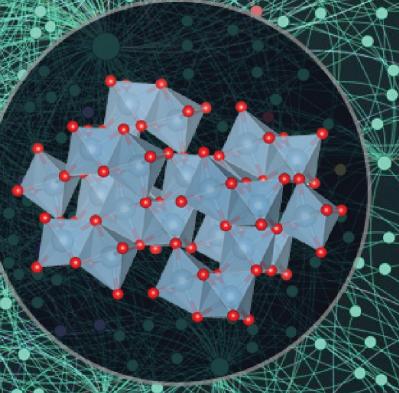
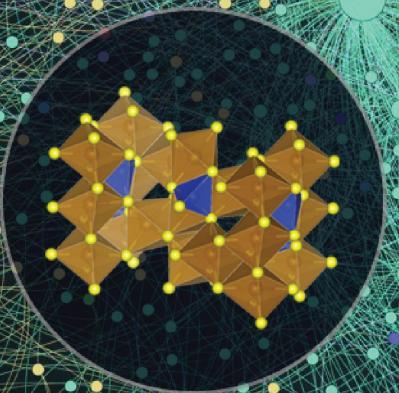
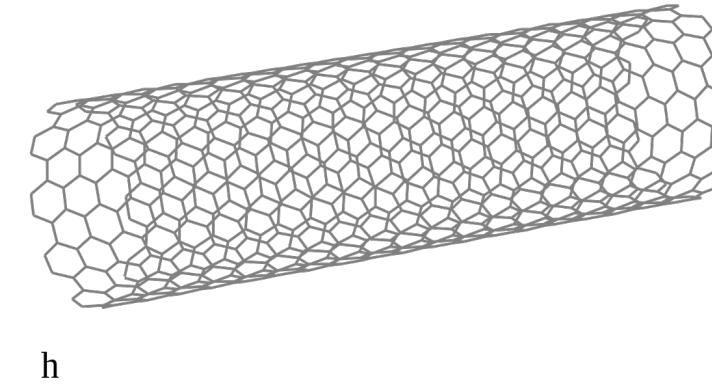
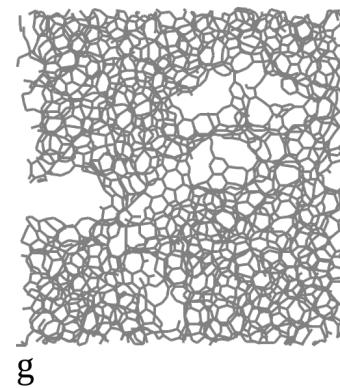
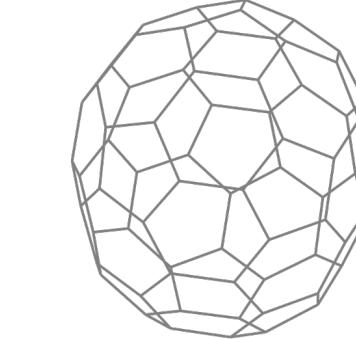
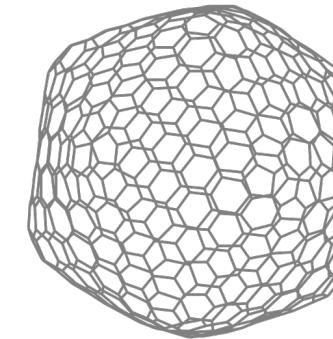
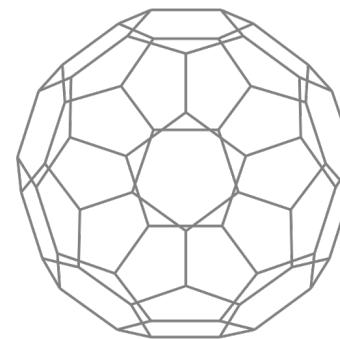
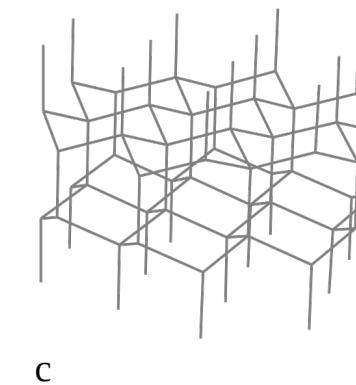
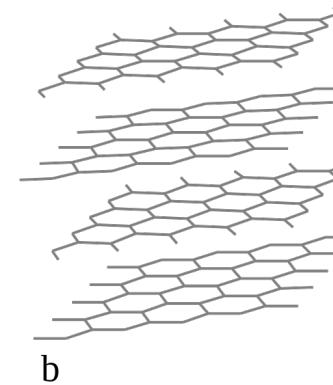
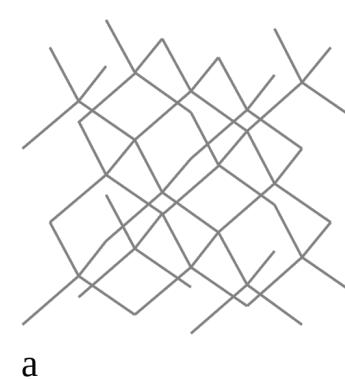


Structure-based feature vector





Carbon is a great example of why we can't ignore structure



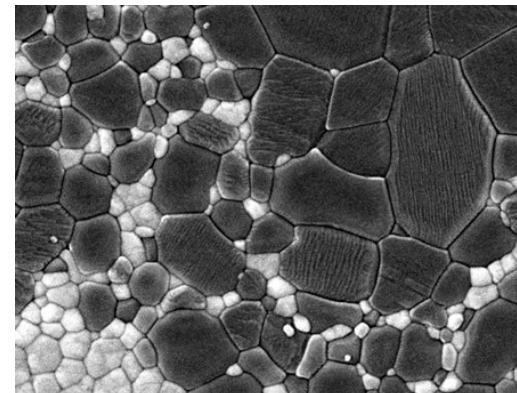
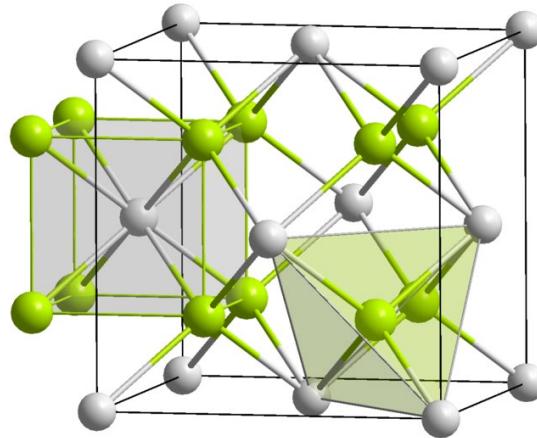
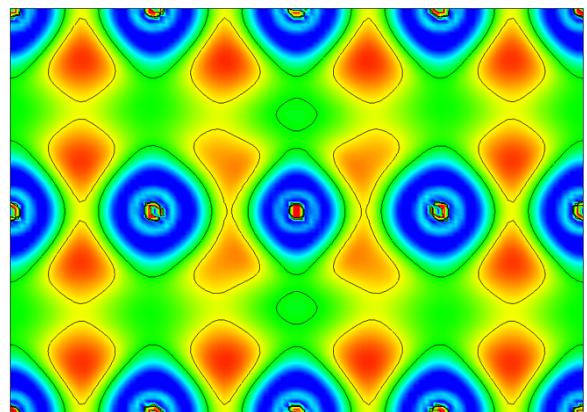
Which structure are we talking about for structural feature?

Electronic

Atomic

Microstructure

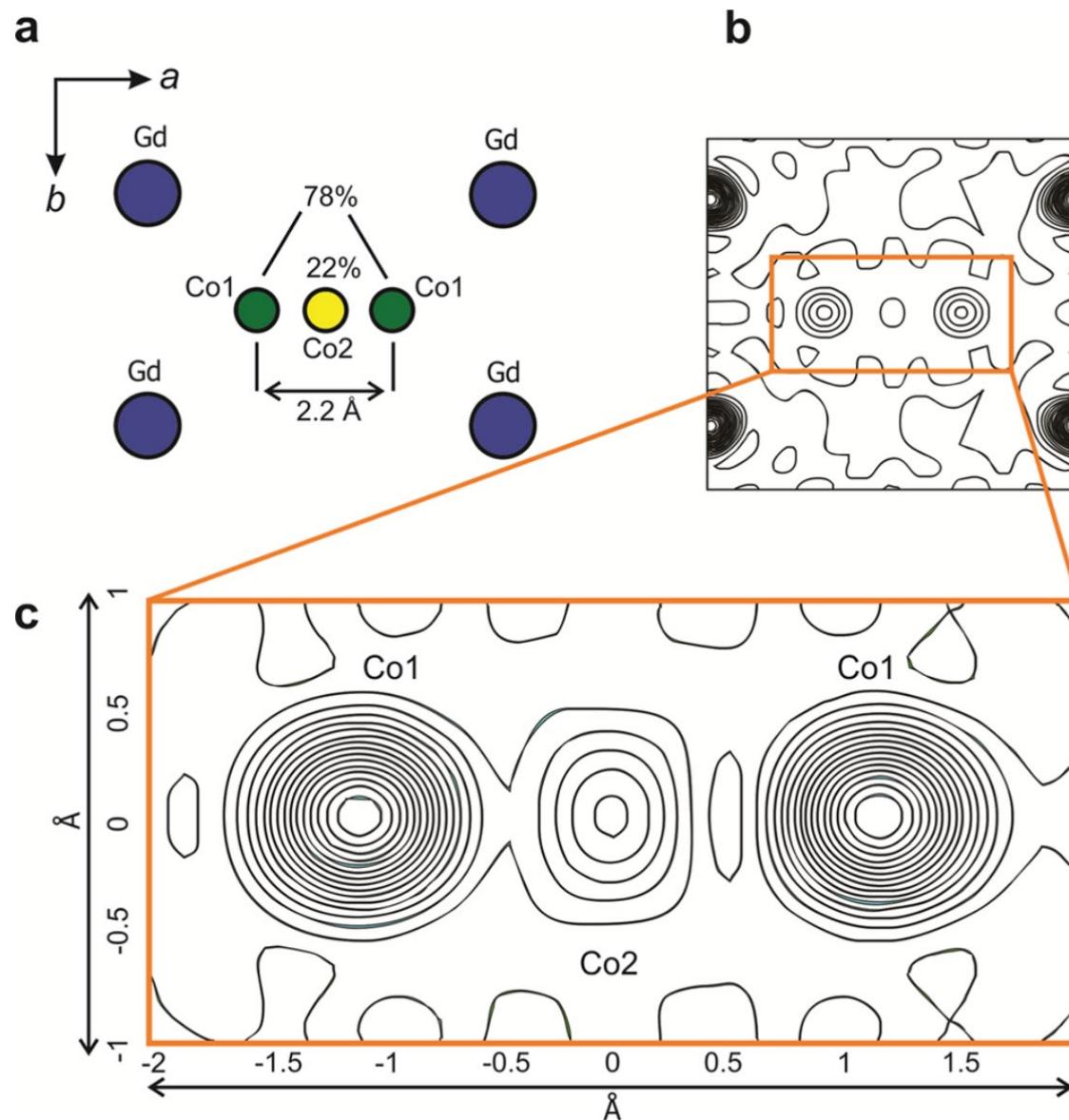
Macroscale



Atomic structure is described with atomic numbers

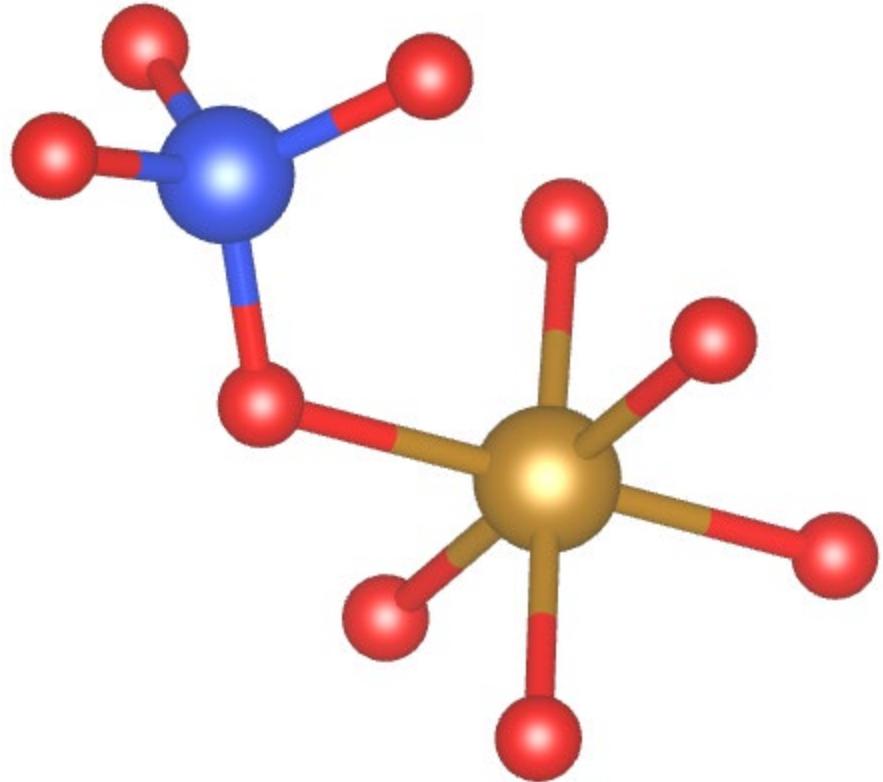


Atomic structure can also be approximated by electron density map



Oliynyk et al, *Inorg Chem*, 55, 13, 6625-6633, (2016)

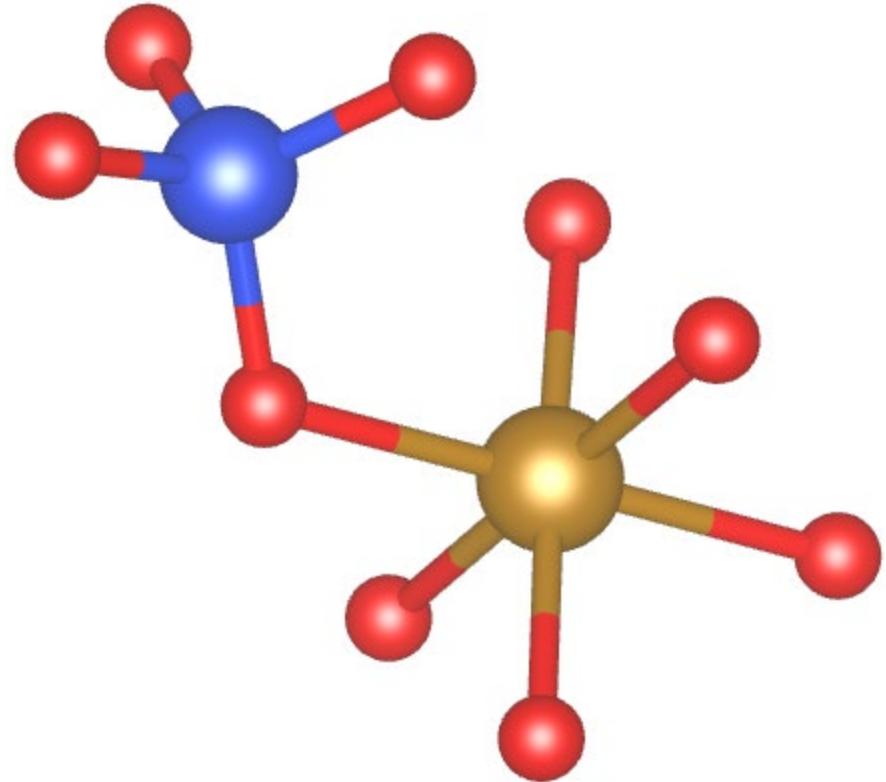
Crystal structure has different features at different scales



Bonds

- Type
- Elements involved
- Length
- Distribution of different bonds

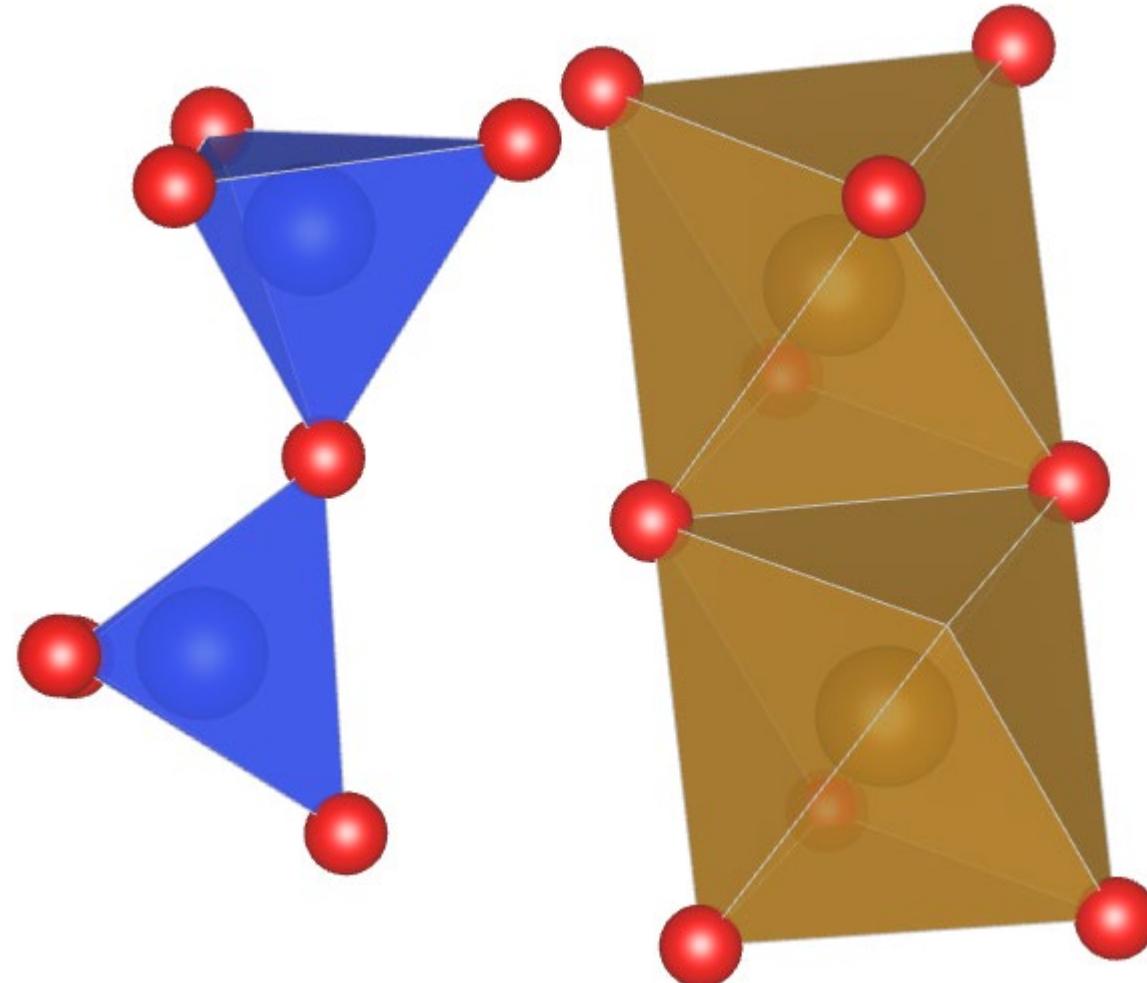
Crystal structure has different features at different scales



Local environment

- Coordination number
- Angles
- Distortions (Jahn-Teller etc)

Crystal structure has different features at different scales

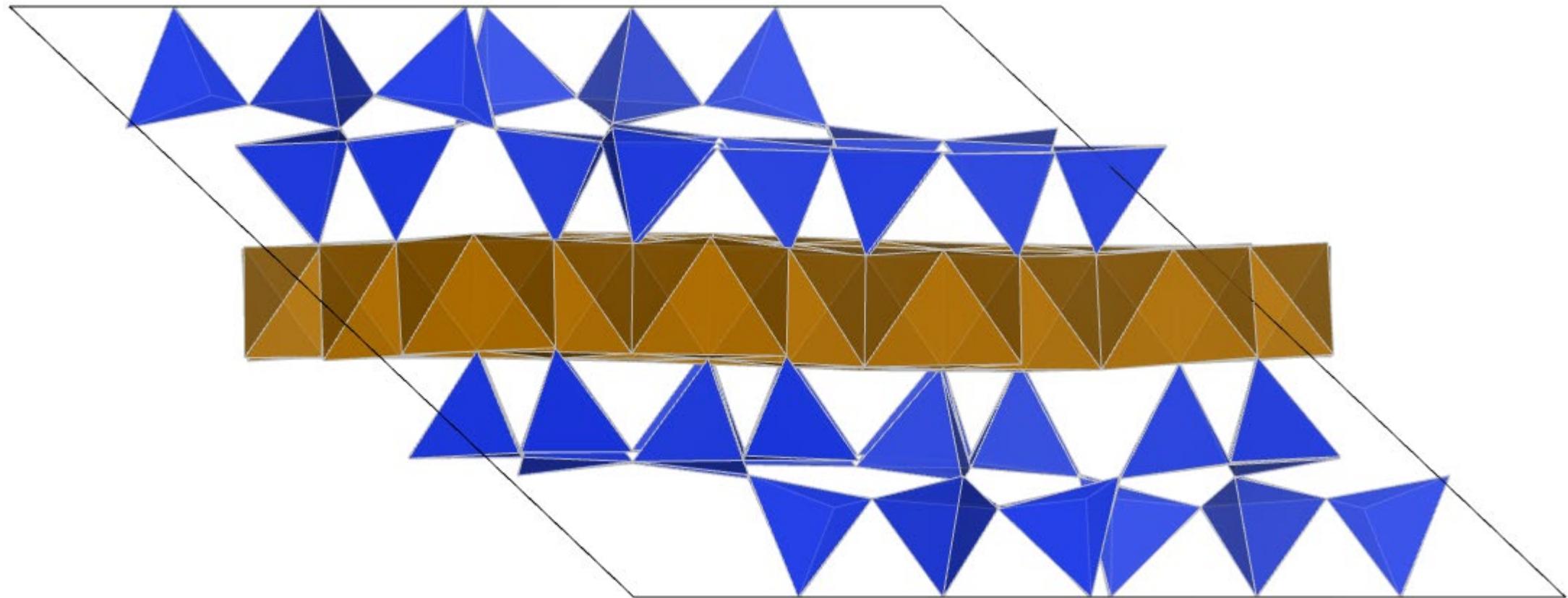


Polyhedra

- isolated vs corner vs edge vs face sharing
- multiple types
- angles
- distances

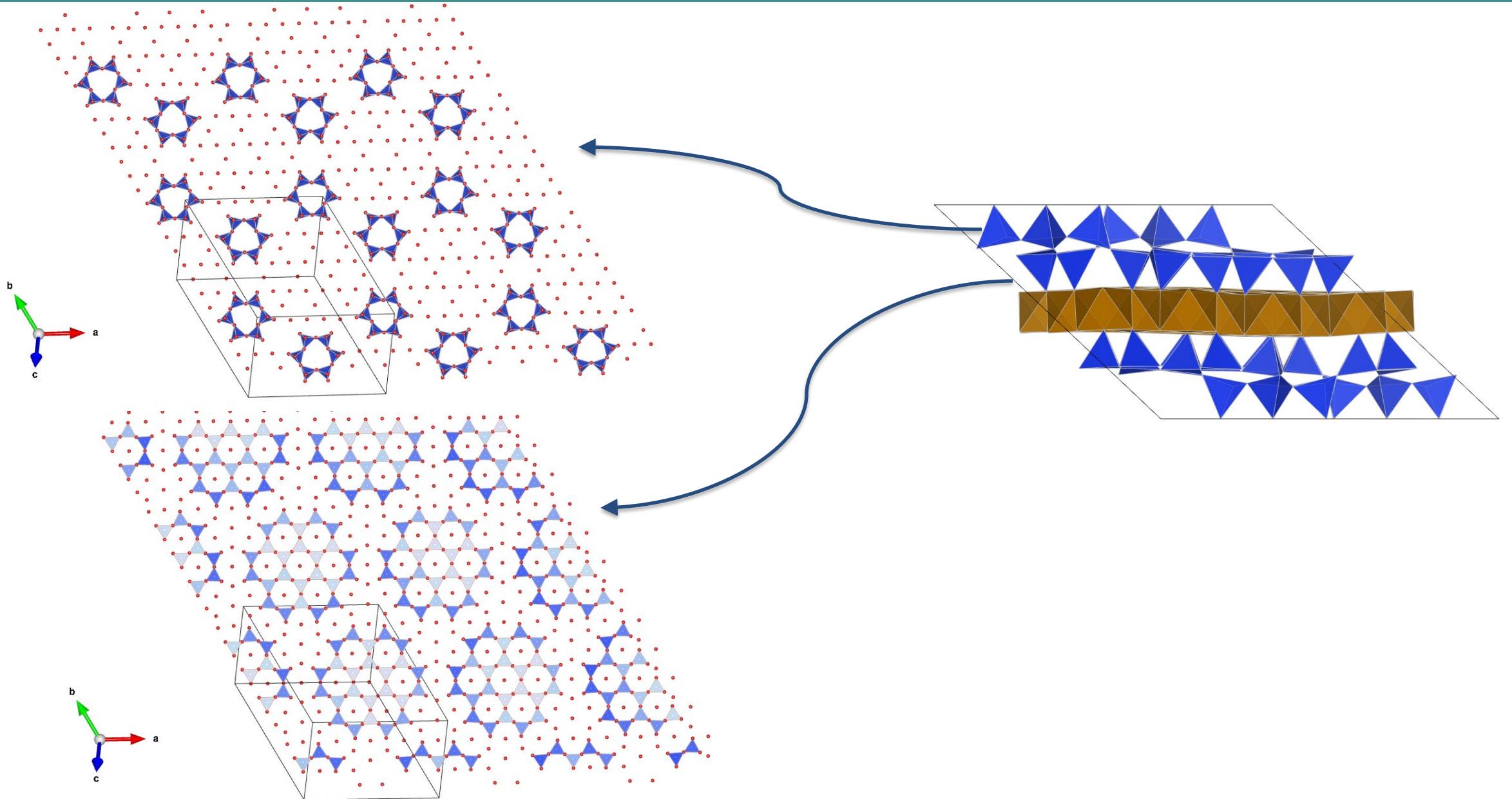


Atomic features can be atomic scale... Or unit cell scale...

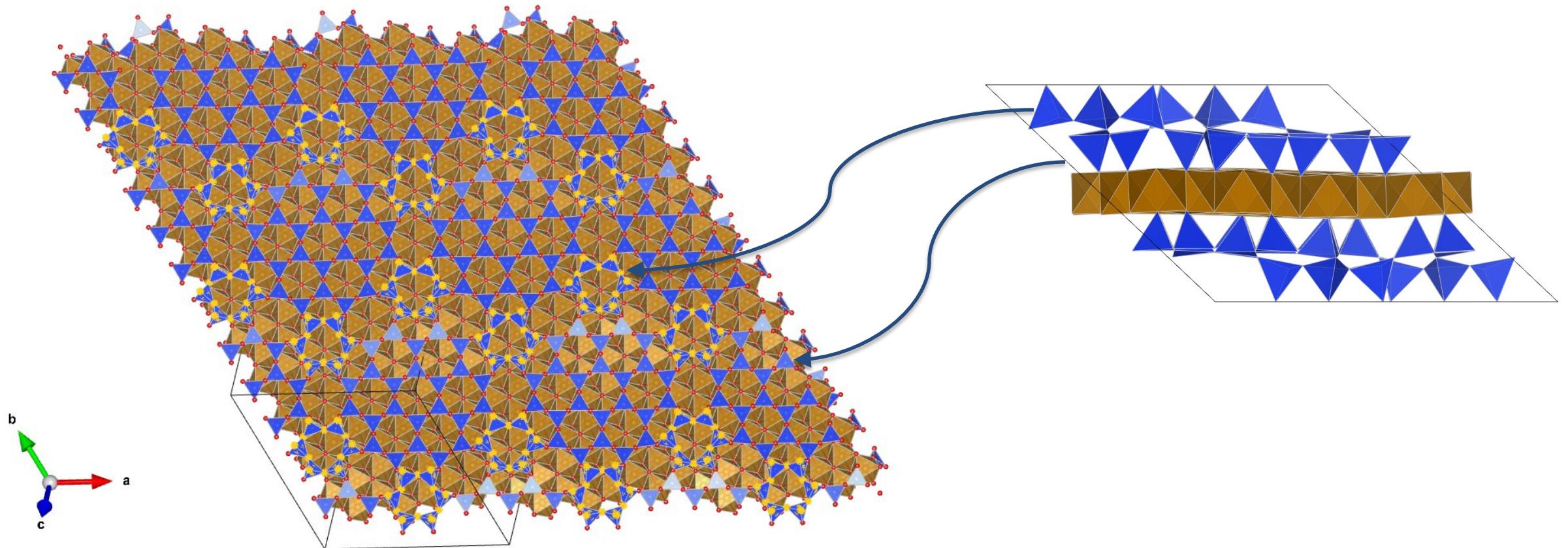




Some unit cell features are “slabs” or layers

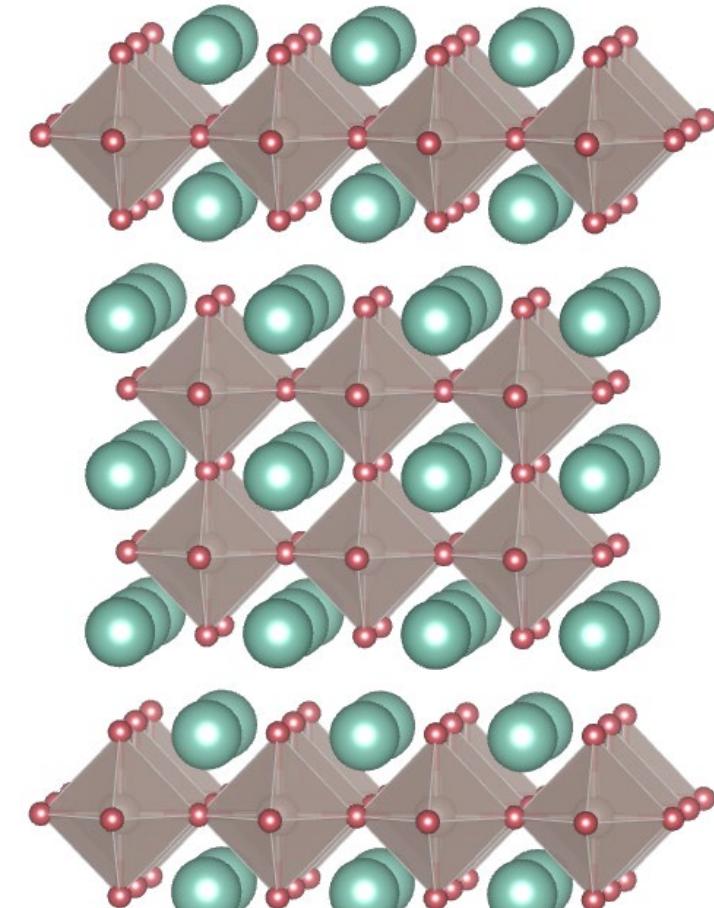
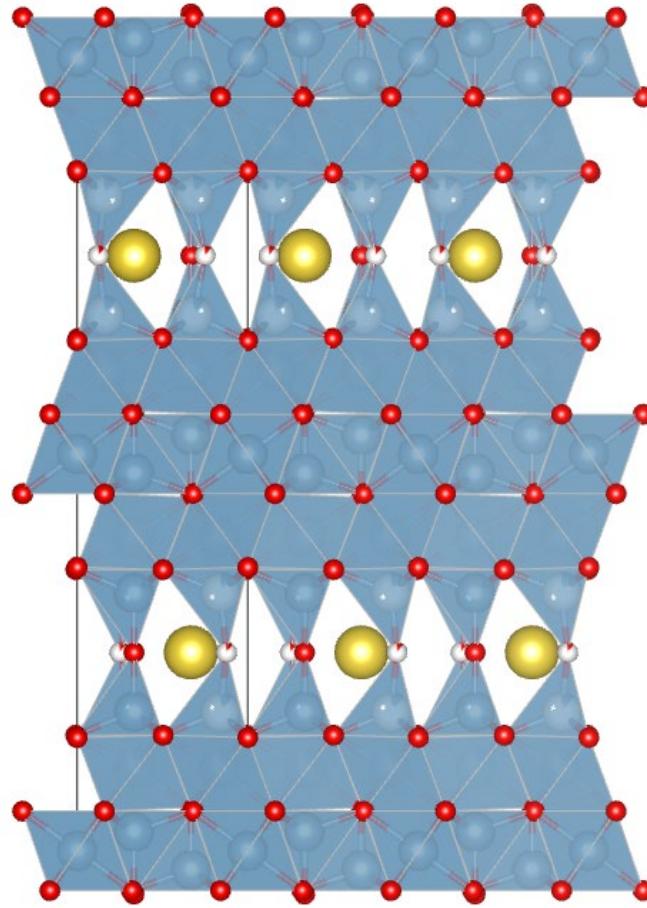
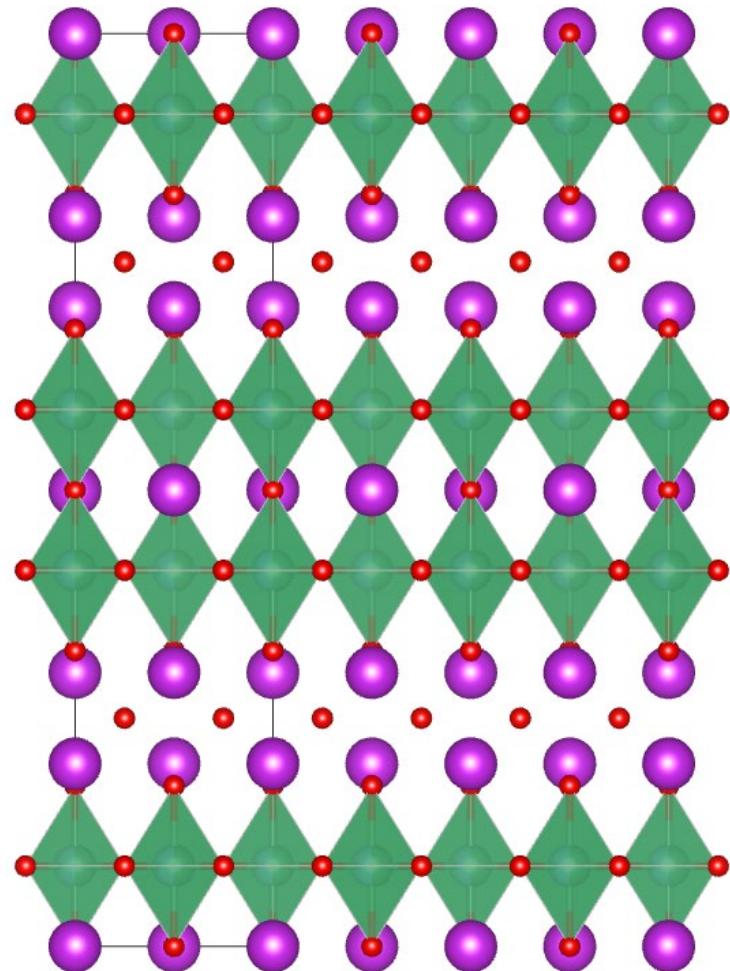


Layers don't need to be continuous



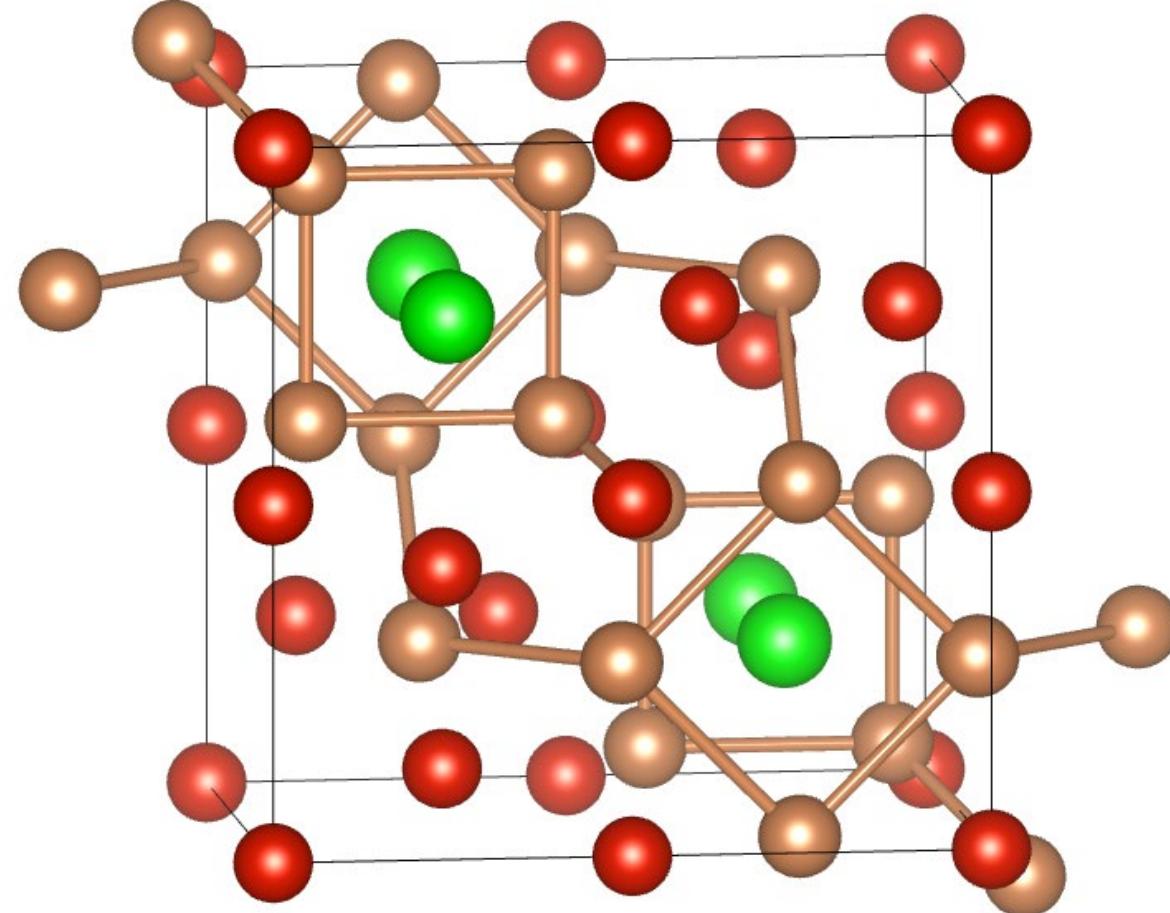


Whole structures are made up of slabs



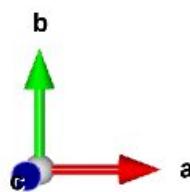
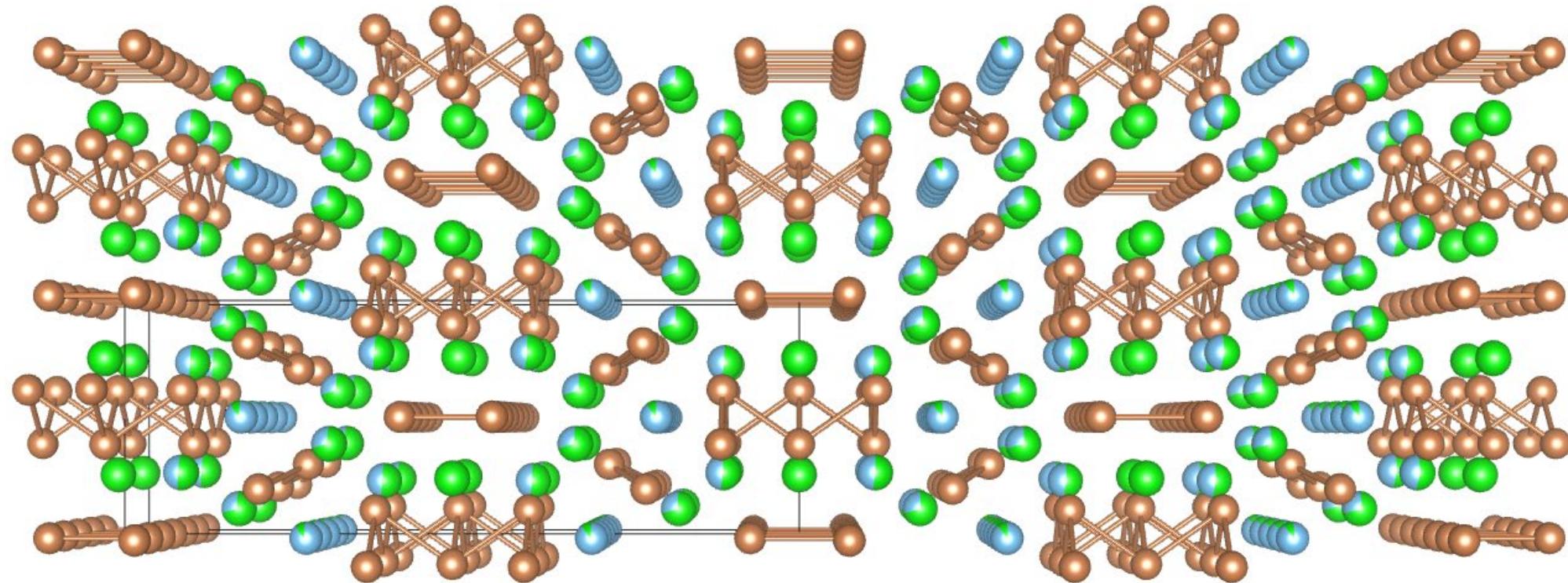


Chains and planes of atoms can define properties



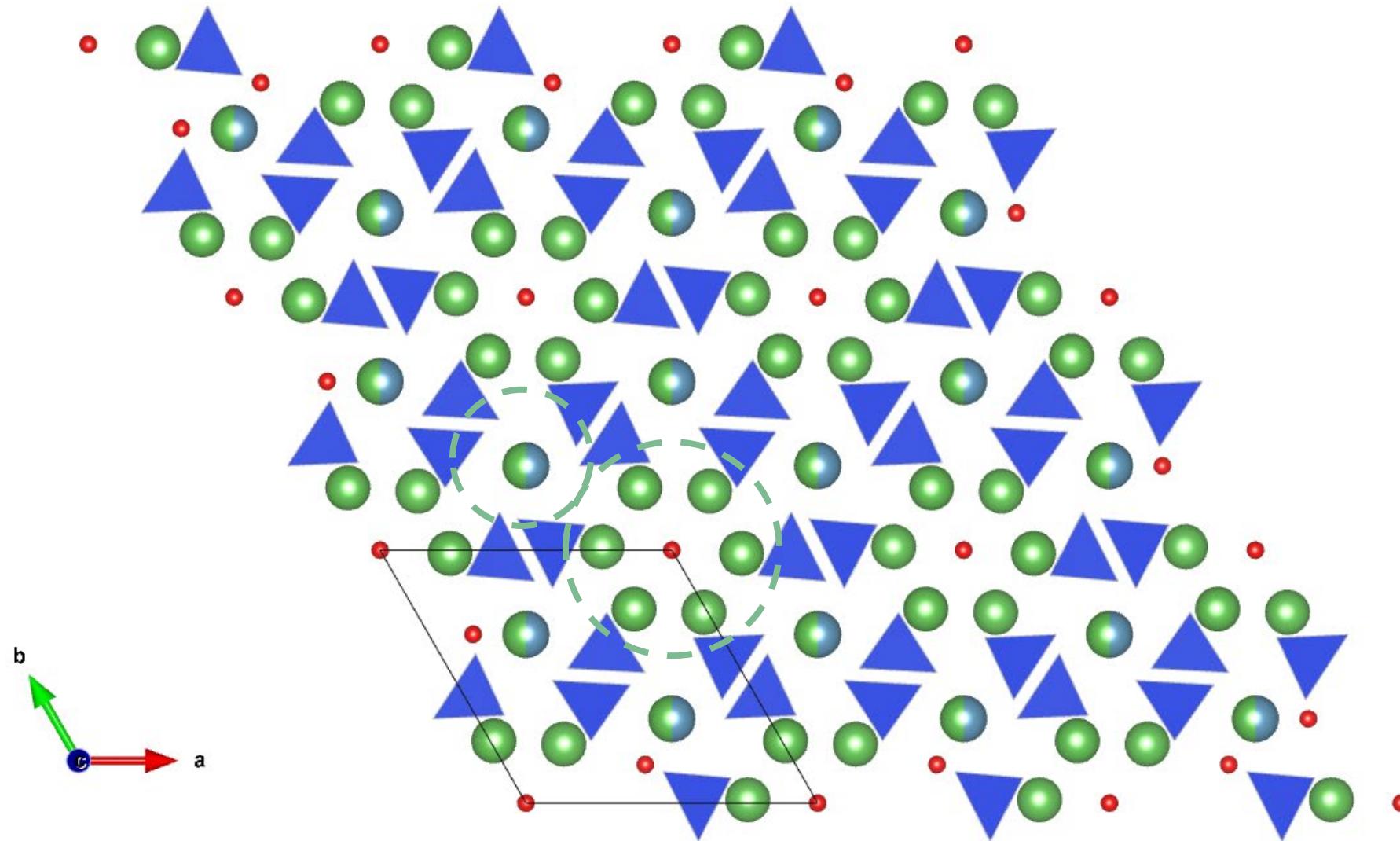


Chains and planes of atoms can define properties



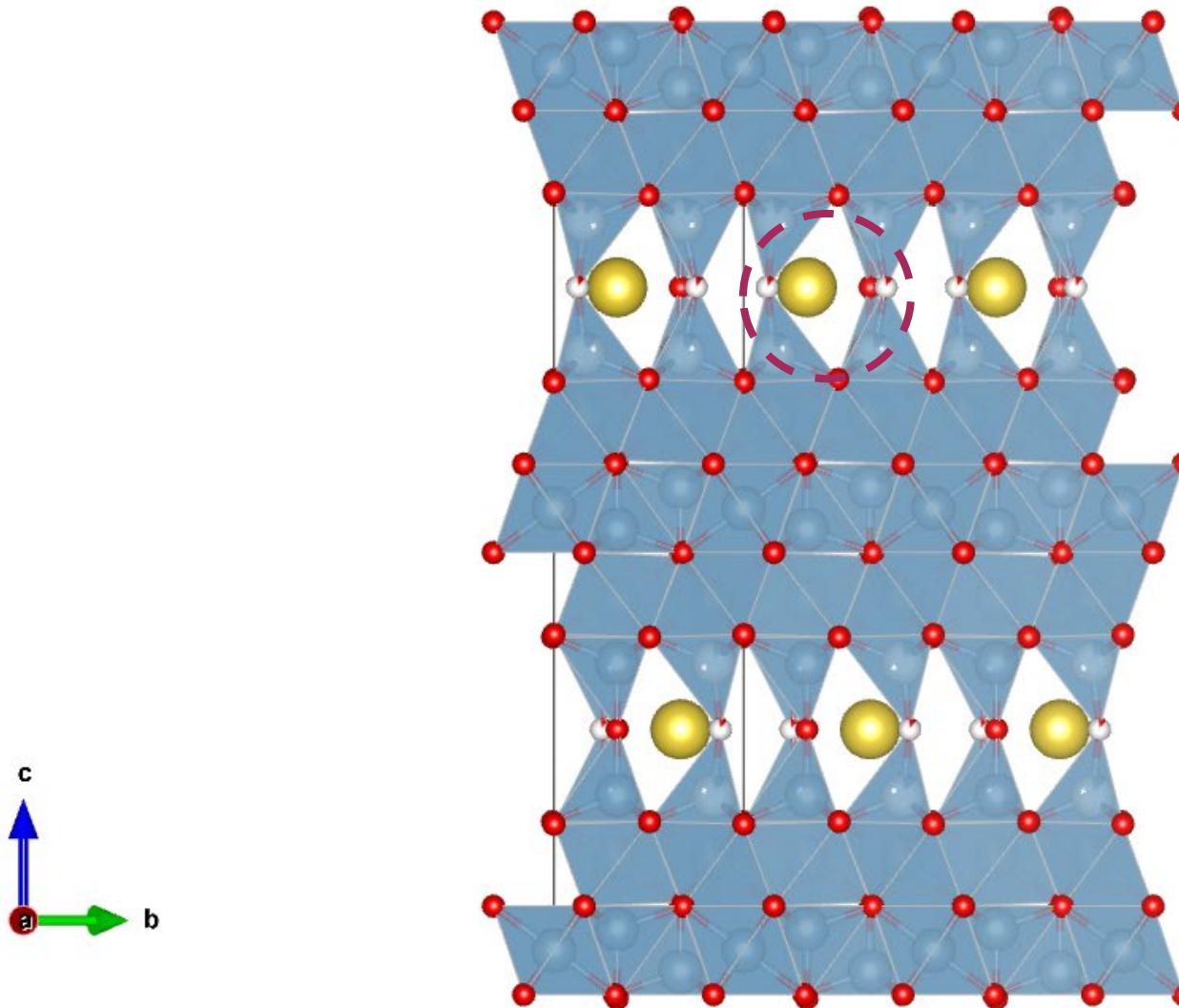


Negative space not occupied by atoms can be the feature



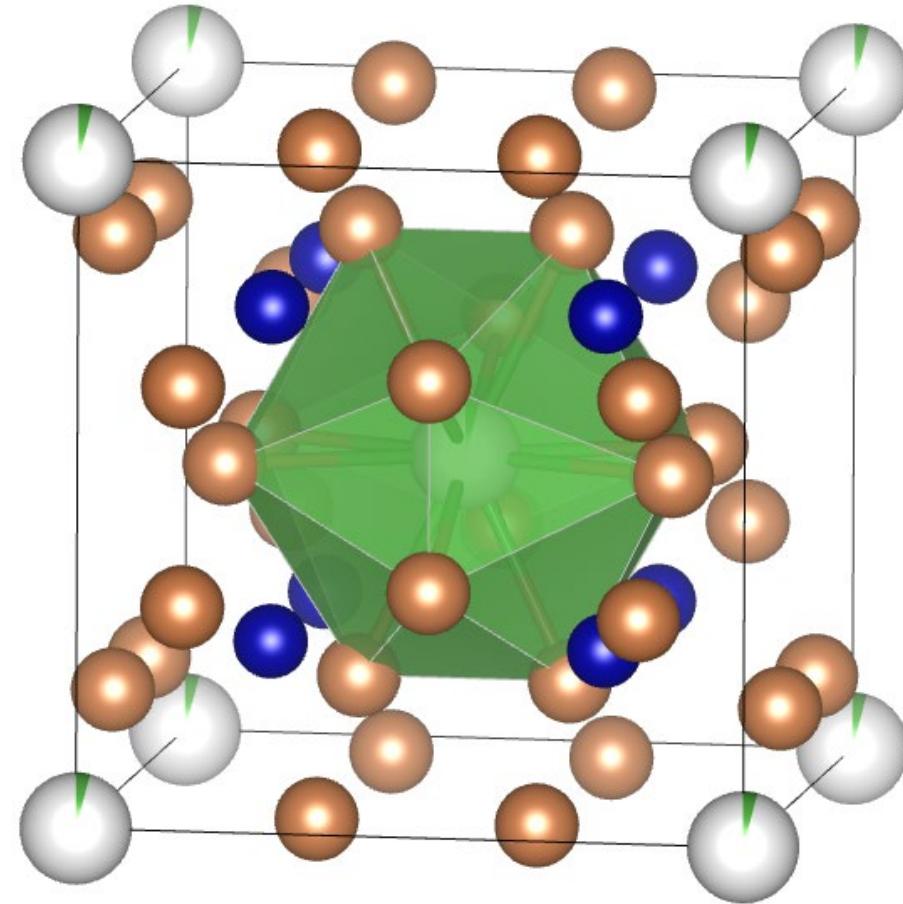
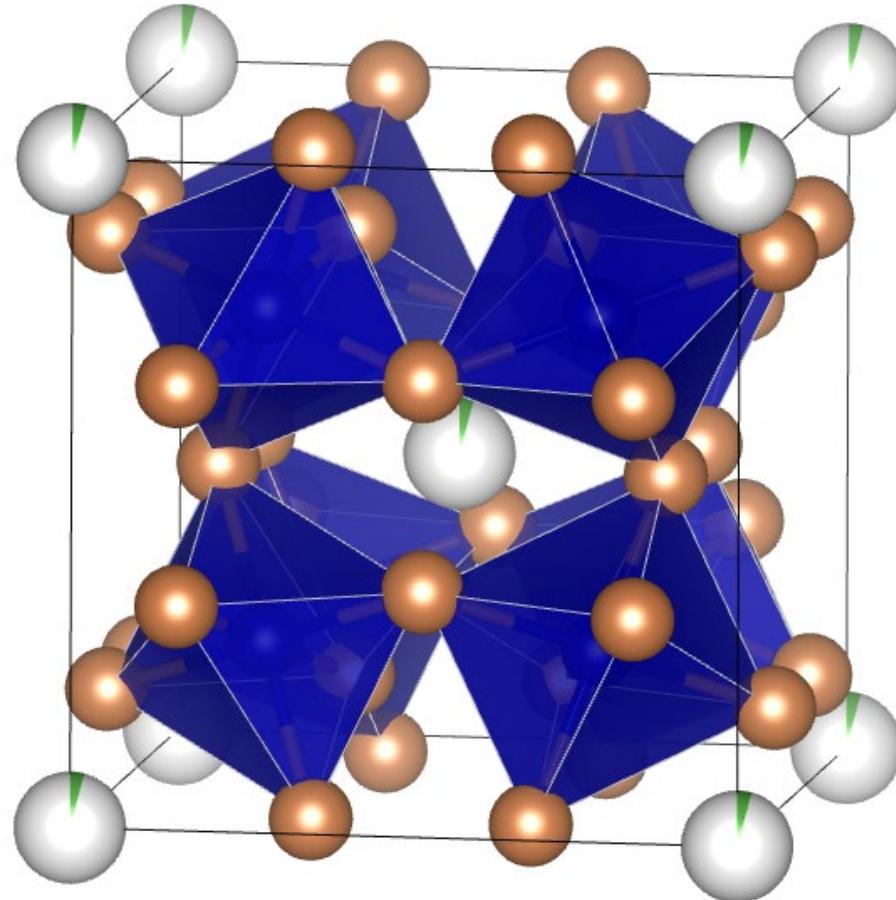


Negative space not occupied by atoms can be the feature

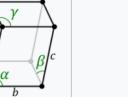
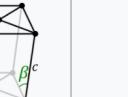
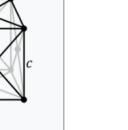
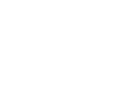




Negative space not occupied by atoms can be the feature



We could define structure with symmetry

Crystal family	Lattice system	Point group (Schönflies notation)	14 Bravais lattices			
			Primitive (P)	Base-centered (S)	Body-centered (I)	Face-centered (F)
Triclinic (a)		C _i				
Monoclinic (m)		C _{2h}				
Orthorhombic (o)		D _{2h}				
Tetragonal (t)		D _{4h}				
Rhombohedral	Rhombohedral	D _{3d}				
	Hexagonal (h)	D _{6h}				
Cubic (c)		O _h				

7 crystal systems

14 bravais lattices

Lattice parameters (a,b,c)

Angles (α, β, γ)

Centering (primitive, base, body, face)

32 Schoenflies point groups

We could define structure with symmetry



230 space groups for unique symmetries
Symmetry is applied to crystal basis

How do we encode all of these features?

Structural feature	Specific structure
Layered	Yes
Channels	No
RS blocks	Yes
Octahedra	Yes
Tetrahedra	No
...	
Corner-sharing	Yes
Edge-sharing	No

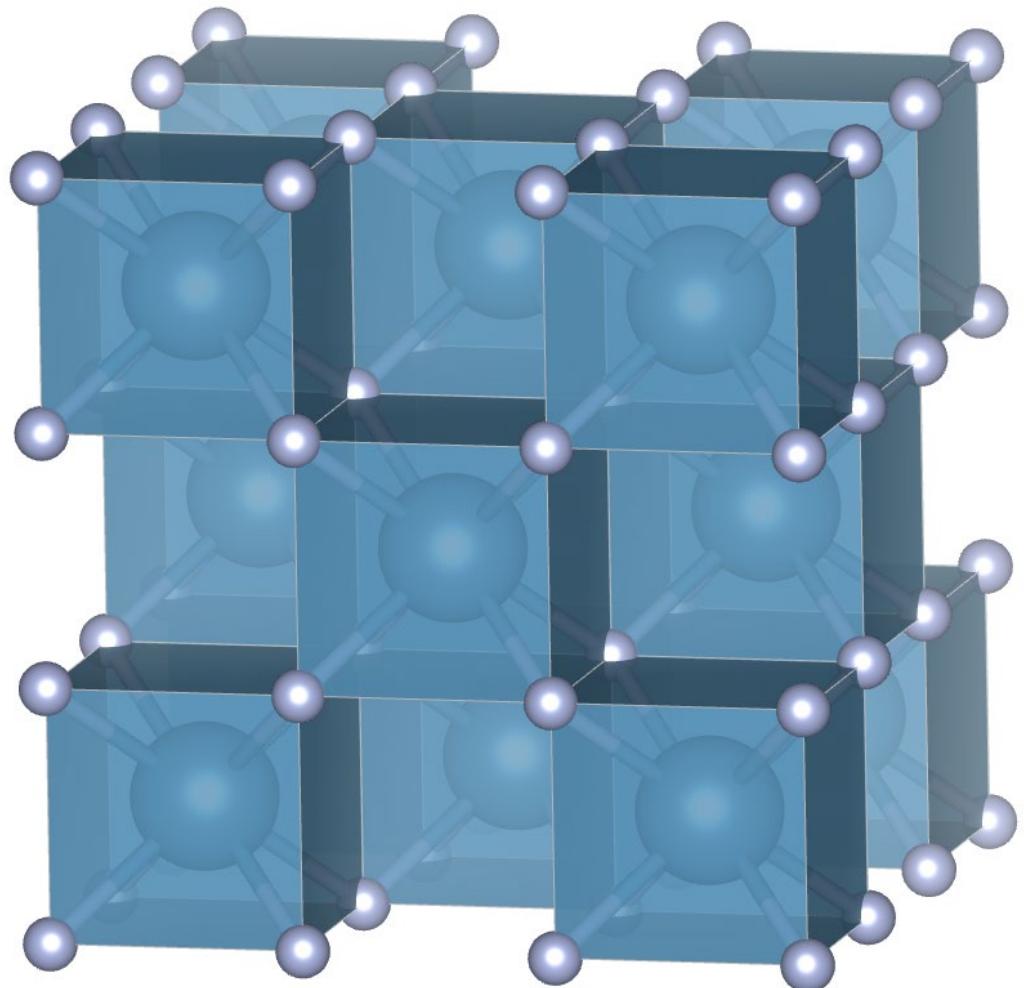
Maybe we could one-hot encode them?

How do we encode all of these features?

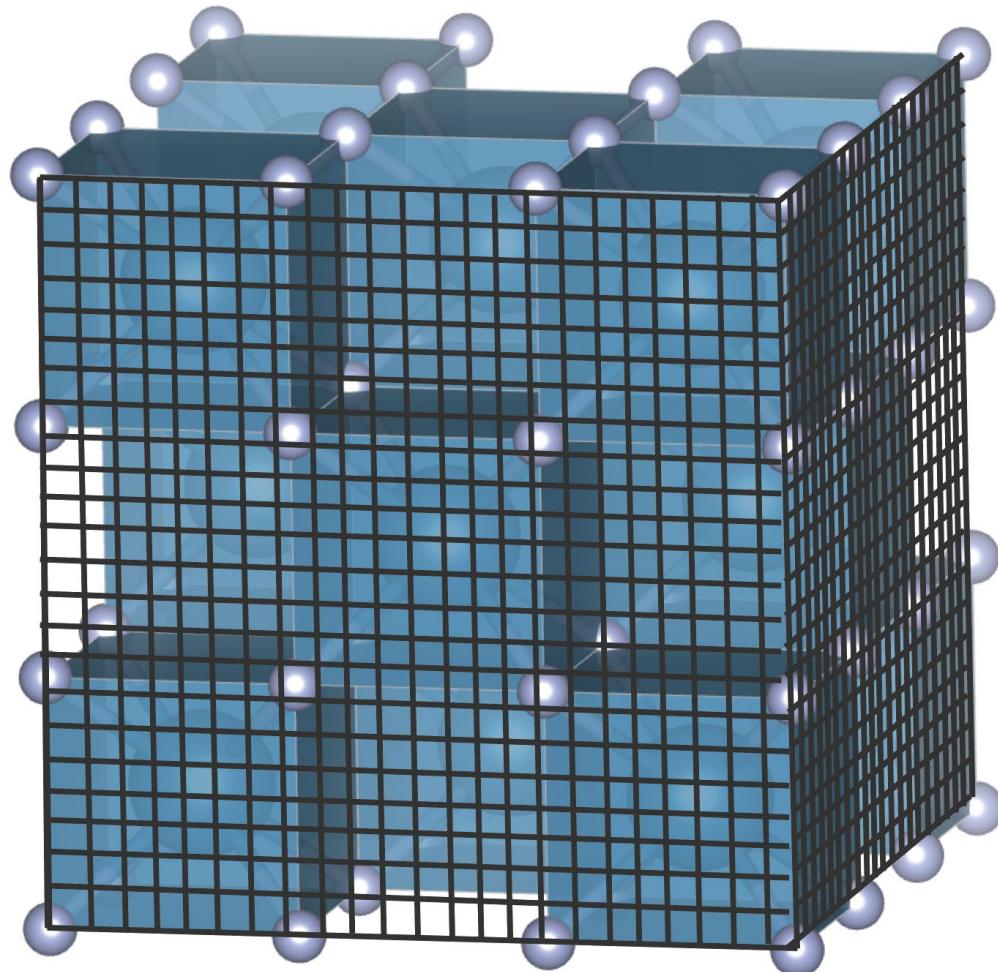
Structural feature	Ruddlesden-Popper	Aurivillius	Spinel	Rock Salt
Layered	1	1	0	0
Channels	0	0	1	0
RS blocks	1	0	0	1
Octahedra	1	1	1	1
Tetrahedra	0	0	1	0
...				
Corner-sharing	1	1	1	1
Edge-sharing	0	0	1	0

Maybe we could one-hot encode them?

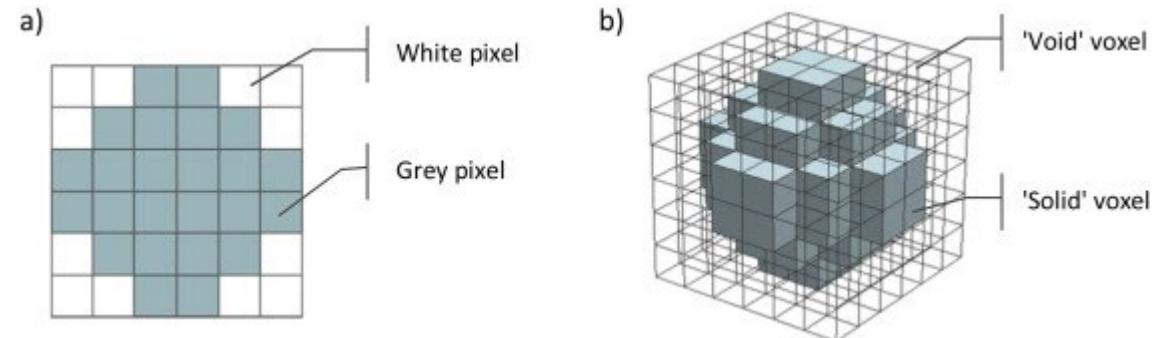
We could chop the crystal up into tiny voxels occupied or not



We could chop the crystal up into tiny voxels occupied or not

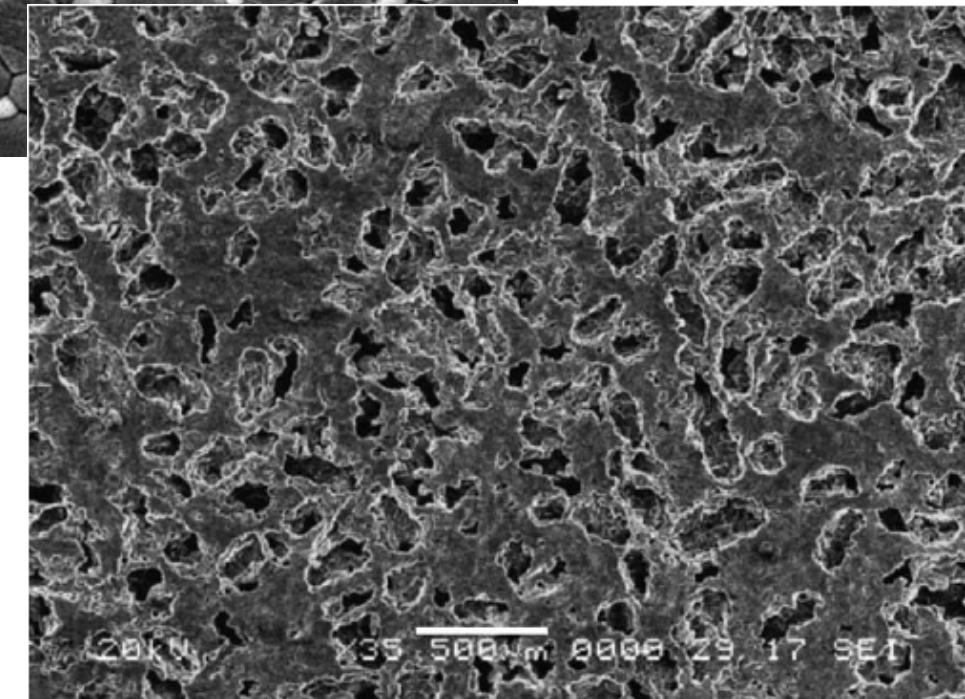
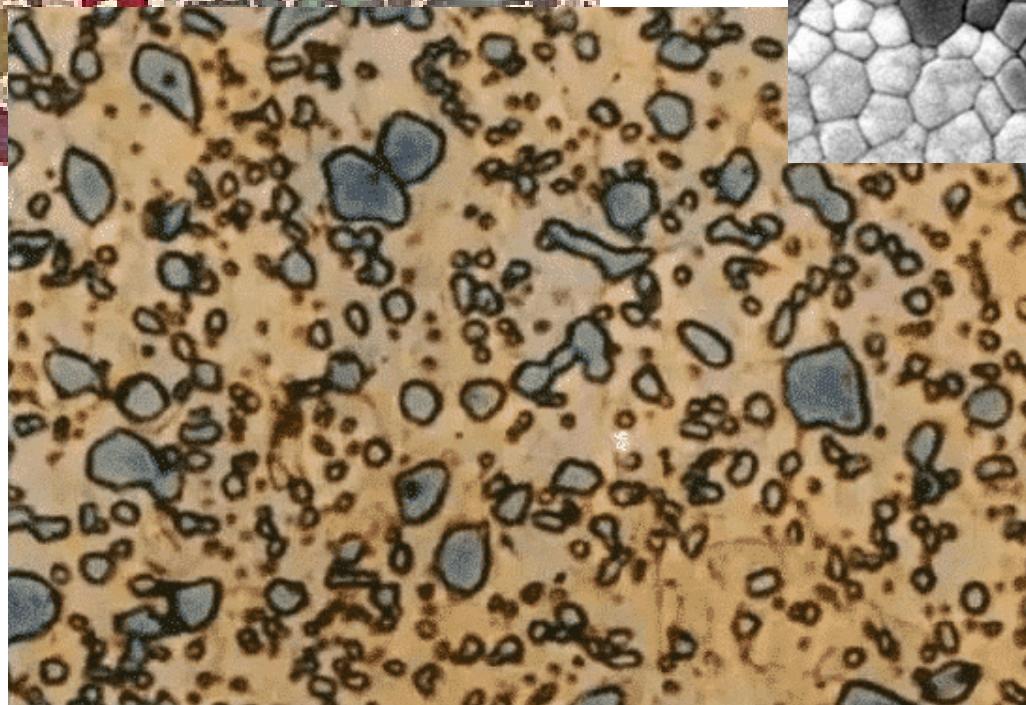
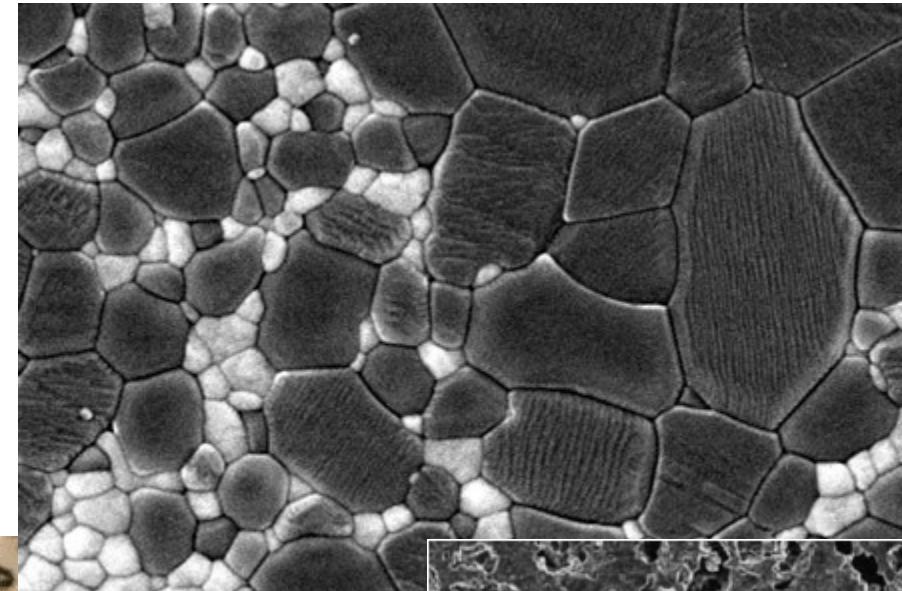
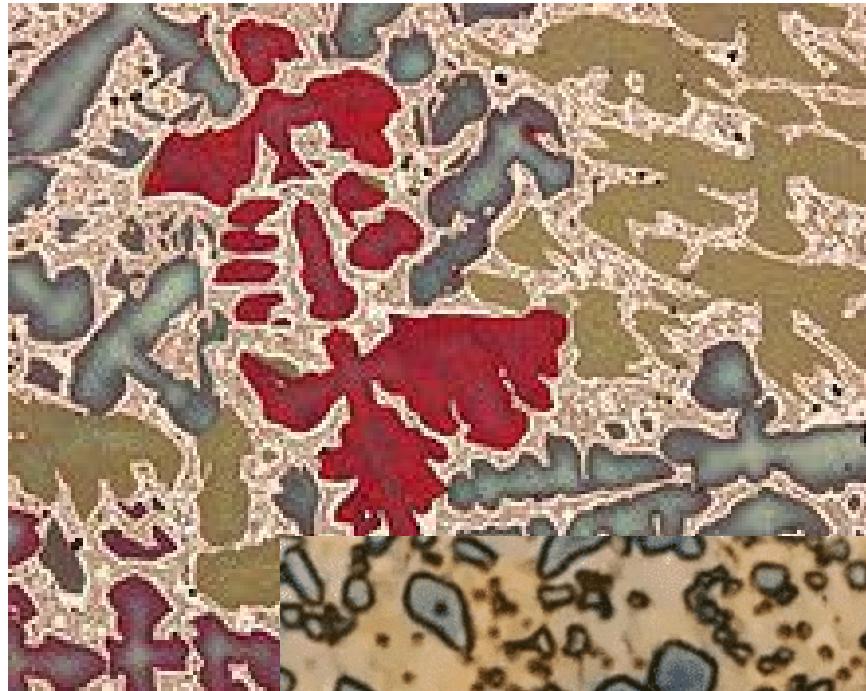


Do we put a point at atom center, or do we try to account for atom size?
How do we encode different atom chemistries?



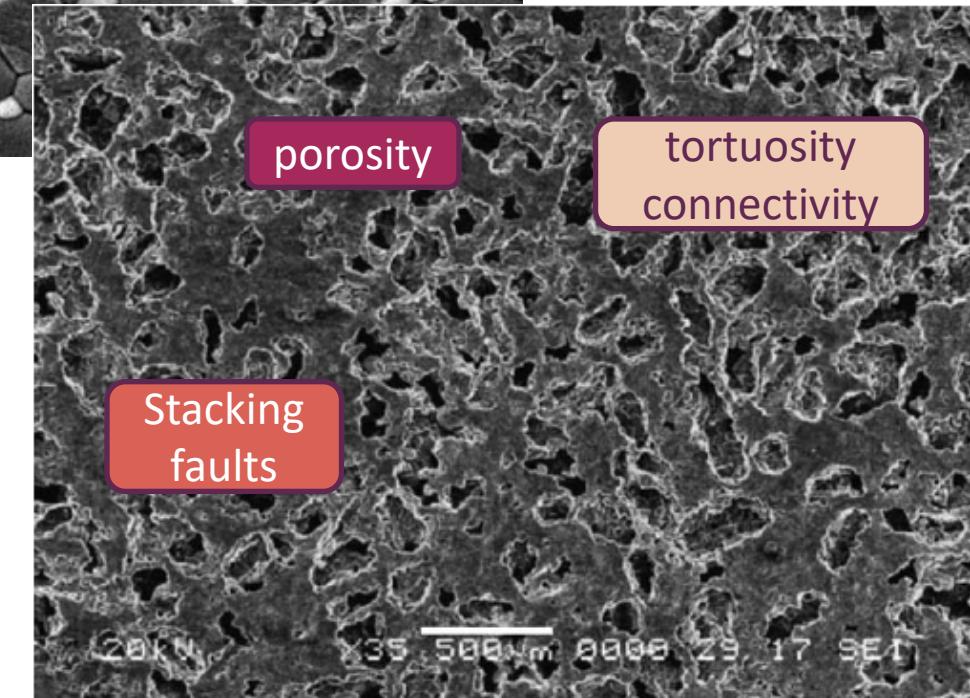
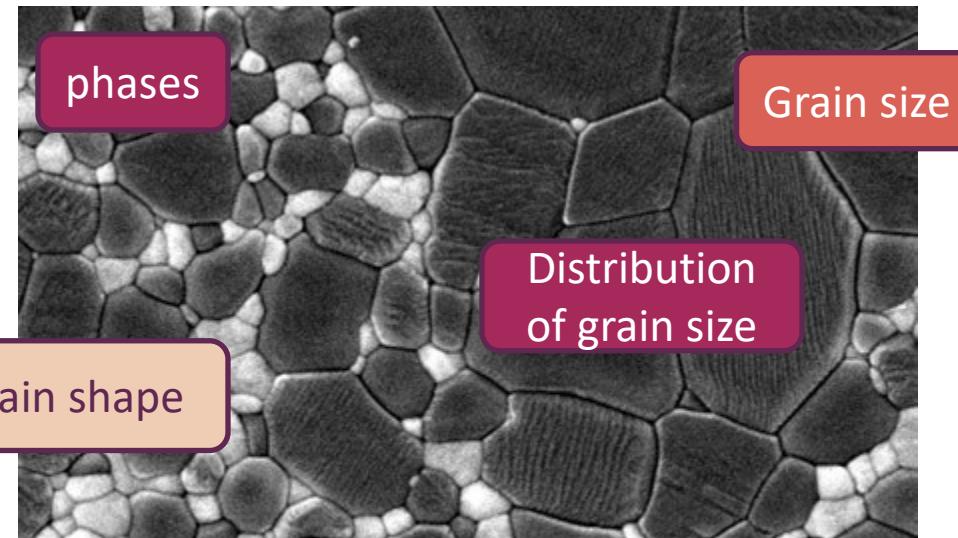
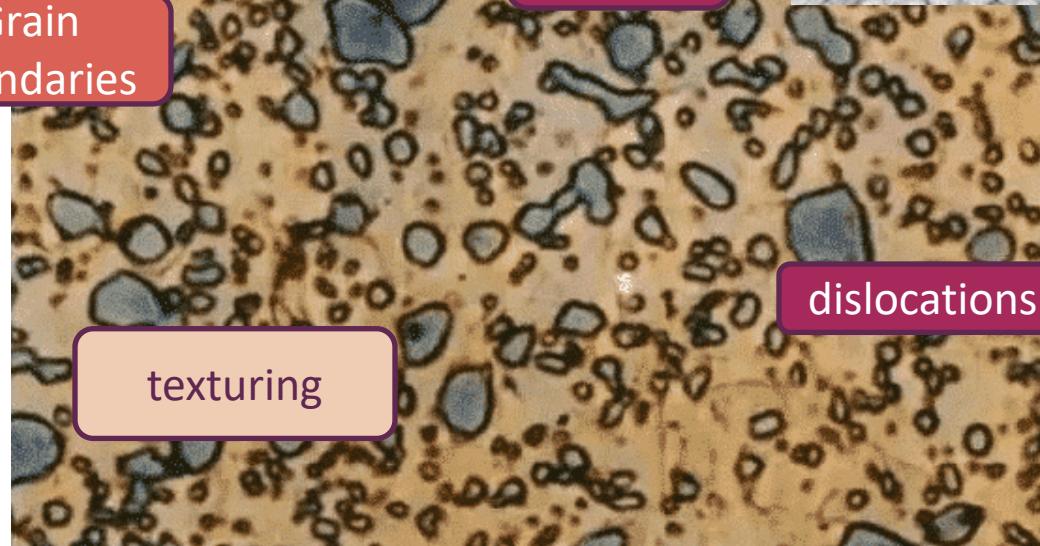
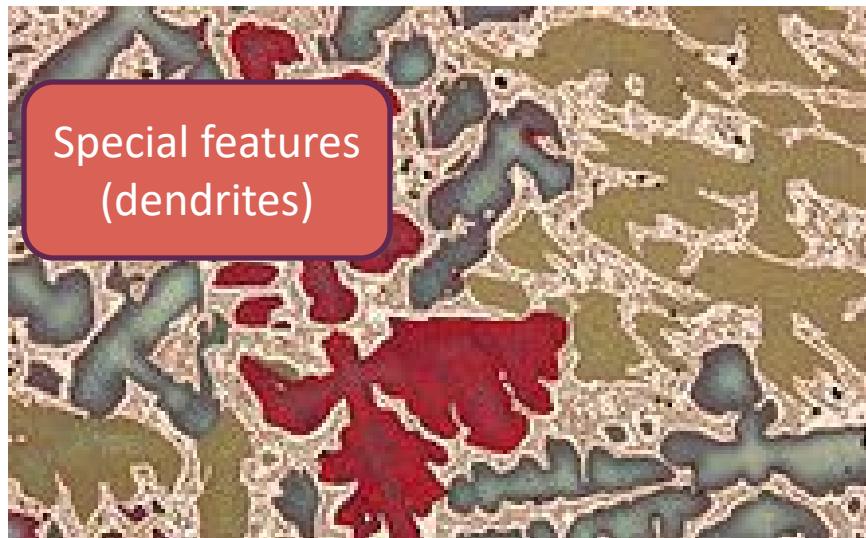


What information is present at the microstructure scale?

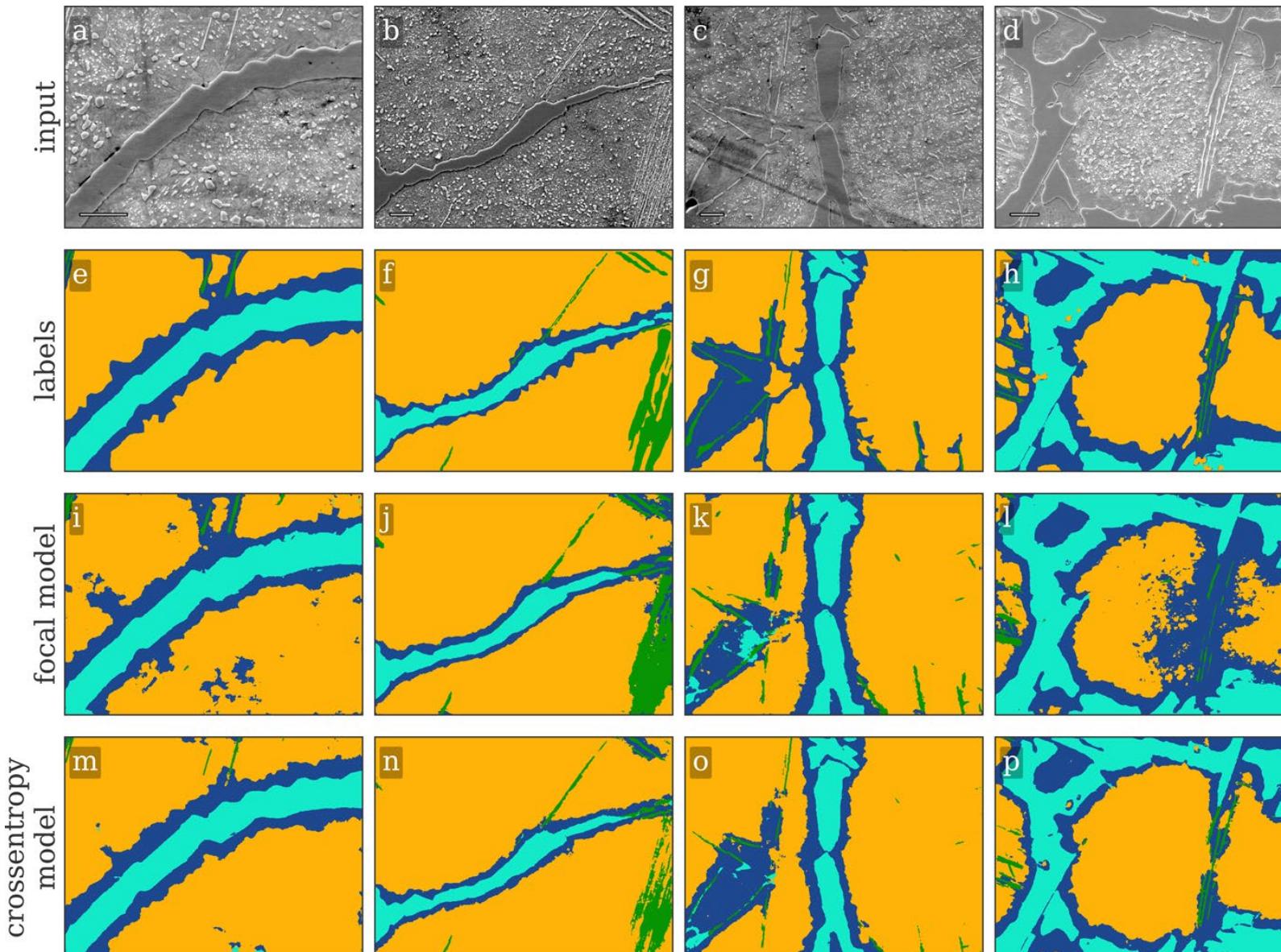




What information is present at the microstructure scale?



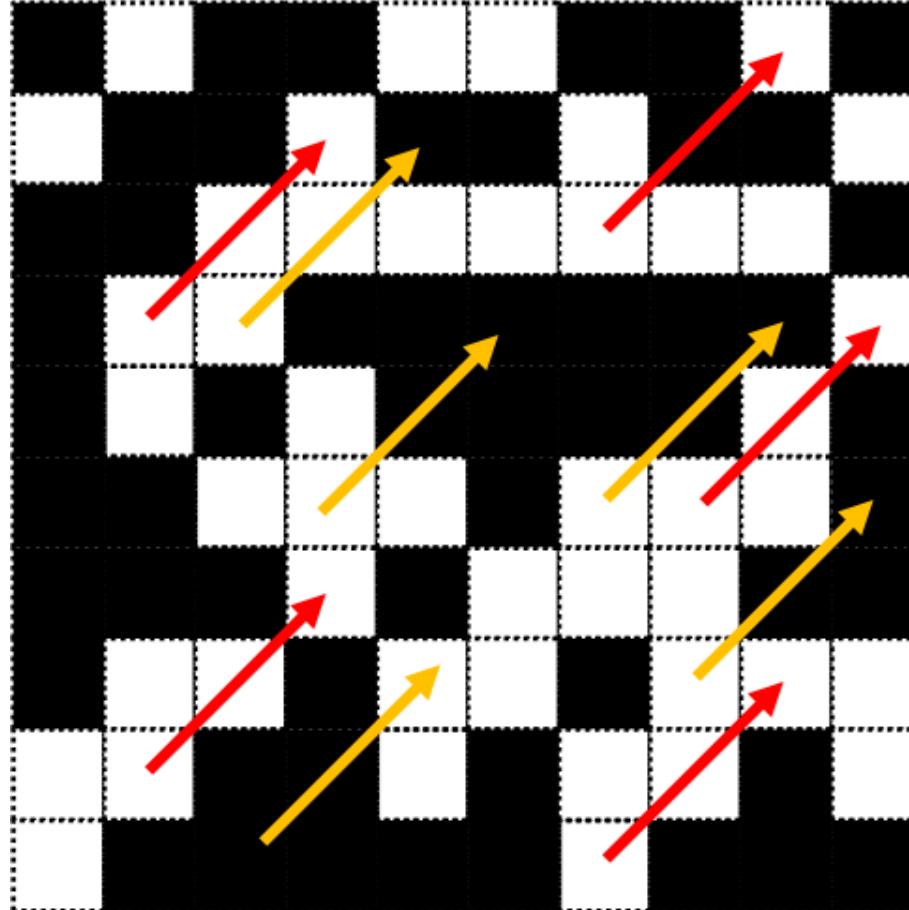
How do we identify and featurize these features?



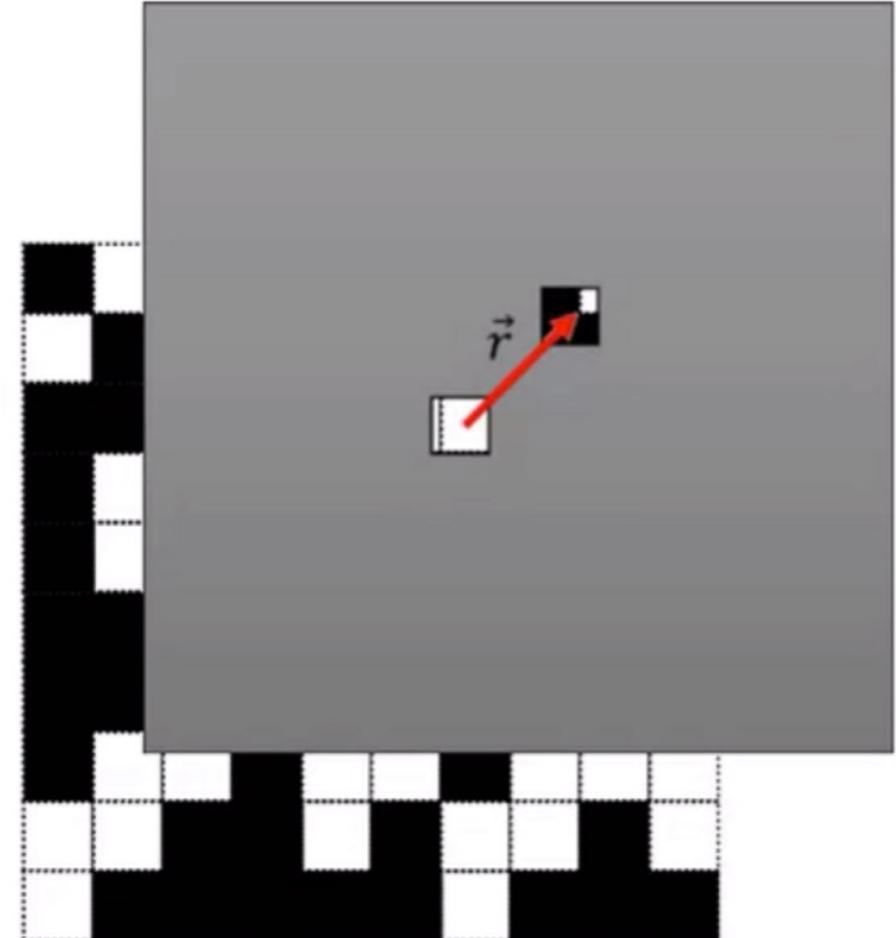


2-pt statistics is one powerful tool for quantifying these features

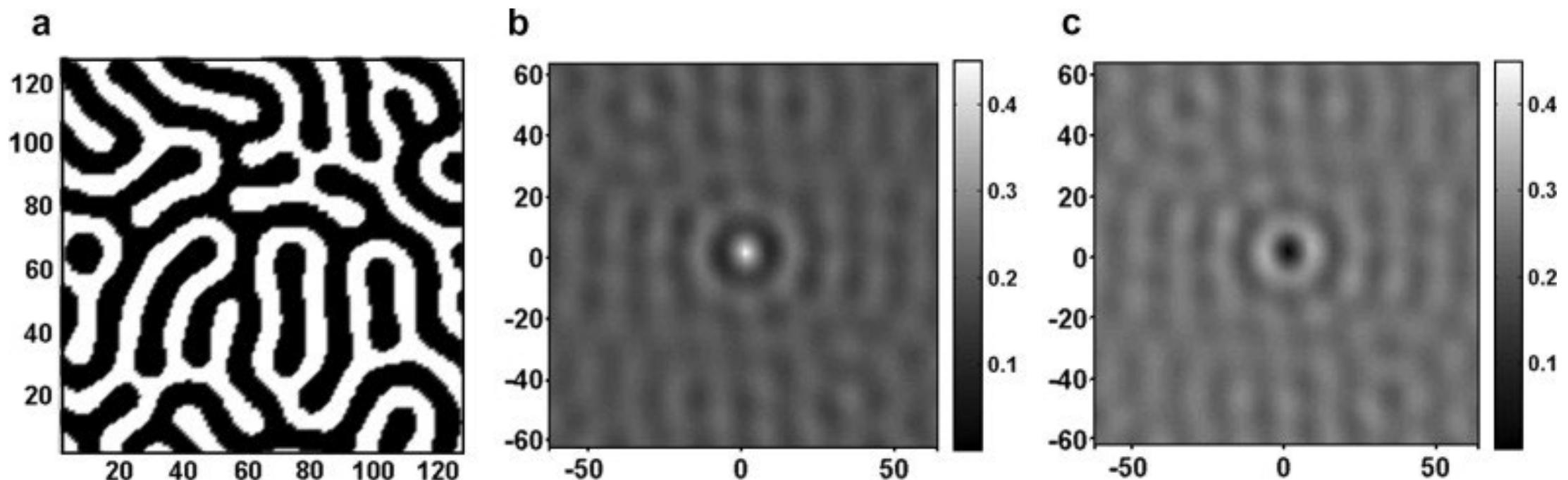
$$f_r^{hh'} = \frac{\# \text{Trials Successful}}{\# \text{Trials Attempted}}$$



$$P(A^h \cap A_{@\vec{r}}^{h'} | \vec{r}) = \frac{P\left(A^h \cap A_{@\vec{r}}^{h'} \cap \vec{r}\right)}{P(\vec{r})}$$



2-pt statistics is one powerful tool for quantifying these features

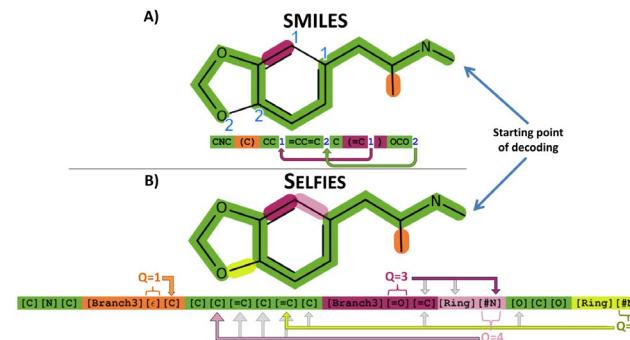


A number of different tools exist for structure featurization!

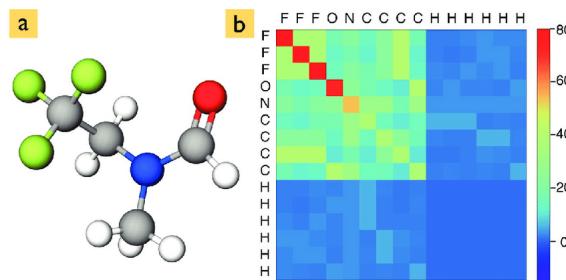
Microstructure segmentation/analysis



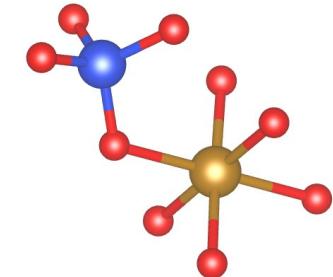
String representations of molecules



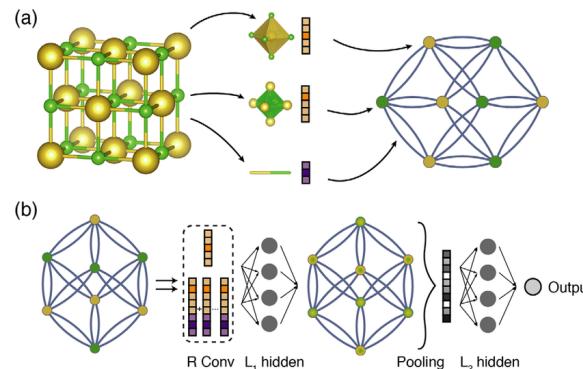
Coulomb matrix



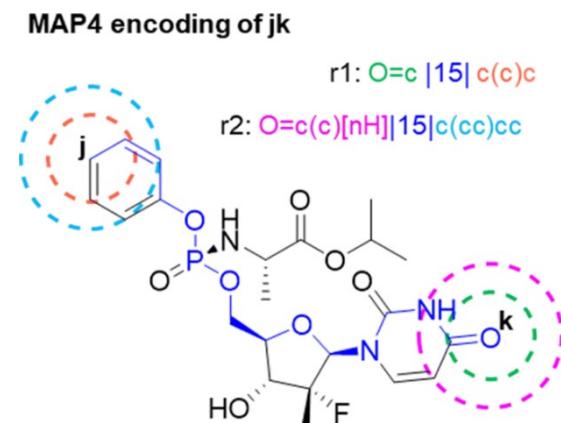
3D coordinates



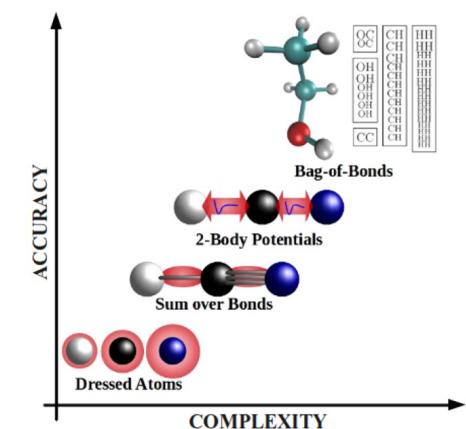
Crystal graphs



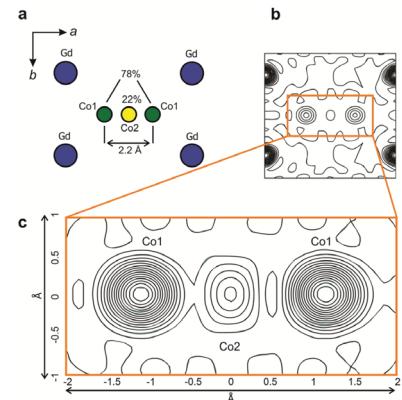
Fingerprints



Bag of Bonds / Fragments



Electron density



Crystal Structure graphs

