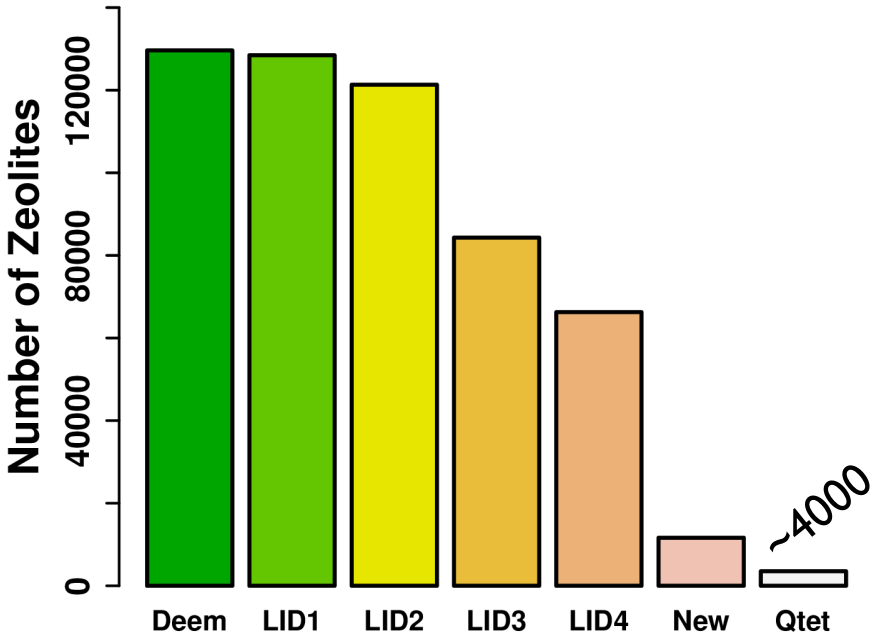
Framework type	Tetrahedrality $\bar{\theta}_{tet}$
MFI	0.979
FAU	0.979
BEA	0.976
MOR	0.971
Average	0.976

High tetrahedrality seems a necessary structural feature to sustain harsh process conditions

Zimmermann, Haranczyk, CG&D, 2016

How many all silicious zeolites are ‘out there’ ?

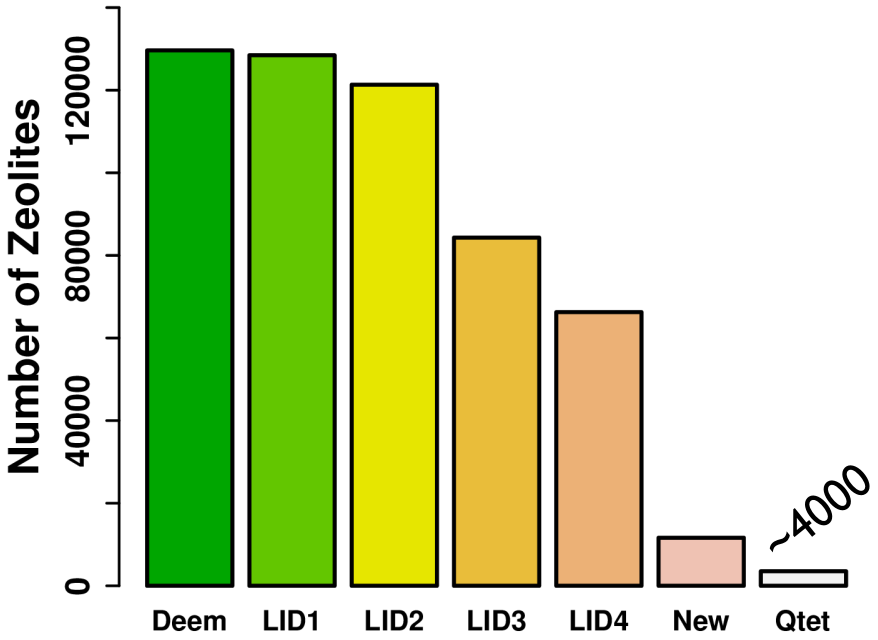
All criteria applied to Deems PCOD database



Descriptor	Percentage filtered
Deem	NA
LID1	0.90
LID2	5.55
LID3	30.51
LID4	21.41
New (5thN)	82.43
Qtet	69.72

How many all silicious zeolites are ‘out there’ ?

All criteria applied to Deems PCOD database



Descriptor	Percentage filtered
Deem	NA
LID1	0.90
LID2	5.55
LID3	30.51
LID4	21.41
New (5thN)	82.43
Qtet	69.72

~120

*Plus including filer based on diffusion criteria (2D channels, large pores)

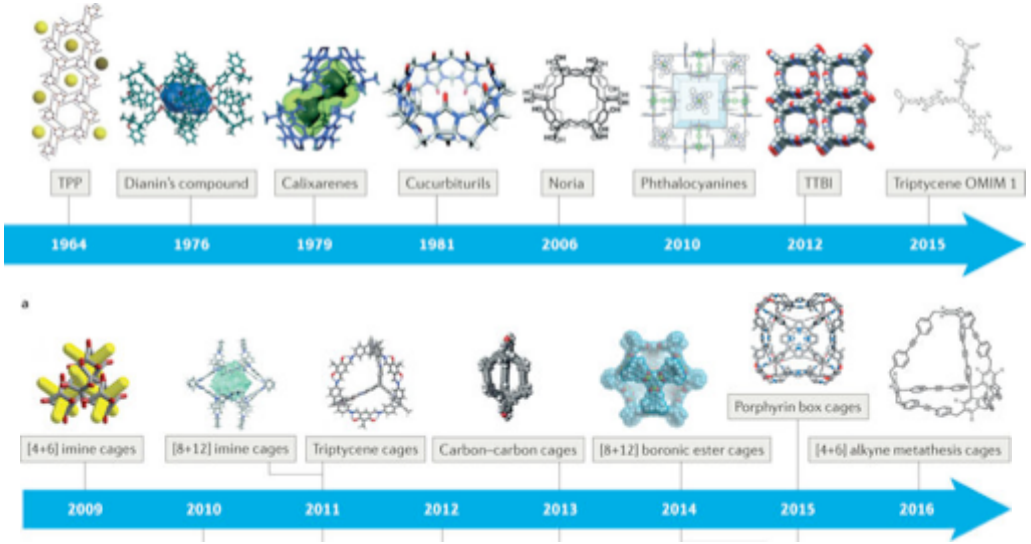
Molecular Barn-finds and their Porous Crystalline Phases

Barn-find

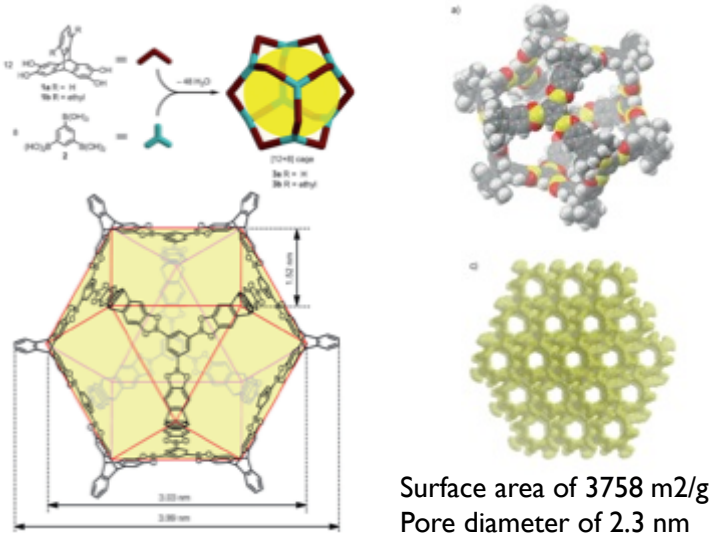
*A **barn find** is a classic car or motorcycle that has been discovered, often in derelict condition. The term comes from their tendency to be found in places such as barns, sheds, carports and outbuildings where they have been stored for many years. The term usually applies to vehicles that are rare and valuable, and which are consequently of great interest to car collectors and enthusiasts despite their poor condition.*

[Wikipedia]

Progress in Discovery of Porous Molecular Materials



Hansell&Cooper,; Nature Rev Mat 2016



Surface area of 3758 m2/g
Pore diameter of 2.3 nm

Zhang et al, Angew. Chem., Int. Ed. 2014

Databases of crystal structures:

- Cambridge Structure Database (CSD)
- Crystallography Open Database

CSD:

- Largest [650k]
- Structures provided by users
- Structures need to fulfill high quality criteria

Porous Molecular Materials from Data Science Perspective

Are there any previously synthesized cage molecules yet to enter the porous materials research?

"Molecular barn finds"

The introduction of a porous molecule to the porous solids community requires three steps: (i) the molecule has to be synthesized, (ii) the corresponding solid material (amorphous or crystalline) needs to be obtained, and (iii) its porosity needs to be confirmed.

Repository of All Known Molecules

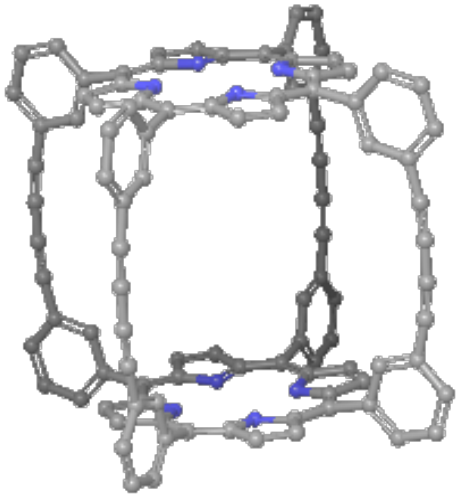
Databases of molecules:

- PubChem
- ChemSpider
- Vendors

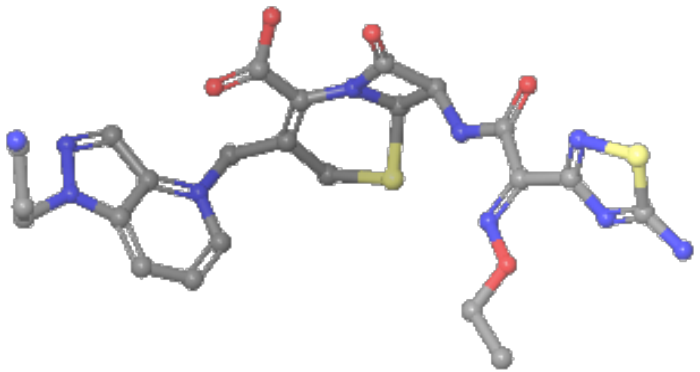
PubChem

- Largest of them [100M]
- **Structures mined from literature, patents, provided by vendors**
- All structures encoded as SMILES, some have 3D coordinates (generated with OpenEye's OMEGA and MMFF94)

Porous ?



CID = 101377082



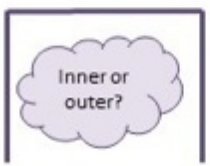
CID = 10008508

Examples from PubChem

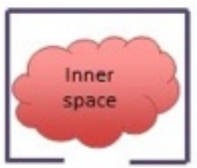
Detection of Molecular Porosity

Linear shape

Open shape



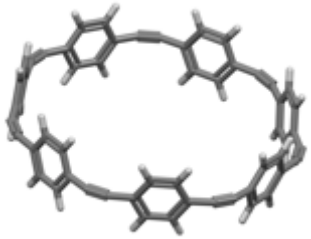
Partially trapped



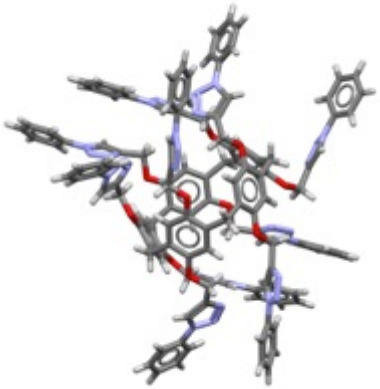
Almost fully trapped

Fully trapped

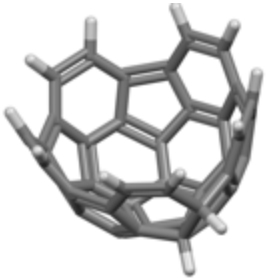
Linear molecule



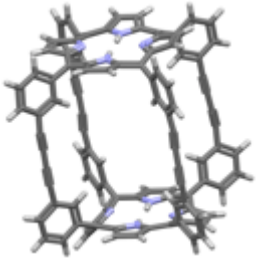
Molecular belt



Complicated shape

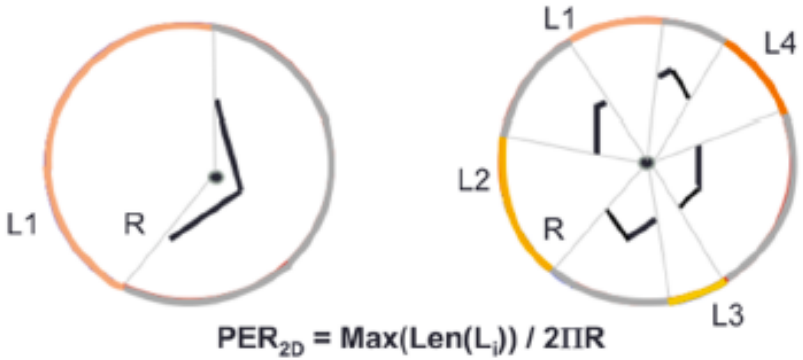


Molecular cup



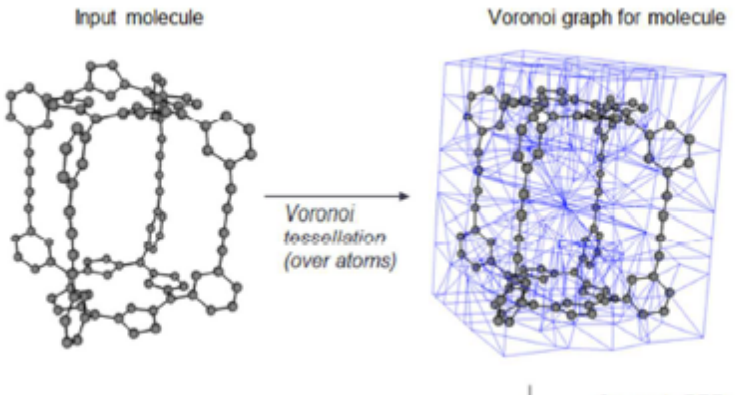
Molecular cage

Pore Exposure Ratio (PER) Molecular Descriptor

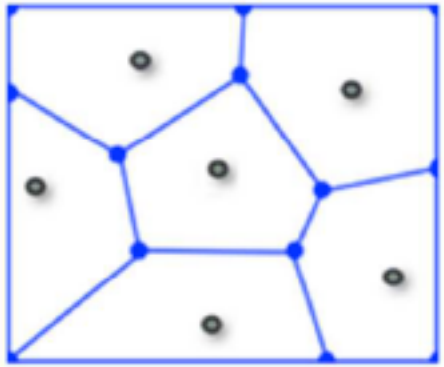


$$PER_{2D} = Max(Len(L_i)) / 2\Pi R$$

Outline of PER Calculation (1)

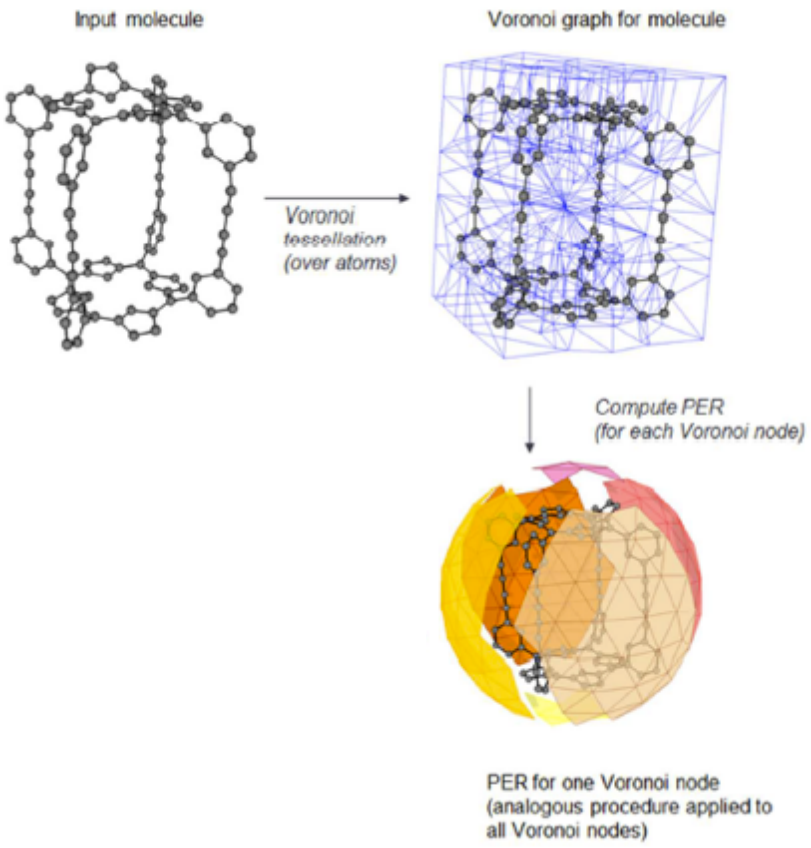


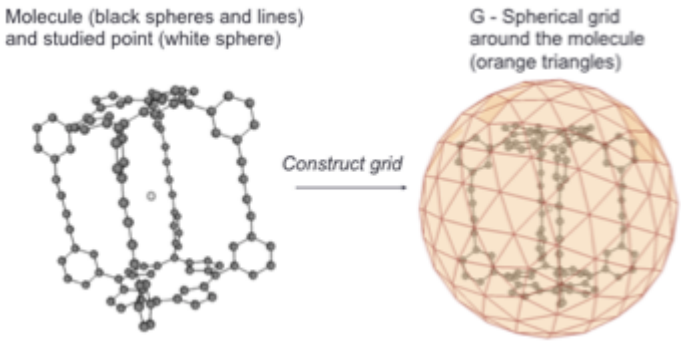
Identification of points highlighting the void space around atoms – Use of Voronoi nodes



Outline of PER Calculation (2)

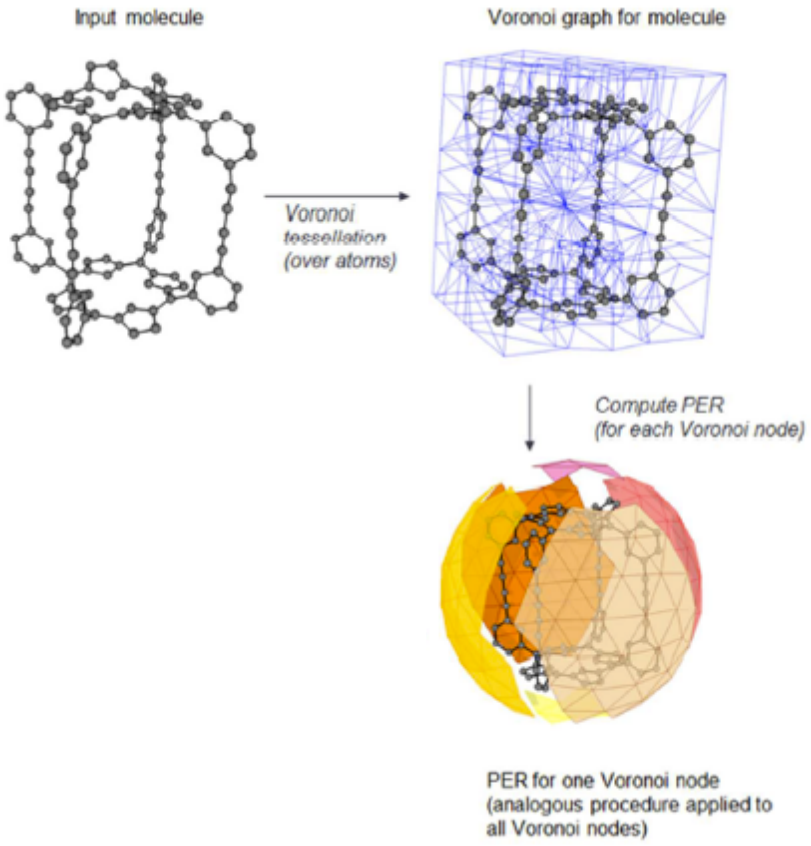
For each node, PER value is calculated

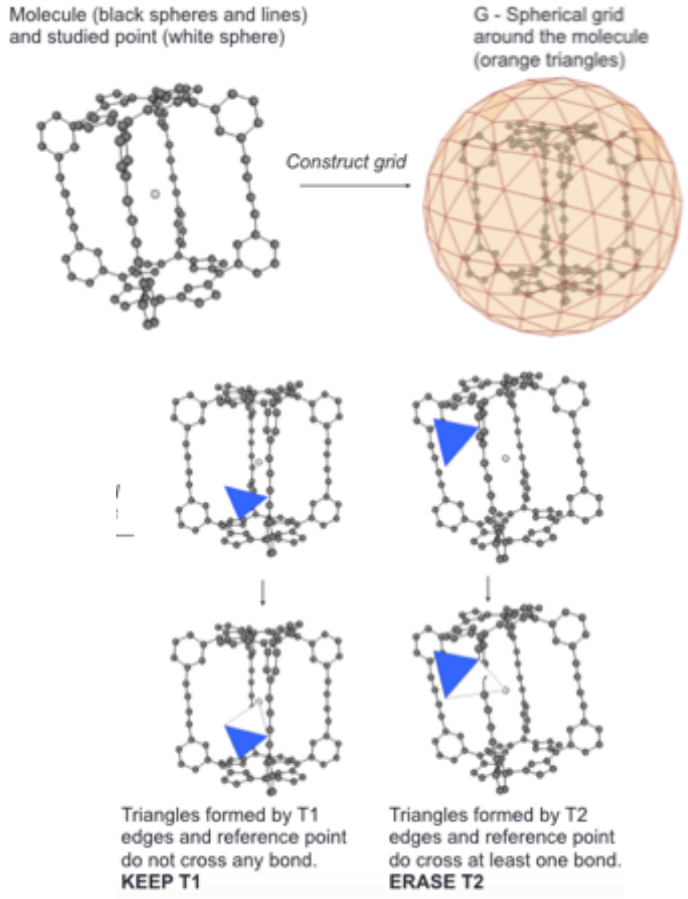




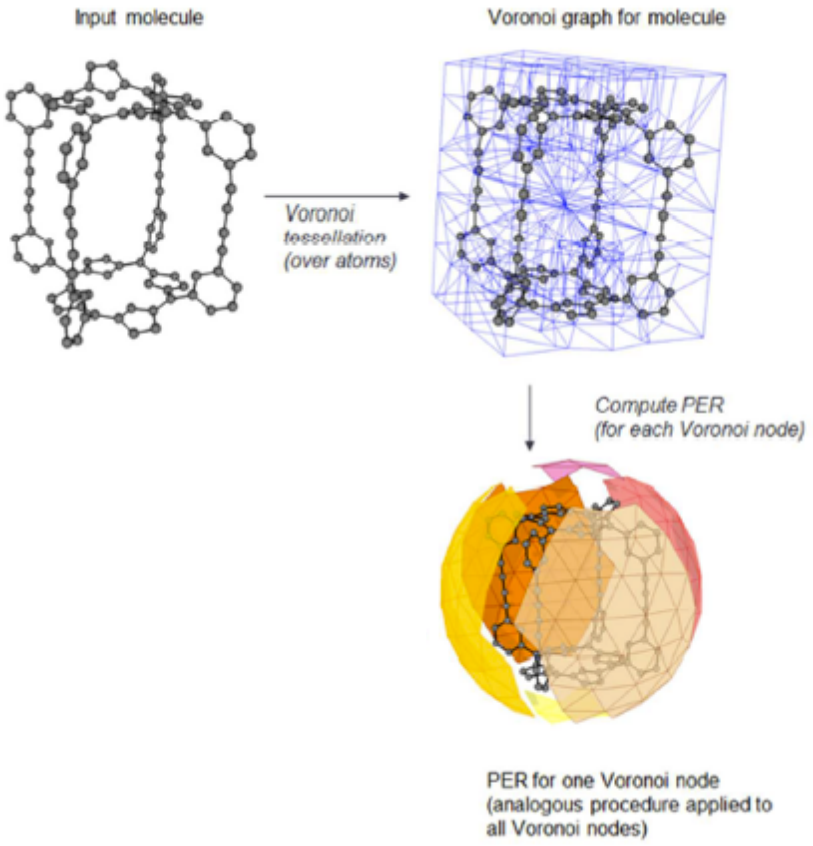
Outline of PER Calculation (3)

For each node, PER value is calculated

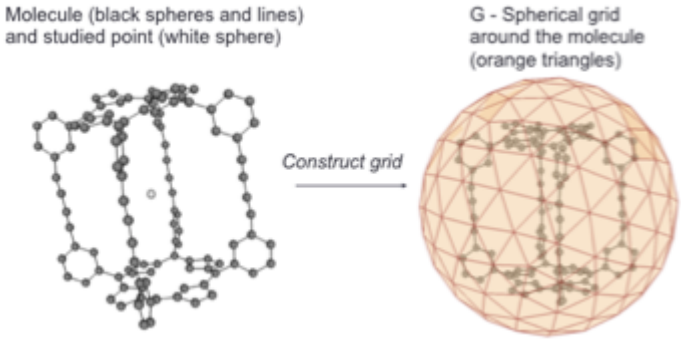




Outline of PER Calculation (4)



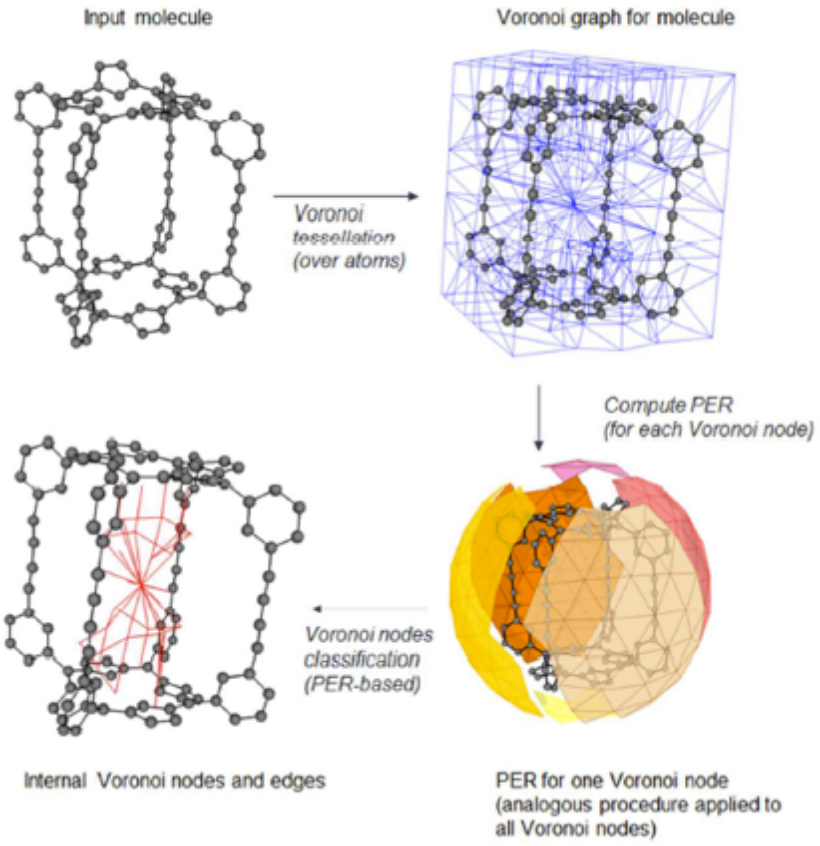
For each node, PER_{node} value is calculated



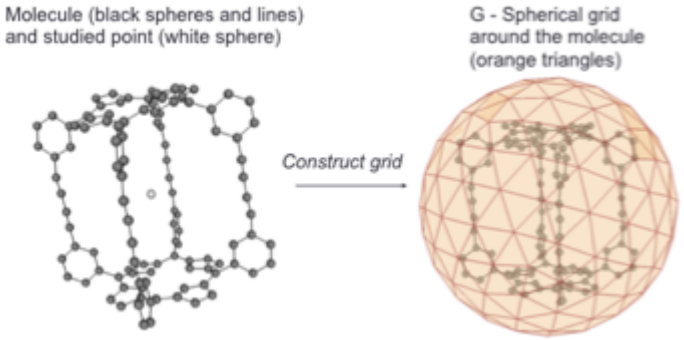
$$PER_{node} = \frac{\text{Area (LCC)}}{\text{Area (Spherical grid)}}$$

LCC – Largest connected component made of triangles

Outline of PER Calculation (5)



For each node, PER_{node} value is calculated

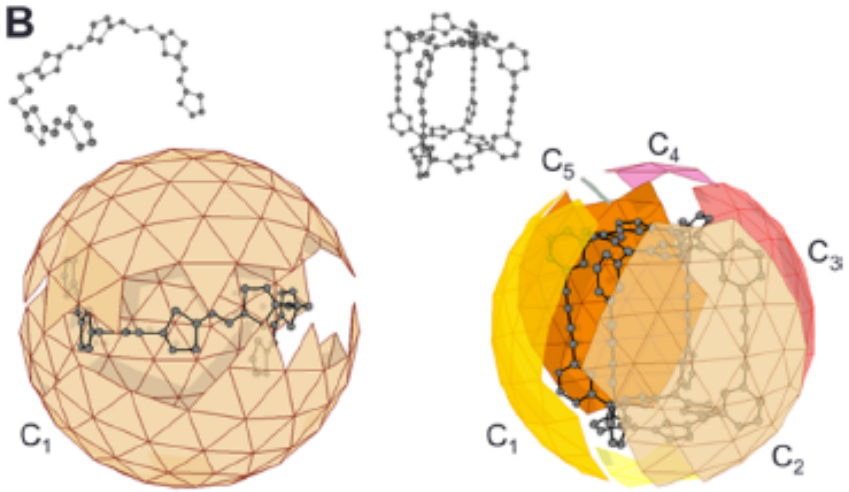


Internal Voronoi Node IF PER_{node} < Threshold

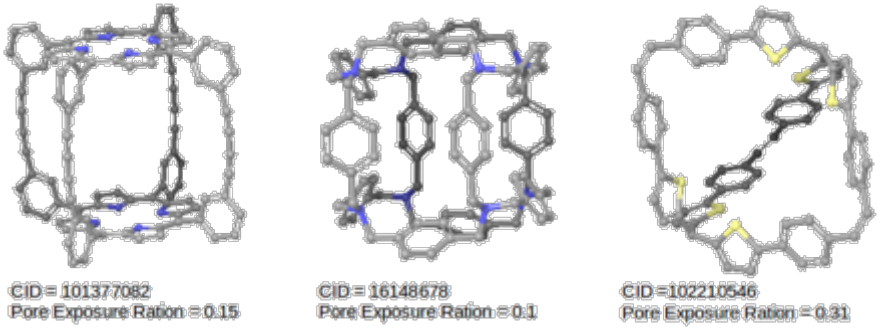
PER = PER_{node} (Largest Internal Voronoi node)

Algorithms being implemented in Zeo++

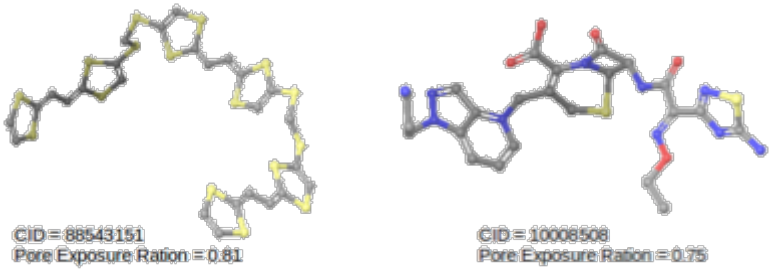
Pore Exposure Ratio (PER) Molecular Descriptor



$PER_{3D} = Max(Surf(C_i)) / Surf(C)$

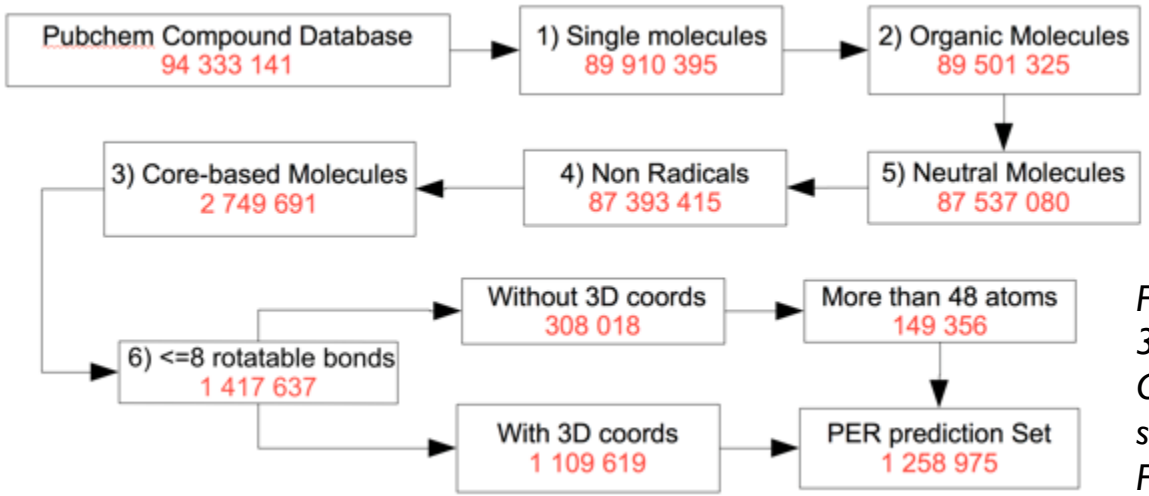


Molecular Cages – PER 0 - 0.36

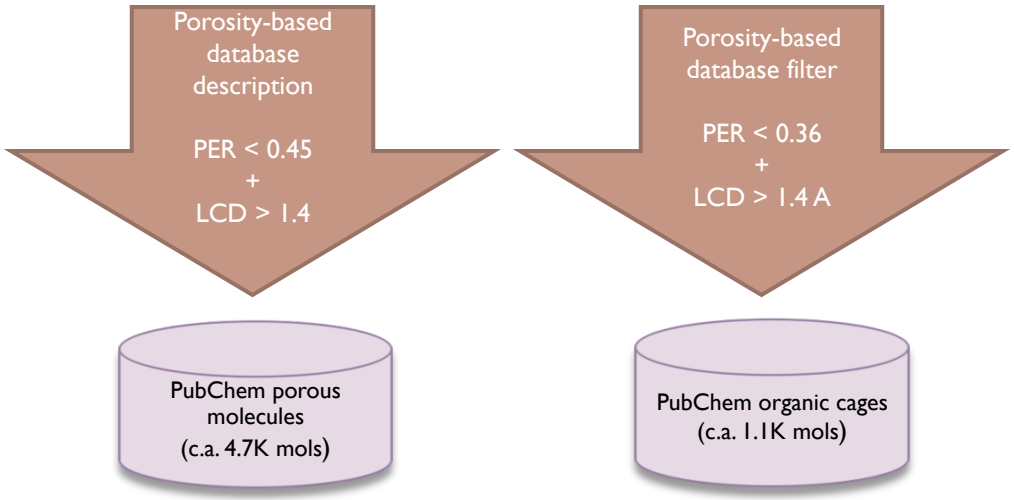


Molecular Non-Cages – PER > 0.36
(Non-porous Molecules - PER > 0.75)

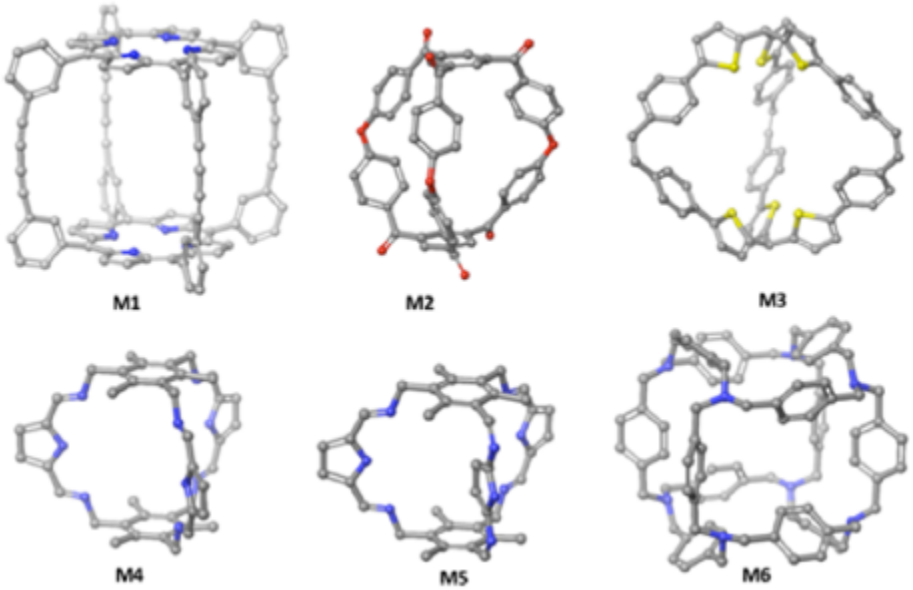
Screening PubChem



Filters: OpenBabel + custom tools
3D Generation: Schrodinger Ligprep
Conformer search: MacroModel, GROMACS, MD/ simulated annealing
Forcefields: OPLS2005, DL_FIELD



Molecular Cage Barn-finds

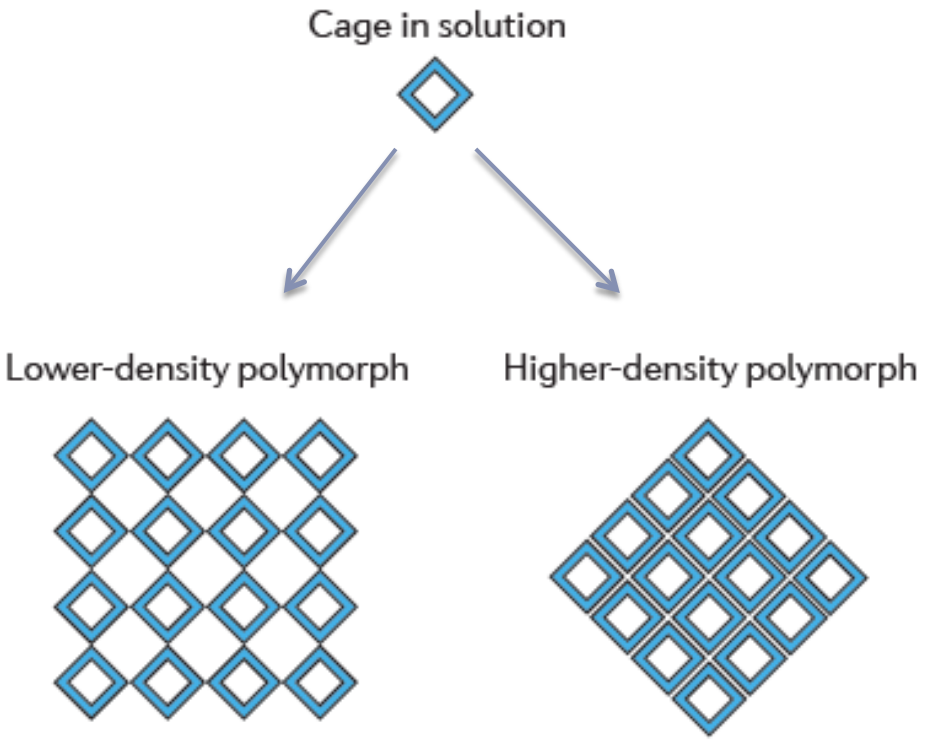


M1	101377082	Japanese	We are not able to obtain a copy of the proceeding article to confirm the synthesis of M1.
M2	102021452	English	The reference article mentions Molecule M2 as a possible outcome of a reaction. The synthesis of M2, however , is not confirmed.
M3	102210546	English	The reference discusses an analogue of M3 in which vertex H atoms are substitute with Me groups. We expect that M3 can be synthesized with the same reaction path as the Me analogue.
M4	102263757	Japanese	We are not able to obtain a copy of the reference. However synthesis of M4 was reported in another article (Tetrahedron Letters 51 (2010) 6521–6525).
M5	102333795	Japanese	We are not able to obtain a copy of the proceeding article to confirm the synthesis of M5.
M6	16148678		We are not able to identify the reference article.

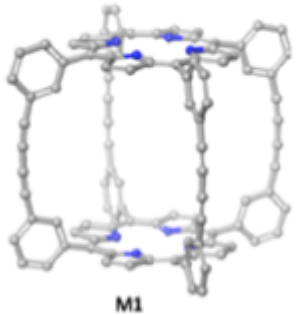
Crystal Structure Prediction

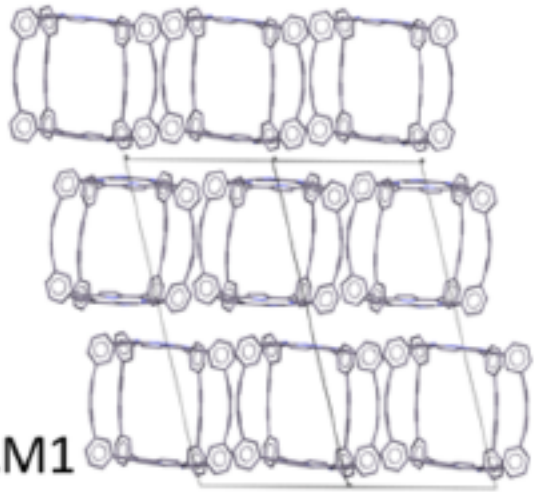
Methodology:

- Systematic search through 13 most common space groups
- 5000 random configurations
- OPLS2005 forcefield
- UPACK code
- Porosity characterization with Zeo++

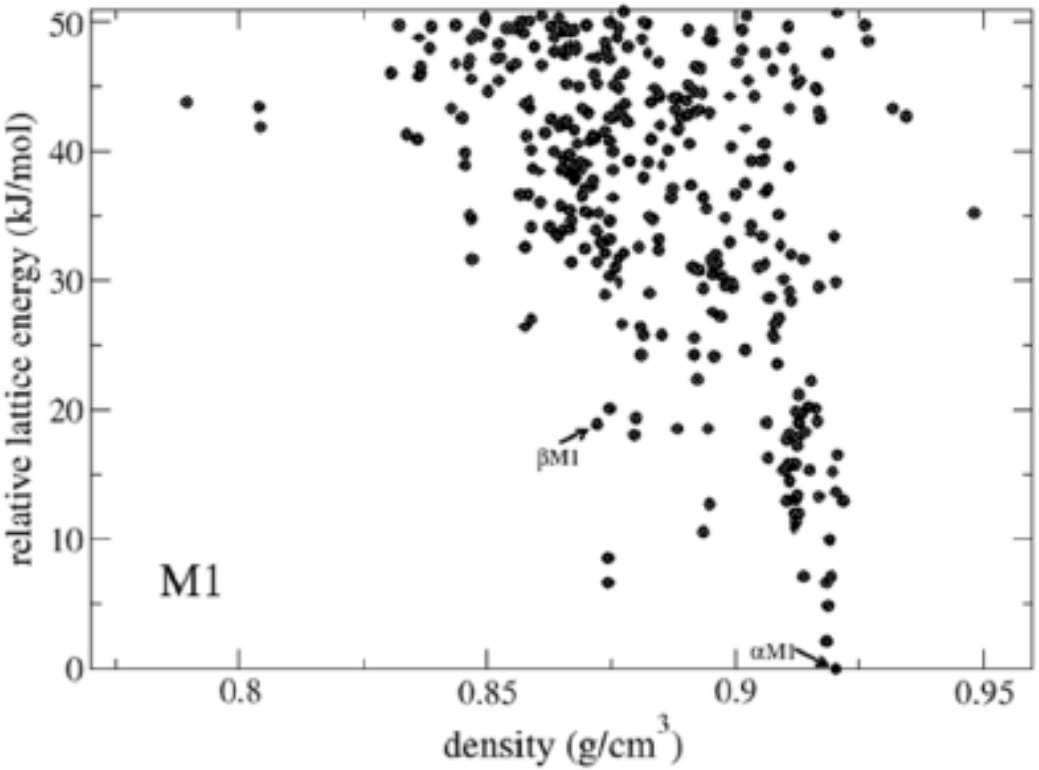


M1 Crystal Structure Prediction (1)

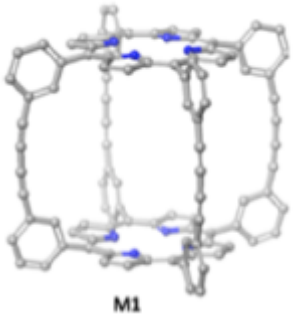


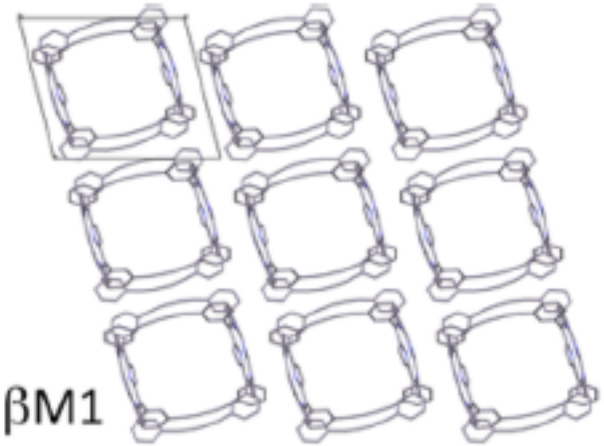


ΔE = 0 kJ/mol ρ = 0.920 g/cm³
PLD = 4.27Å LCD = 8.17Å
SA = 1071 m²/g

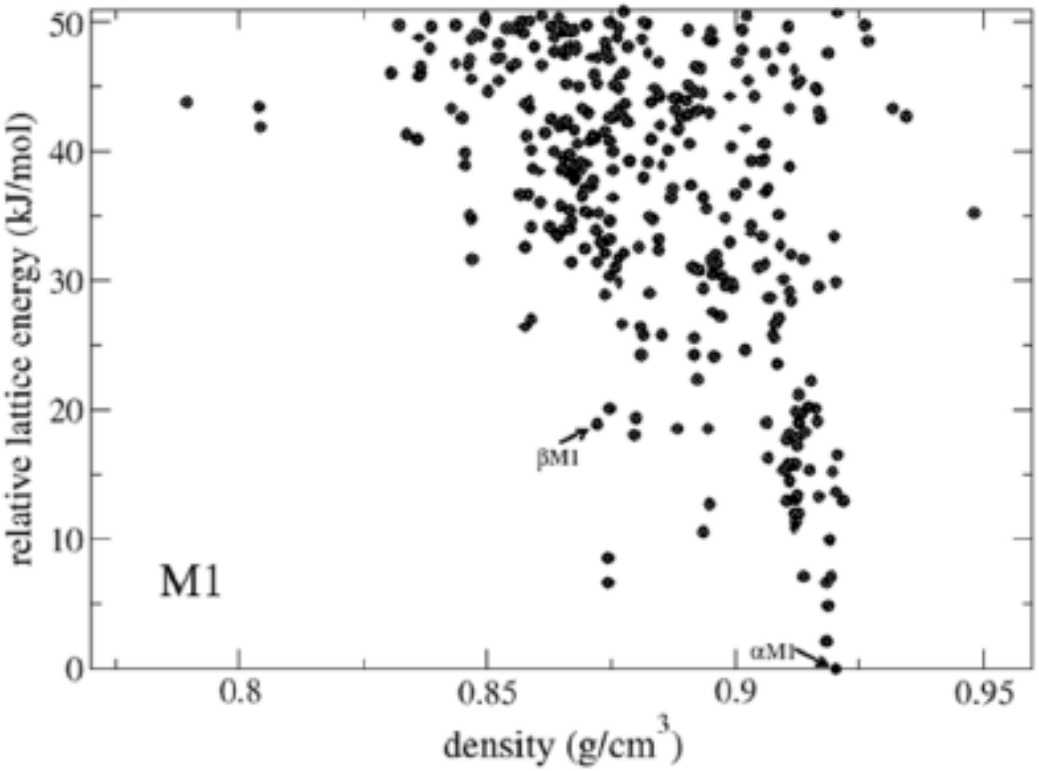


M1 Crystal Structure Prediction (2)

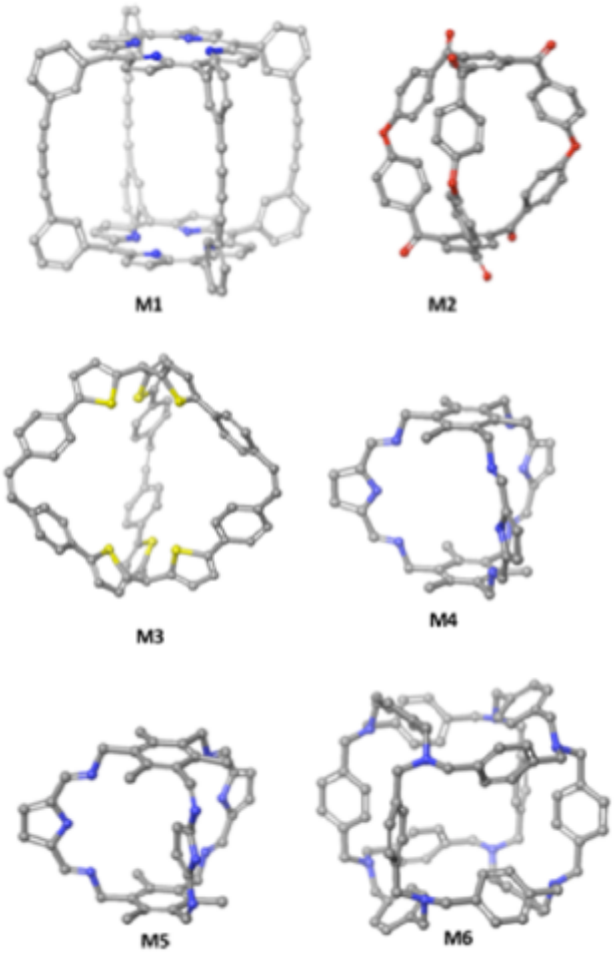


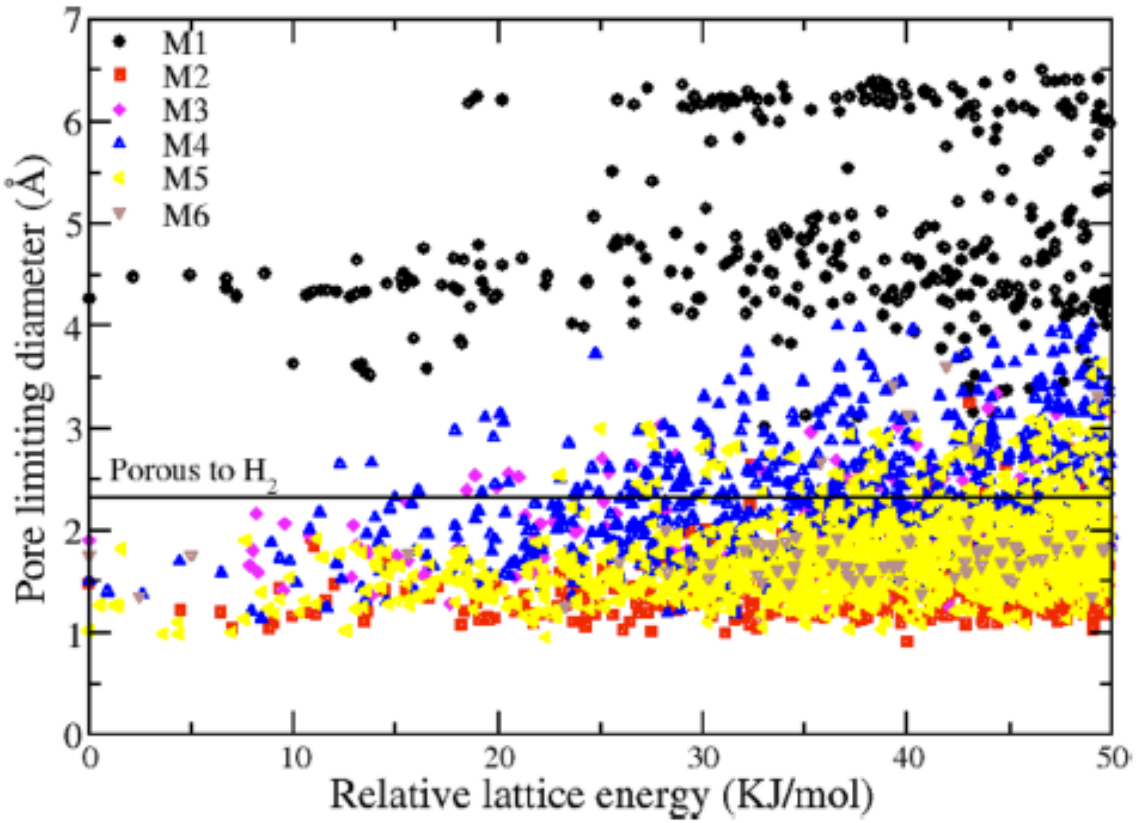


ΔE = 19.0 kJ/mol ρ = 0.872 g/cm³
PLD = 6.25 Å LCD = 8.32 Å
SA = 1207 m²/g



Crystalline Phases of M1-M6





I. Gomez Garcia, CG&D, 2017

Conclusions

- Nanoporous materials, with their chemical diversity and industrial importance, represent an exciting and enormous search space – novel algorithms to execute searches are needed
- In many cases shape/geometry is a sufficient proxy for molecular interactions; therefore applications of computational geometry can have a huge impact
- Opportunities for material discovery via optimization, search algorithms and machine learning are clear
- Feature engineering (esp. physics-based descriptors) and statistical-based analysis of large sets of structures offer collaboration opportunities between mathematicians and domain scientists

Thank you