Imperial College London

19 March 2019

Updating the database of POCs

Andrew Tarzia

Imperial College London

Search process

- Extract all REFCODEs associated with list of 23 authors using **ConQuest** OR CSD Python API
- Filter by:
  - Has 3D coordinates
  - Not organometallic
  - Crystal structures only
  - Published post 2015
- Visualize in ConQuest and keep only 'cage-like' structures
- Produces 108 cage structures (CDB41 has 41 structures) although this may not be exhaustive

Imperial College London

Properties of crystal structures

- Visualize all structures and determine if:
  - Cage backbone is disordered
  - Solvent is present
  - Squeeze/Mask is used

	A	B	C	D	E	F	G	H	I	J	K	L
1	CIF name	REFCODE	Disorder in cage?	Solvent?	Squeeze/Mask?	DOI	Note on structure	Synonyms				
2	ABIKUU.cif	ABIKUU	N	Y	N	10.1038/nchem.2663		TCC1-R methanol solvate hydrate				
3	ABILAB.cif	ABILAB	N	Y	N	10.1038/nchem.2663		rac-TCC1 diethyl ether solvate				
4	ABILEF.cif	ABILEF	Y	N	N	10.1038/nchem.2663		TCC2-R bis(CC3-S)				
5	ABILIJ.cif	ABILIJ	Y	Y	Y	10.1038/nchem.2663		TCC2-R bis(CC3-S) 1,4-dioxane dichloromethane solvate				

Imperial College London

Next steps

- Automatically clean up disorder
- Automatically remove solvent and extract cage molecules (pyWindow)
- The above process is also being applied to collect metal–organic cages from the CSD