



# How to analyse molecular geometry with Mogul

CCDC Virtual Workshop Summer 2021 – Session 1

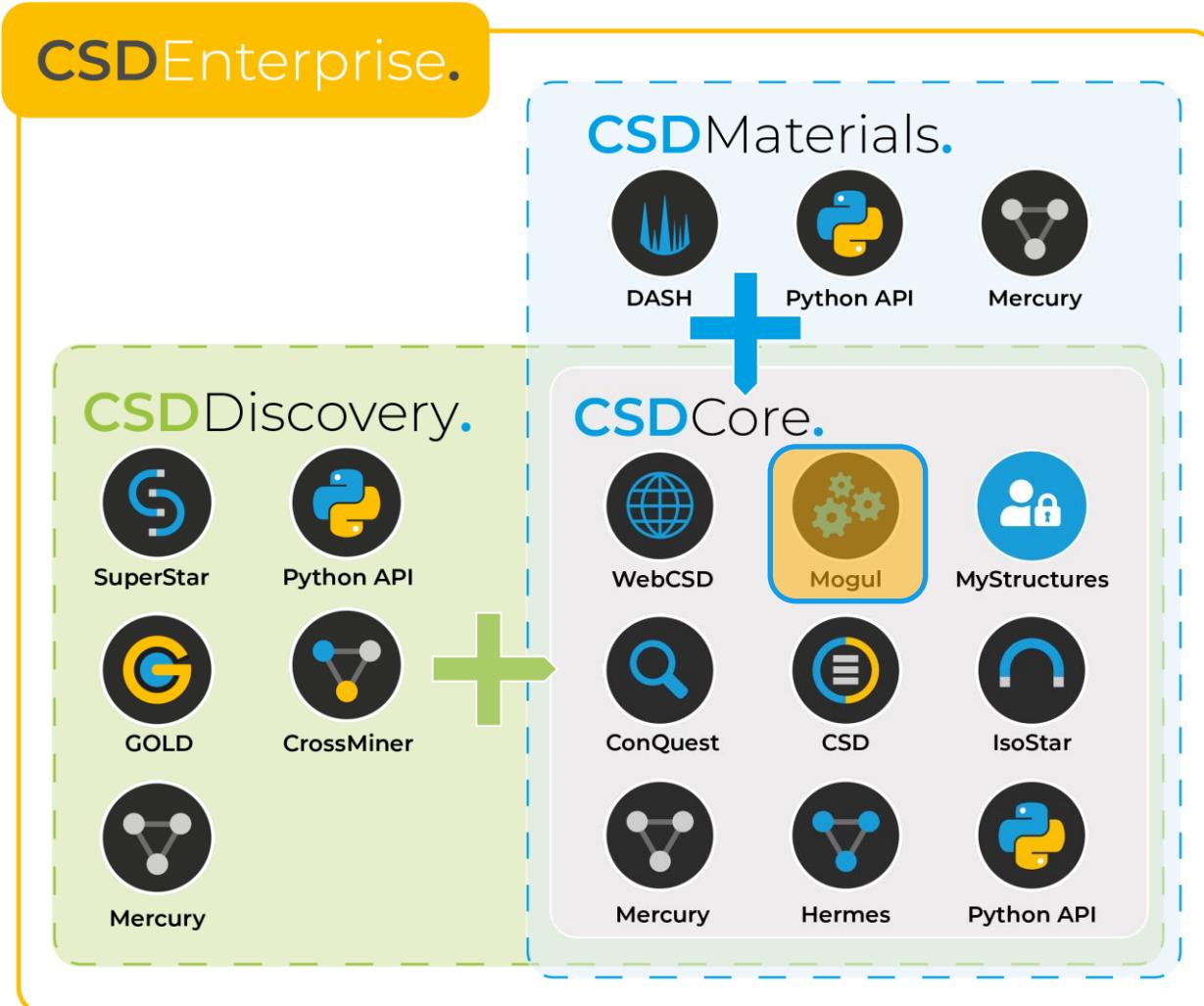
Suzanna Ward, Ilaria Gimondi, Lily Hunnisett,  
Natalie Johnson, Paul McKeown

July 2021

# Learning outcomes for today

- How to use the Mogul geometry check from Mercury to assess the geometry of a molecule.
- How to interpret the results obtained.
- How to launch Mogul from Mercury.
- How to run a geometry check on specific features (for example a torsional angle) of a molecule that you load.

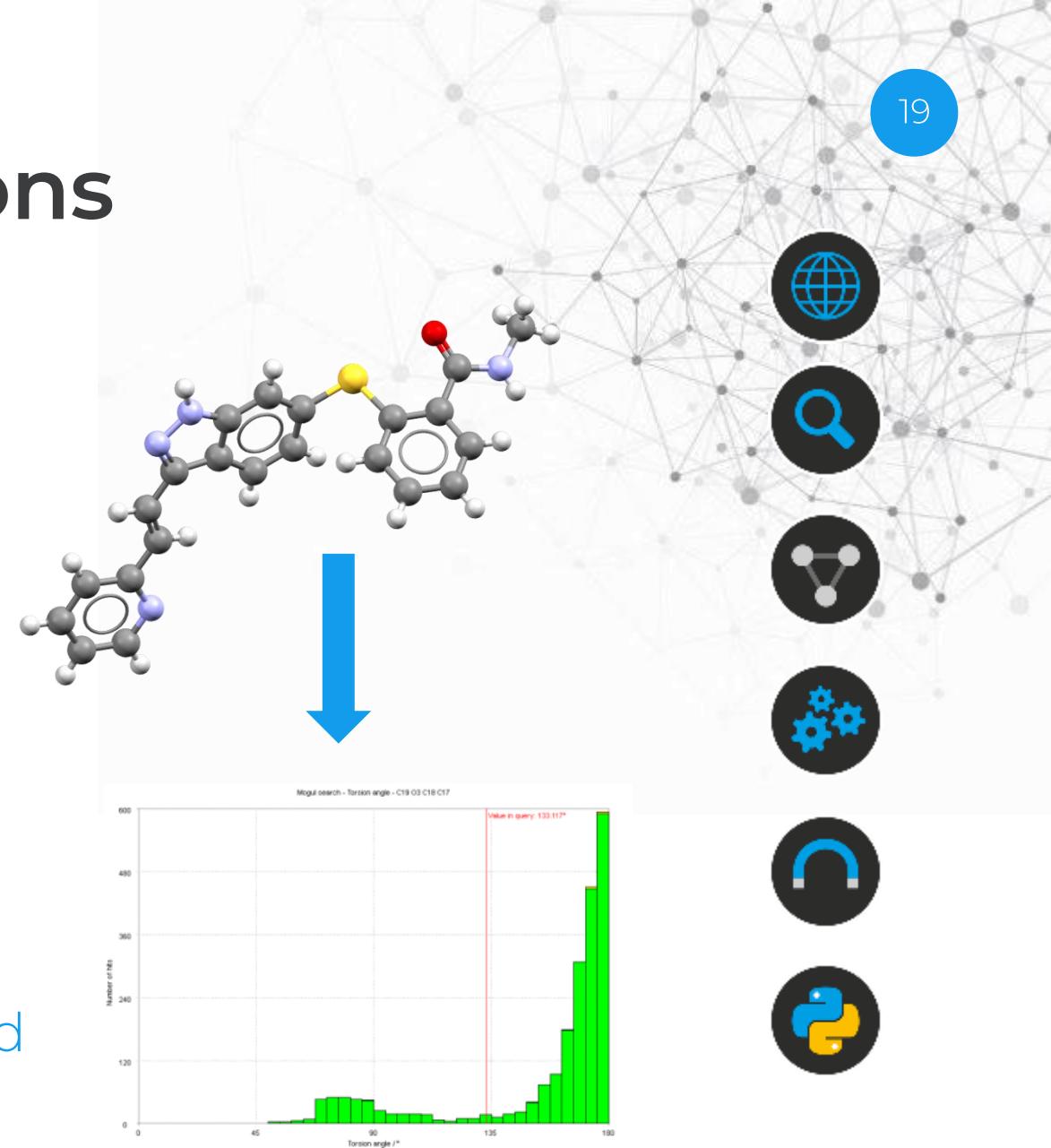
# The CSD software



CCDC

# CSD-Core – Conformations

- Mogul provides precise information on preferred molecular geometries
- Validate molecular geometries rapidly using interactive plots & statistics
- Identify any unusual features of a given query structure
- Mine millions of chemically classified bond lengths, angles, torsion angles and ring conformations in the CSD



CCDC

# Mogul

- Incorporates **pre-computed libraries** of bond lengths, valence angles, torsion angles and ring conformations **derived entirely from the CSD**
- **Validate complete geometry:** retrieve distributions, and figures of merit for all bonds, angles, torsions and rings in the molecule
- **Fragment Generalisation:** If fragment specified is rare, Mogul will include CSD results from the most similar fragments
- **Hyperlinking to the CSD:** View CSD entries in specific areas of histogram
- **Integration with other software:** Instruction file permits other programs to interact with Mogul

# When could we use Mogul?

- In teaching to help us understand fundamental chemistry
- To help design and assess new materials such as drug molecules
- To validate conformations
  - To analyse calculated conformations
  - To filter out protein-ligand docking solutions with unlikely ligand conformations
- To validate geometries
  - To check molecular dimensions of new crystal structures
- To create restraint data/ligand dictionaries
  - For protein refinement
  - To guide small molecule structure solution from powder diffraction data

# Using Mogul to teach chemistry

[A2-introduction to "Access Structures"](#)

Nuffield Foundation Newcastle University

The full CSD database, maintained by the Cambridge Crystallographic Data Centre, contains over 800,000 real crystal structures. The full CSD software requires the purchase of an annual site licence, but individual structures may be viewed and manipulated free of charge using the "Access Structures" option on the CCDC website. It uses a JSView viewer, thus enabling them to be viewed on both Windows and Apple devices.

## Getting Started

- To use the CSD "Access Structures" function you will require a computer or laptop or tablet with access to the internet.
- You will then need to find CSD "Access Structures" page using one of these two methods:
  - Type in the web address <https://www.ccdc.cam.ac.uk/structures>.
  - Type into a search engine such as Google "csd access structures" and click on the link that says "Access Structures".

## Basics



To view a particular structure, you need to type its refcode (provided) into the 'CSD refcode' box. For the purpose of this demo, type in the refcode 'CAFINE' for the crystal structure of caffeine.

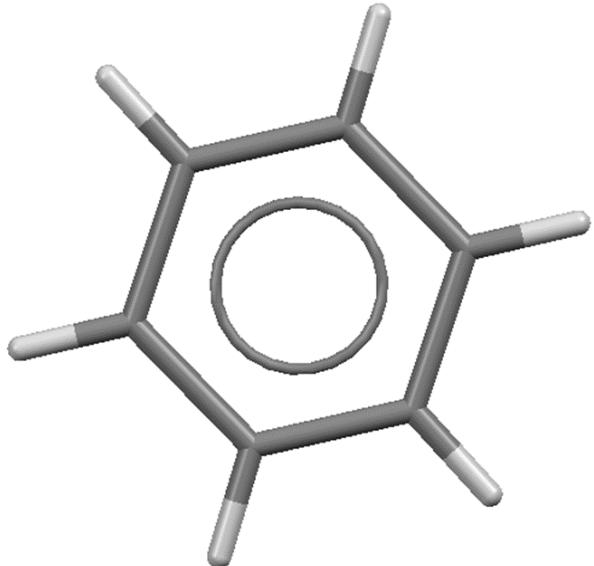
With the mouse you can hold left click to manually rotate the molecule and zoom using the middle button.

The drop-down headed 'Style' can be used to change the display.

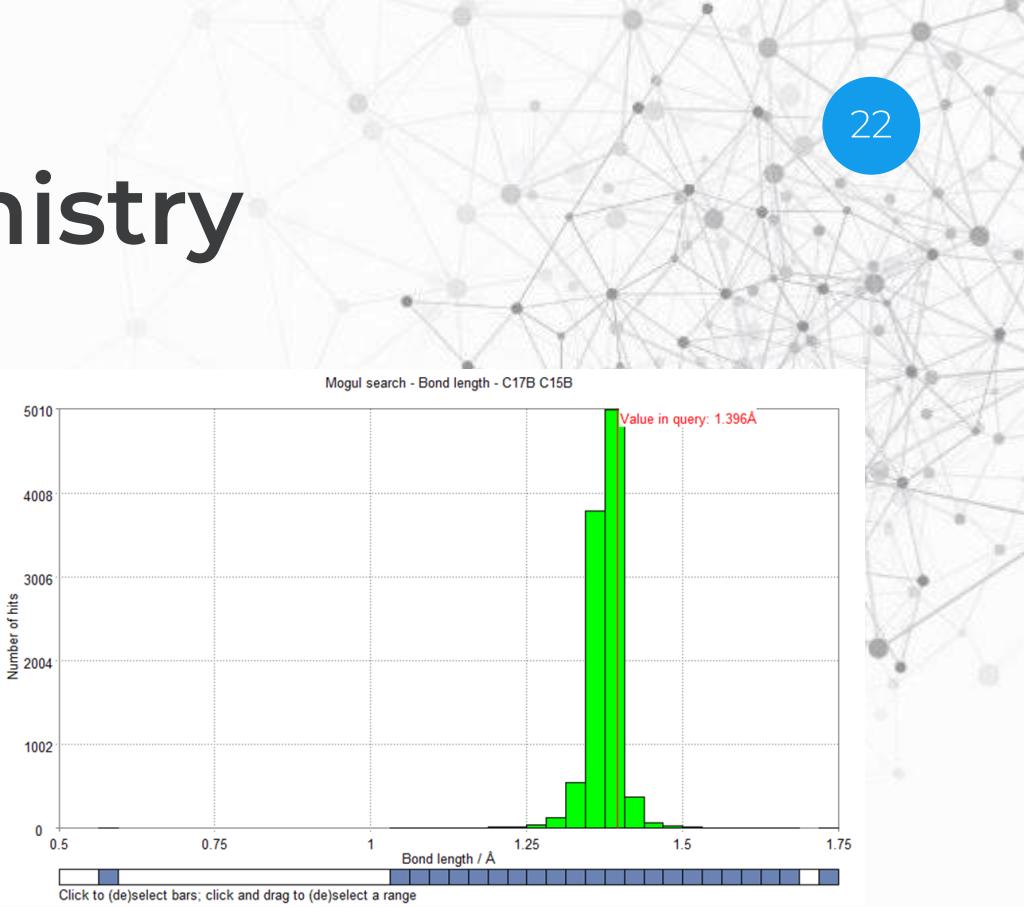
The 'Labels' drop-down can be used to label different atoms in the molecule.

'Packing' options can be used to look at the larger crystal structure, e.g. the unit cell and the bulk structure of 3x3x3 unit cells.

**TOP TIP!**  
When looking at some crystal structures on the database there will be two species on a refcode, you may only need to look at one of them. The two species are ions; one is a positive ion and the other is a negative ion. It is of course not possible to have one ion without the opposite 'counter ion'.

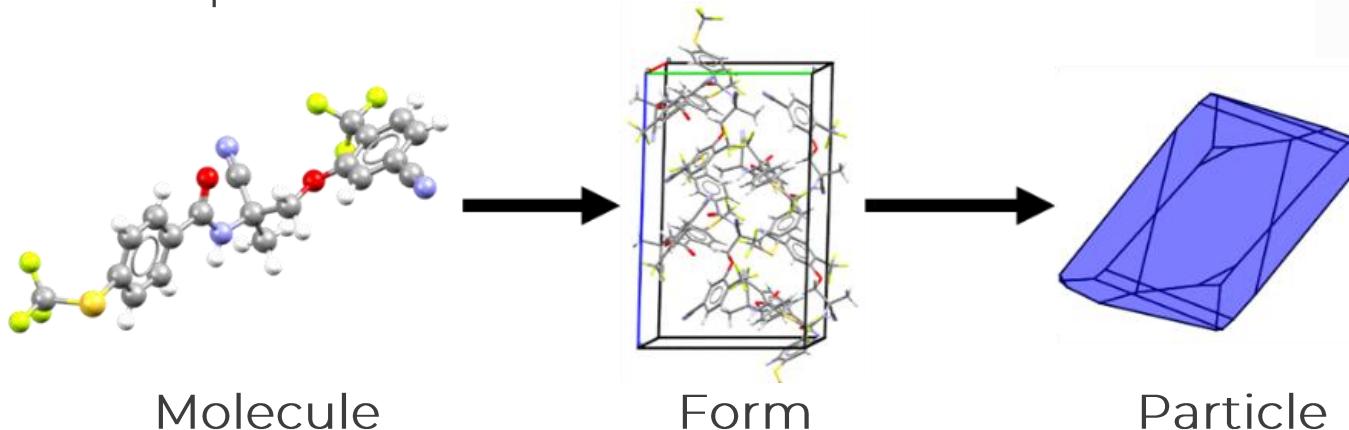


- Distance is about  $1.38\text{\AA}$
- Taken from over 450,000 benzene rings in reported in the CSD



# Using Mogul to inform key decisions

- The term “solid form informatics” was first introduced in mid-2000s
  - Use of structural knowledge to inform key decisions in pharmaceutical development*
- Now a key part of the solid form development workflow at most major pharmaceutical companies



Robert Docherty et al *Journal of Pharmacy and Pharmacology*, (2015) 67, 857. doi:10.1111/jphp.12394.  
 Taylor et al. *J. Chem. Inf. Model.*, (2014) 54 (9), 2500. Wood, P. A. et al. *CrystEngComm* (2013) 15, 65

**JPP** JOURNAL OF  
Pharmacy and Pharmacology

**JPP**  
Journal of Pharmacy  
And Pharmacology

**Research Paper**

**The integration of solid-form informatics into solid-form selection**

Neil Feeder<sup>a</sup>, Elna Pidcock<sup>a</sup>, Anthony M. Reilly<sup>a</sup>, Ghazala Sadiq<sup>a</sup>, Cheryl L. Doherty<sup>b</sup>, Kevin R. Back<sup>b</sup>, Paul Meenan<sup>a</sup> and Robert Docherty<sup>b</sup>

<sup>a</sup>The Cambridge Crystallographic Data Centre, Cambridge, <sup>b</sup>Pharmaceutical Science, Pfizer Global R&D, Sandwich, UK and <sup>c</sup>Pharmaceutical Science, Pfizer Global R&D, Groton, USA

**Keywords**  
Cambridge Structural Database; quality-by-design; solid-form informatics; solid-form selection

**Correspondence**  
Robert Docherty, Pharmaceutical Science, Pfizer Global R&D, Sandwich CT13 9NJ,

**Abstract**  
Objectives To demonstrate how the use of structural informatics during drug development assists with the assessment of the risk of polymorphism and the selection of a commercial solid form.  
Methods The application of structural chemistry knowledge derived from the hundreds of thousands of crystal structures contained in the Cambridge Structural Database to the rational design of pharmaceutical solid forms. This work shows the comparison of the use of Full hydrogen bonding software tools and the Solid Form Informatics software tool.

**CrystEngComm**

**PAPER**

**View Article Online**

**Knowledge-based approaches to co-crystal design†**

Cite this: CrystEngComm, 2014, 16, 5839  
Peter A. Wood,\* Neil Feeder, Matthew Furlow, Peter T. A. Galek, Colin R. Groom and Elna Pidcock

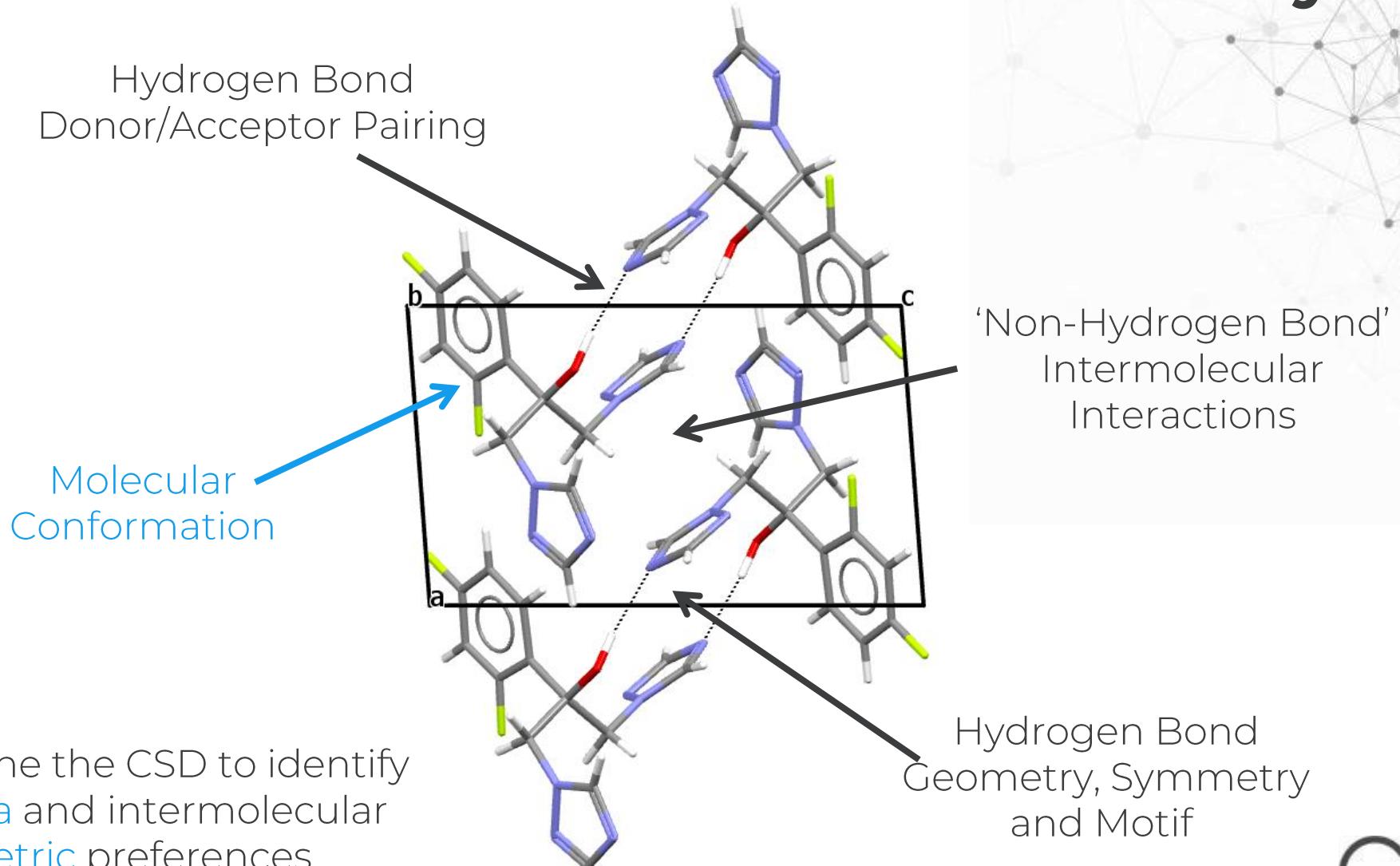
Received 12th February 2014,  
Accepted 10th March 2014  
DOI: 10.1039/c4ce00316k  
[www.rsc.org/crystengcomm](http://www.rsc.org/crystengcomm)

**1. Introduction**  
The definition of a co-crystal,<sup>1,2</sup> whether this is the correct term to use<sup>3</sup> and even whether it should contain a hyphen has been debated at length in the literature. At the most basic level, a co-crystal is a crystalline form composed from two or more components with a particular stoichiometry. Generally, the definitions used in the Crystalline Engineering field also place some or all of the following requirements on the solid forms in question:  
 1. All components are organic species (ruling out inorganics or organometallics).  
 2. None of the components are charged (otherwise classified as salts).  
 3. None of the components are water (otherwise classified as hydrates).

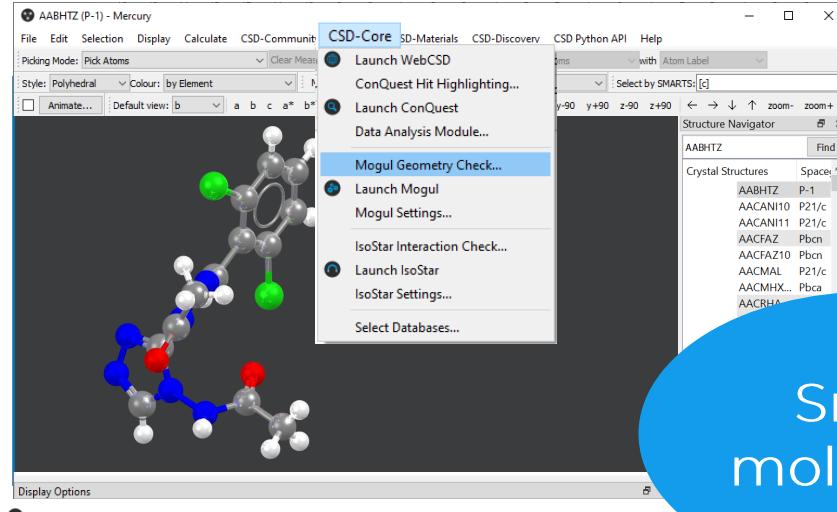
The use of knowledge-based methods has been intimately connected with the field of co-crystal design since the seminal papers of Etter and Desiraju in the 1990s. Here we explain and exemplify how rational co-crystal design has been carried out in the past using crystal structure knowledge as well as presenting emerging methodologies for knowledge-based co-former selection.

CCDC

# Characteristics that influence stability



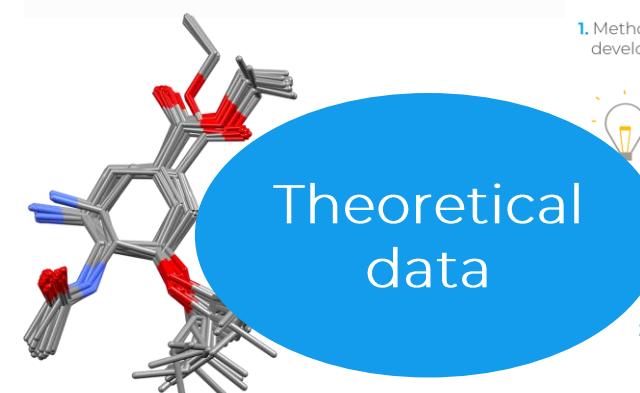
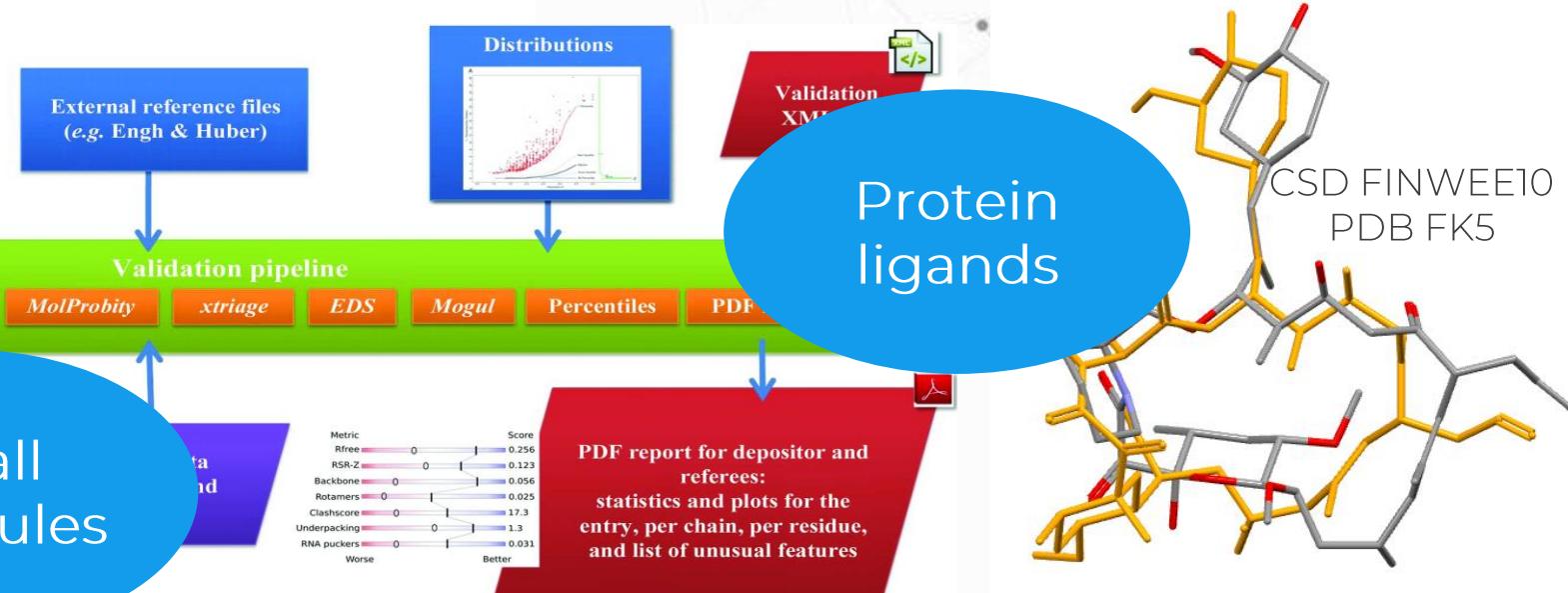
# Using Mogul to check new structural data



Small molecules

A screenshot of the CCDC Mogul software interface. The top navigation bar includes File, Searches, Databases, and Help. The 'Build query' tab is selected. The results navigator shows 'All hits: 137' and 'Any: 137'. The classification column includes 'Not unusual (enough hits)', 'Unusual (enough hits)', and 'Unusual (few hits)'. The table lists various torsion angles and their hit counts. To the right, there are two histograms: one for 'Mogul search - Torsion angle - C21 C22 C24 N27' showing a distribution of torsion angles from 0 to 180 degrees, and another histogram for 'Histogram display' showing the number of hits versus a range of values.

Type	Molecule	Fragment	Classification	No. of hits	Query value
> bond					
> angle					
✓ torsion	LIM_K21_901-B_pdb1hak_1				
	C59 O58 C1 C2		Not unusual (enough hits)	13461	169.071
	C59 O58 C1 C6		Not unusual (enough hits)	13461	-10.086
	O23 C21 N11 C12		Not unusual (enough hits)	45	11.263
	C24 C22 C21 N11		Not unusual (enough hits)	40	-75.917
	C22 C24 N27 C30		Not unusual (enough hits)	938	-155.722
	C22 C24 N27 C34		Not unusual (enough hits)	938	76.119
	C45 C52 C43 C32		Not unusual (enough hits)	2357	-71.842
	C51 C52 C43 C32		Not unusual (enough hits)	2357	105.277
	C22 C21 N11 C12		Not unusual (few hits)	3	-165.684
	O23 C21 N11 C10		Unusual (enough hits)	45	-158.238
	O23 C21 C22 C24		Unusual (enough hits)	42	107.065
	C21 C22 C24 N27		Unusual (enough hits)	137	-22.055
	C31 C32 C43 C52		Unusual (enough hits)	131	-140.163
	C33 C32 C43 C52		Unusual (enough hits)	131	97.369
	C22 C21 N11 C10		Unusual (few hits)	3	24.815



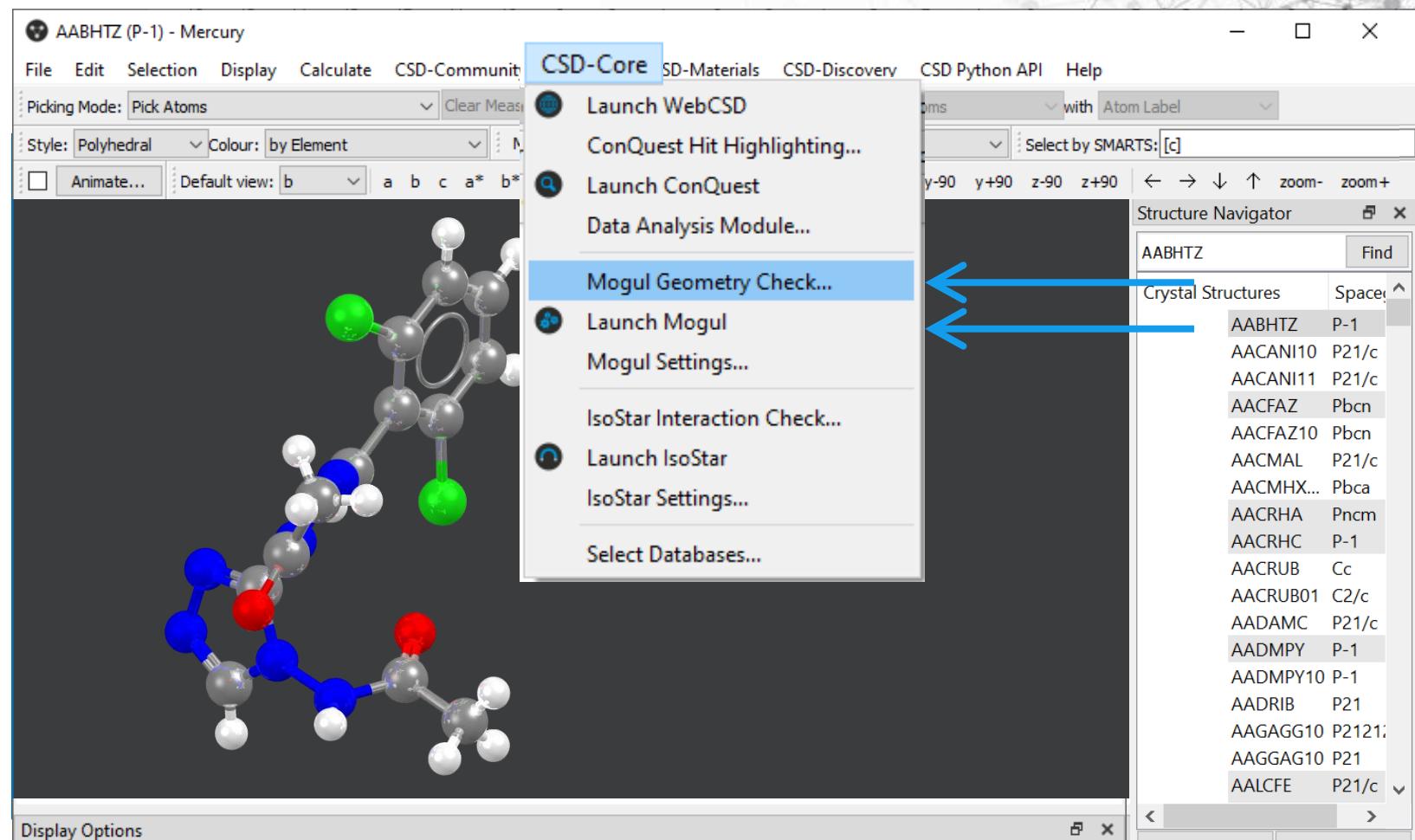
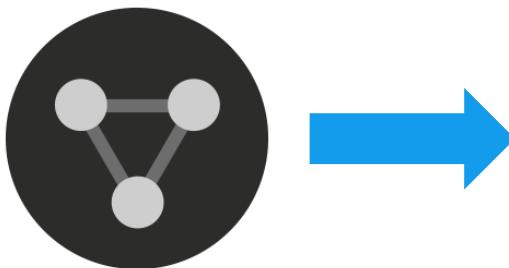
Theoretical data

1. Methods development.
  2. Test methods blind with unpublished targets.
  3. Compare predictions to experimental observations.
- Crystal Structure Prediction (CSP) Blind Test.

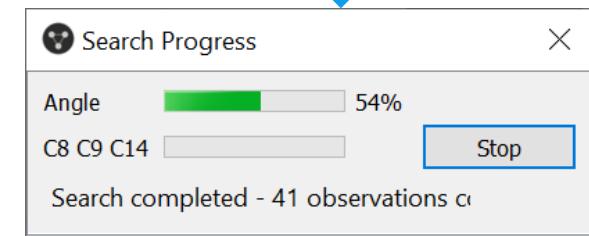
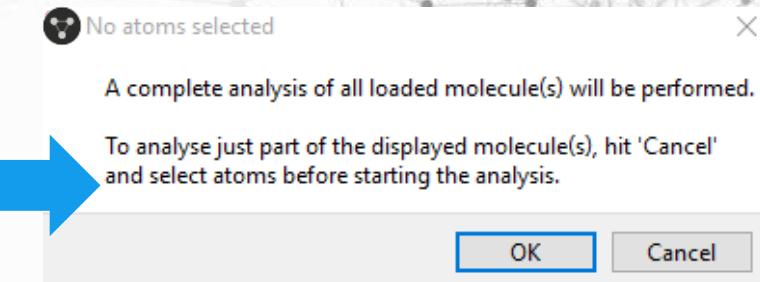
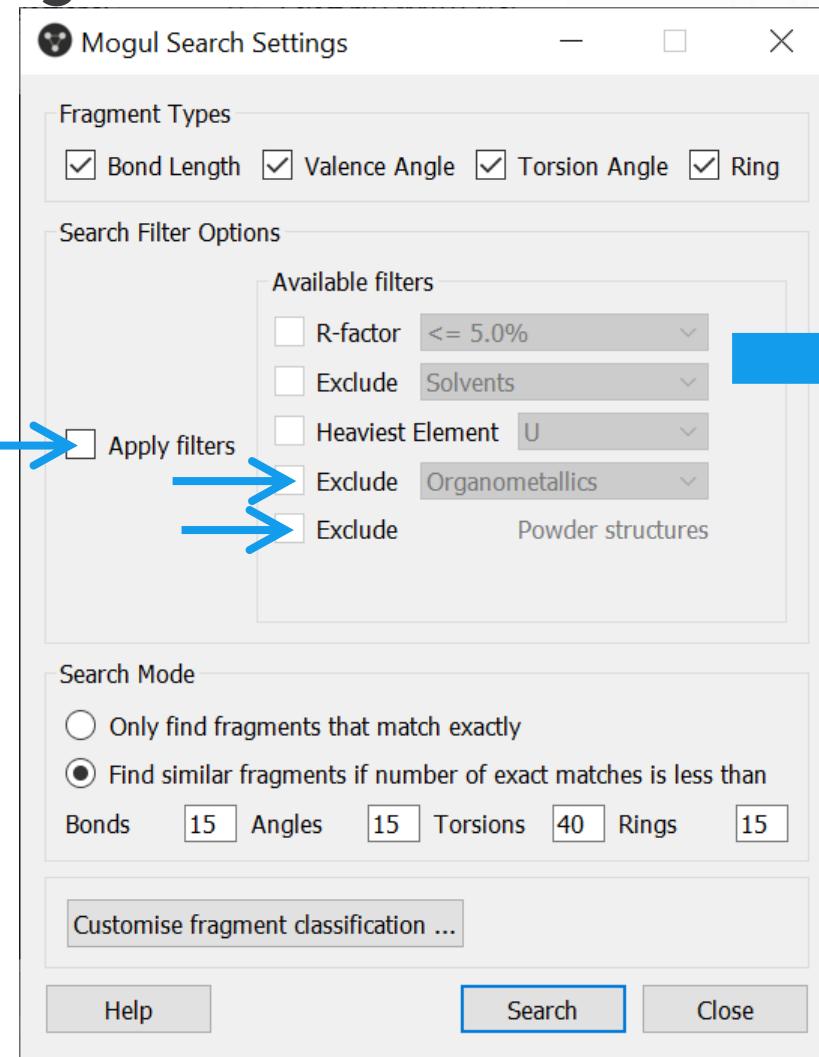
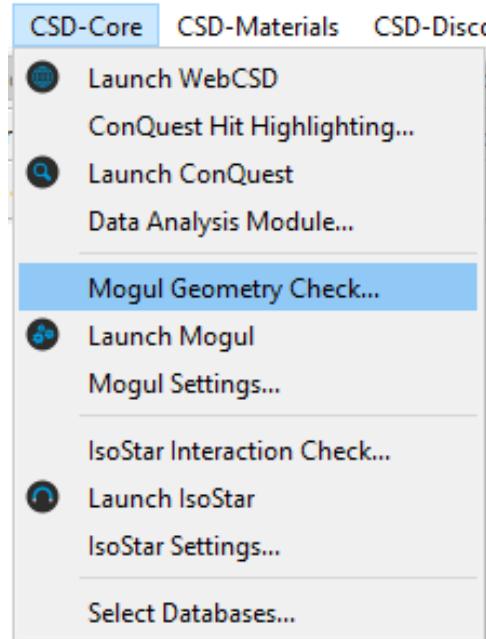
# Chemical coverage - the details....

- Large distributions are reduced by random selection.
- Bond fragments and rings involving metals are characterised in exactly the same way as organic fragments
- The following are not included:
  - Bonds, angles & torsions involving H atoms
  - Valence angles & torsions involving metals
  - Cyclic torsions of rings less than 8 atoms
  - Rings containing less than 5 atoms (fused and bridged rings are included)

# Show One: Mogul – Launch from Mercury



# Mogul Geometry Check



# Mogul Geometry Check - Results

Mogul Results Viewer

Show / hide : Columns Fragments... Deselect all fragments Export...

Help Double click to view result in Mogul

Type	Molecule	Fragment	Classification	No. of hits	Query value	Mean
> bond						
> angle						
✓ torsion						
✓ LIM_K21_901-B_pdb1hak_1						
	C59 O58 C1 C2		Not unusual (enough hits)	13461	169.071	
	C59 O58 C1 C6		Not unusual (enough hits)	13461	-10.086	
	O23 C21 N11 C12		Not unusual (enough hits)	45	11.263	
	C24 C22 C21 N11		Not unusual (enough hits)	40	-75.917	
	C22 C24 N27 C30		Not unusual (enough hits)	938	-155.722	
	C22 C24 N27 C34		Not unusual (enough hits)	938	76.119	
	C45 C52 C43 C32		Not unusual (enough hits)	2357	-71.842	
	C51 C52 C43 C32		Not unusual (enough hits)	2357	105.277	
	C22 C21 N11 C12		Not unusual (few hits)	3	-165.684	
	O23 C21 N11 C10		Unusual (enough hits)	45	-158.238	
	O23 C21 C22 C24		Unusual (enough hits)	42	107.065	
	C21 C22 C24 N27		Unusual (enough hits)	137	-22.055	
	C31 C32 C43 C52		Unusual (enough hits)	131	-140.163	
	C33 C32 C43 C52		Unusual (enough hits)	131	97.369	
	C22 C21 N11 C10		Unusual (few hits)	3	24.815	

## Colour code:

- **Blue**: It is **not unusual** value in the CSD and there are **enough data** in the CSD for comparison.
- **Pink**: It is **not unusual** value in the CSD, but there are only **few data** in the CSD for comparison.
- **Orange**: It is an **unusual** value in the CSD and there are **enough data** in the CSD for comparison.
- **Brown**: It is an **unusual** value in the CSD, but there are only **few data** in the CSD for comparison.

# Mogul Geometry Check - Results

Mogul Results Viewer

Show / hide : Columns Fragments... Deselect all fragments Export...

Help Double click to view result in Mogul

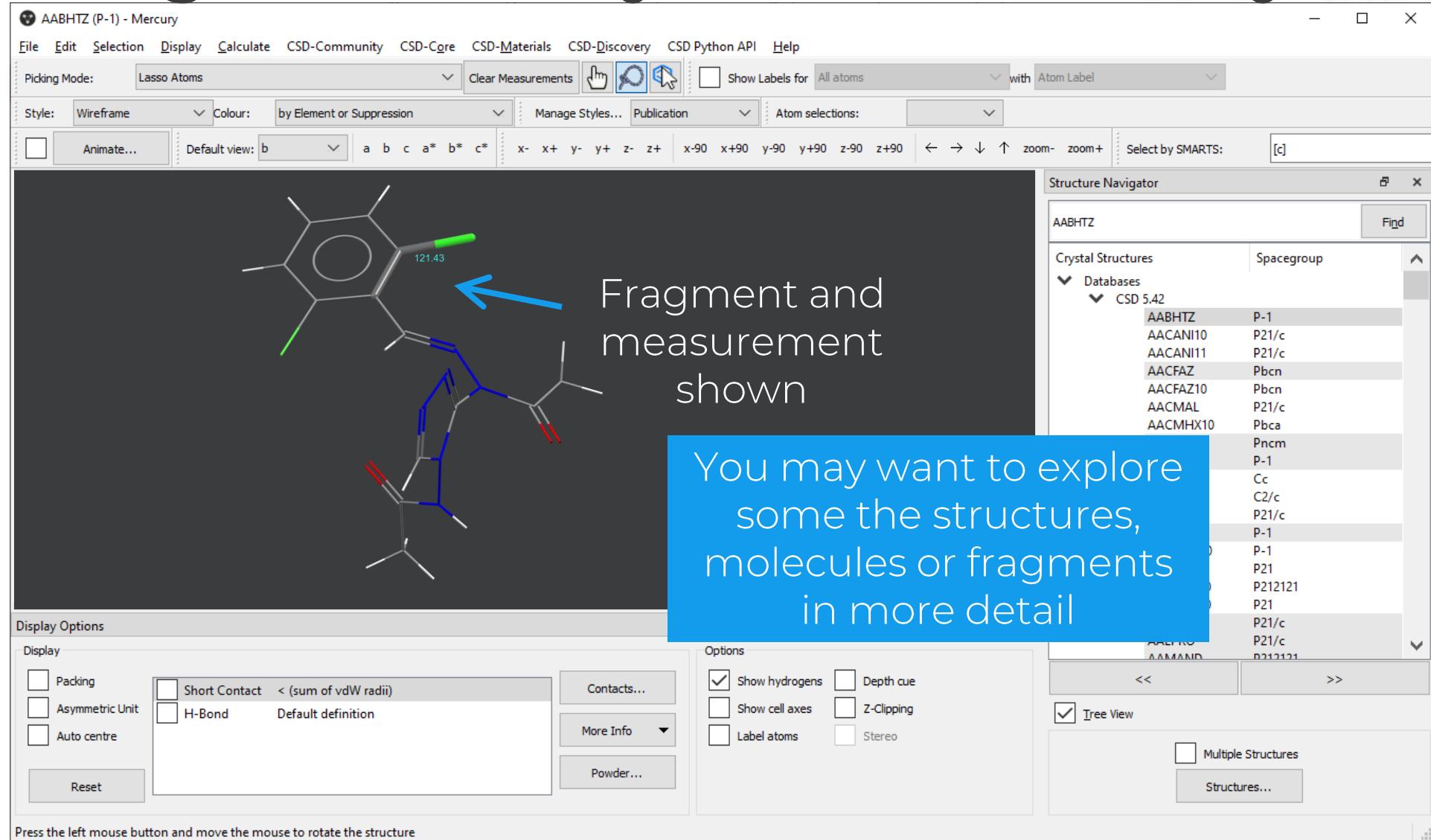
Type	Molecule	Fragment	Classification	No. of hits	Query value	Mean	Std. dev.	z-score	x - mean	Minimum	Maximum	Median	d(min)
✓ bond	AABHTZ												
➤ angle	AABHTZ	C2 C1 Cl1	Not unusual (enough hits)	10637	117.372	118.409	1.966	0.528	1.038	86.905	152.624	118.518	0.000
✓		C6 C1 Cl1	Not unusual (enough hits)	449	119.477	119.883	1.790	0.227	0.406	111.452	136.611	119.752	0.000
		C2 C1 C6	Not unusual (enough hits)	421	123.139	122.057	2.692	0.402	1.082	79.010	129.445	122.401	0.000
		C3 C2 C1	Not unusual (enough hits)	10491	119.515	119.405	1.501	0.073	0.110	89.047	141.490	119.478	0.000
		C4 C3 C2	Not unusual (enough hits)	20098	119.999	120.239	1.325	0.181	0.240	70.643	140.815	120.159	0.000
		C3 C4 C5	Not unusual (enough hits)	10491	119.854	119.405	1.501	0.299	0.449	89.047	141.490	119.478	0.000
		C4 C5 C12	Not unusual (enough hits)	10637	116.212	118.409	1.966	1.118	2.197	86.905	152.624	118.518	0.000
		C6 C5 C12	Not unusual (enough hits)	449	121.434	119.883	1.790	0.867	1.551	111.452	136.611	119.752	0.000
		C4 C5 C6	Not unusual (enough hits)	421	122.346	122.057	2.692	0.107	0.288	79.010	129.445	122.401	0.000
		C1 C6 C5	Not unusual (enough hits)	170	115.125	116.276	3.163	0.364	1.152	111.918	151.002	115.875	0.000
		C5 C6 C7	Not unusual (enough hits)	142	126.222	122.066	3.194	1.301	4.156	106.172	133.828	121.766	0.000
		C1 C6 C7	Not unusual (enough hits)	142	118.637	122.066	3.194	1.074	3.429	106.172	133.828	121.766	0.000
		C6 C7 N1	Not unusual (enough hits)	63	120.071	121.580	1.633	0.924	1.509	118.121	125.394	121.710	0.000
		C7 N1 N2	Not unusual (enough hits)	304	117.522	117.173	2.634	0.133	0.349	112.028	123.777	117.122	0.000
		C8 N2 N1	Not unusual (enough hits)	16	124.440	121.412	5.026	0.603	3.028	110.211	126.188	123.650	0.000
		C12 N2 N1	Not unusual (enough hits)	15	116.721	115.241	2.607	0.568	1.480	111.791	120.406	114.780	0.000
		N2 C8 N3	Not unusual (enough hits)	15	125.283	124.096	1.376	0.863	1.187	121.050	125.790	124.354	0.000
		N5 C8 N3	Not unusual (enough hits)	18	110.781	110.483	0.852	0.349	0.298	109.567	112.225	110.237	0.000
		C8 N3 N4	Not unusual (enough hits)	47	106.679	105.504	2.395	0.491	1.175	100.785	114.792	105.591	0.000
		C9 N4 N3	Not unusual (enough hits)	190	107.177	107.613	1.828	0.238	0.436	101.988	116.826	107.467	0.000
		N5 C9 N4	Not unusual (enough hits)	61	110.746	110.185	0.564	0.995	0.561	108.650	111.600	110.233	0.000
		C8 N5 N6	Not unusual (enough hits)	15	128.047	125.894	2.689	0.801	2.154	118.623	129.059	126.475	0.000
		C9 N5 C8	Not unusual (enough hits)	15	104.615	105.953	1.319	1.015	1.338	103.742	108.706	105.802	0.000
		C9 N5 N6	Not unusual (enough hits)	47	126.994	126.891	1.763	0.058	0.102	121.092	133.246	127.067	0.000
		C10 N6 N5	Not unusual (enough hits)	15	117.419	118.913	1.088	1.373	1.494	117.419	120.816	118.696	0.000
		C11 C10 N6	Not unusual (enough hits)	37	114.417	114.933	1.651	0.312	0.515	112.356	120.514	114.515	0.000
		O1 C10 N6	Not unusual (enough hits)	37	121.348	121.538	1.427	0.133	0.189	116.489	123.542	121.784	0.000
		O1 C10 C11	Not unusual (enough hits)	4907	124.224	122.048	1.806	1.205	2.177	91.903	161.131	121.976	0.000
		C13 C12 N2	Not unusual (enough hits)	34	117.582	116.154	0.932	1.532	1.428	114.091	117.893	116.022	0.000
		O2 C12 N2	Not unusual (enough hits)	34	117.664	119.326	0.983	1.690	1.662	117.131	121.407	119.472	0.000
		O2 C12 C13	Not unusual (enough hits)	4135	124.750	122.007	3.849	0.713	2.743	85.274	175.076	121.832	0.000
		N5 C8 N2	Not unusual (few hits)	4	123.936	123.637	0.517	0.578	0.299	123.145	124.209	123.597	0.000
		C12 N2 C8	Unusual (enough hits)	8	118.784	120.955	1.018	2.132	2.171	118.784	122.010	120.992	0.000
✓ torsion	AABHTZ												



Single click  
on a row to  
view result in  
Mercury

CCDC

# Mogul Geometry Check – Mercury view



CCDC

# Mogul Geometry Check - Results

Mogul Results Viewer

Show / hide : Columns Fragments... Deselect all fragments Export...

Help Double click to view result in Mogul

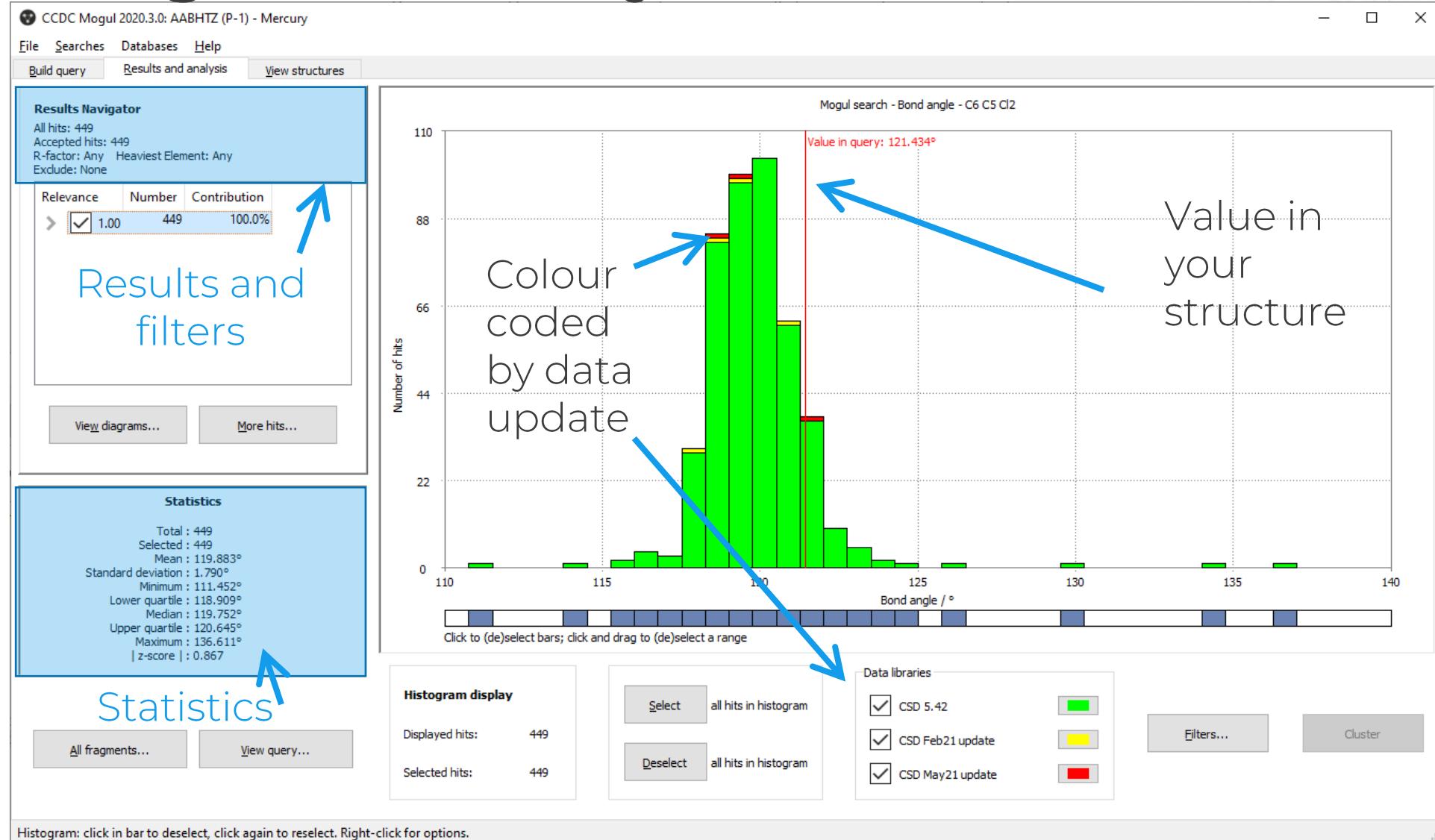
Type	Molecule	Fragment	Classification	No. of hits	Query value	Mean	Std. dev.	z-score	x - mean	Minimum	Maximum	Median	d(min)
✓ bond	AABHTZ												
➤ angle	AABHTZ	C2 C1 Cl1	Not unusual (enough hits)	10637	117.372	118.409	1.966	0.528	1.038	86.905	152.624	118.518	0.000
✓ angle	AABHTZ	C6 C1 Cl1	Not unusual (enough hits)	449	119.477	119.883	1.790	0.227	0.406	111.452	136.611	119.752	0.000
		C2 C1 C6	Not unusual (enough hits)	421	123.139	122.057	2.692	0.402	1.082	79.010	129.445	122.401	0.000
		C3 C2 C1	Not unusual (enough hits)	10491	119.515	119.405	1.501	0.073	0.110	89.047	141.490	119.478	0.000
		C4 C3 C2	Not unusual (enough hits)	20098	119.999	120.239	1.325	0.181	0.240	70.643	140.815	120.159	0.000
		C3 C4 C5	Not unusual (enough hits)	10491	119.854	119.405	1.501	0.299	0.449	89.047	141.490	119.478	0.000
		C4 C5 C12	Not unusual (enough hits)	10637	116.212	118.409	1.966	1.118	2.197	86.905	152.624	118.518	0.000
		C6 C5 C12	Not unusual (enough hits)	449	121.434	119.883	1.790	0.867	1.551	111.452	136.611	119.752	0.000
		C4 C5 C6	Not unusual (enough hits)	421	122.346	122.057	2.692	0.107	0.288	79.010	129.445	122.401	0.000
		C1 C6 C5	Not unusual (enough hits)	170	115.125	116.276	3.163	0.364	1.152	111.918	151.002	115.875	0.000
		C5 C6 C7	Not unusual (enough hits)	142	126.222	122.066	3.194	1.301	4.156	106.172	133.828	121.766	0.000
		C1 C6 C7	Not unusual (enough hits)	142	118.637	122.066	3.194	1.074	3.429	106.172	133.828	121.766	0.000
		C6 C7 N1	Not unusual (enough hits)	63	120.071	121.580	1.633	0.924	1.509	118.121	125.394	121.710	0.000
		C7 N1 N2	Not unusual (enough hits)	304	117.522	117.173	2.634	0.133	0.349	112.028	123.777	117.122	0.000
		C8 N2 N1	Not unusual (enough hits)	16	124.440	121.412	5.026	0.603	3.028	110.211	126.188	123.650	0.000
		C12 N2 N1	Not unusual (enough hits)	15	116.721	115.241	2.607	0.568	1.480	111.791	120.406	114.780	0.000
		N2 C8 N3	Not unusual (enough hits)	15	125.283	124.096	1.376	0.863	1.187	121.050	125.790	124.354	0.000
		N5 C8 N3	Not unusual (enough hits)	18	110.781	110.483	0.852	0.349	0.298	109.567	112.225	110.237	0.000
		C8 N3 N4	Not unusual (enough hits)	47	106.679	105.504	2.395	0.491	1.175	100.785	114.792	105.591	0.000
		C9 N4 N3	Not unusual (enough hits)	190	107.177	107.613	1.828	0.238	0.436	101.988	116.826	107.467	0.000
		N5 C9 N4	Not unusual (enough hits)	61	110.746	110.185	0.564	0.995	0.561	108.650	111.600	110.233	0.000
		C8 N5 N6	Not unusual (enough hits)	15	128.047	125.894	2.689	0.801	2.154	118.623	129.059	126.475	0.000
		C9 N5 C8	Not unusual (enough hits)	15	104.615	105.953	1.319	1.015	1.338	103.742	108.706	105.802	0.000
		C9 N5 N6	Not unusual (enough hits)	47	126.994	126.891	1.763	0.058	0.102	121.092	133.246	127.067	0.000
		C10 N6 N5	Not unusual (enough hits)	15	117.419	118.913	1.088	1.373	1.494	117.419	120.816	118.696	0.000
		C11 C10 N6	Not unusual (enough hits)	37	114.417	114.933	1.651	0.312	0.515	112.356	120.514	114.515	0.000
		O1 C10 N6	Not unusual (enough hits)	37	121.348	121.538	1.427	0.133	0.189	116.489	123.542	121.784	0.000
		O1 C10 C11	Not unusual (enough hits)	4907	124.224	122.048	1.806	1.205	2.177	91.903	161.131	121.976	0.000
		C13 C12 N2	Not unusual (enough hits)	34	117.582	116.154	0.932	1.532	1.428	114.091	117.893	116.022	0.000
		O2 C12 N2	Not unusual (enough hits)	34	117.664	119.326	0.983	1.690	1.662	117.131	121.407	119.472	0.000
		O2 C12 C13	Not unusual (enough hits)	4135	124.750	122.007	3.849	0.713	2.743	85.274	175.076	121.832	0.000
		N5 C8 N2	Not unusual (few hits)	4	123.936	123.637	0.517	0.578	0.299	123.145	124.209	123.597	0.000
		C12 N2 C8	Unusual (enough hits)	8	118.784	120.955	1.018	2.132	2.171	118.784	122.010	120.992	0.000



Double click  
on a row to  
view result in  
Mogul

CCDC

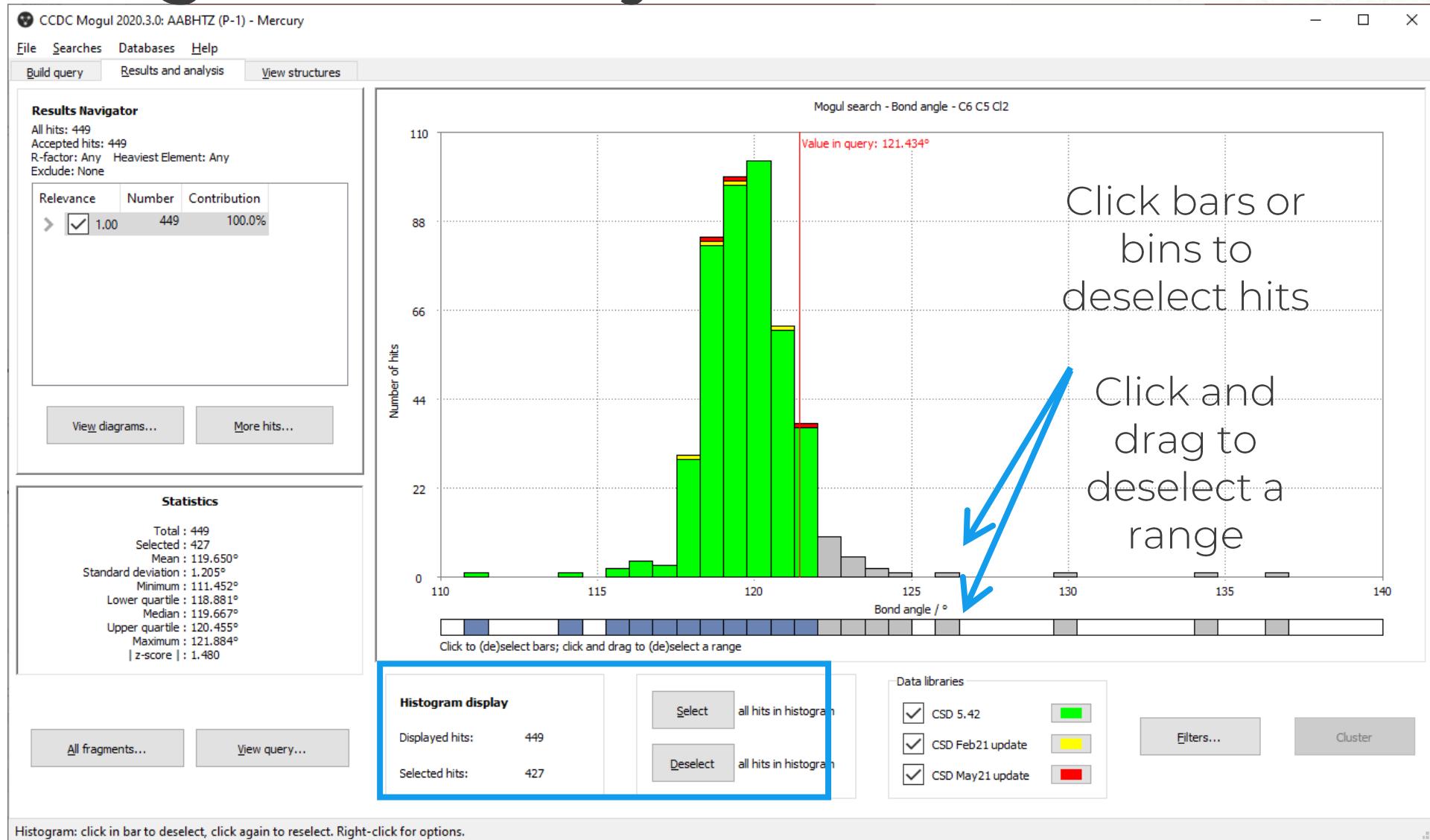
# Mogul Geometry Check – Results



Histogram built with data from structures in the CSD

CCDC

# Mogul Geometry Check – Results



CCDC

# Mogul Geometry Check – View structures

CCDC Mogul 2020.3.0: AABHTZ (P-1) - Mercury

File Searches Databases Help

Build query Results and analysis View structures

Refcode: AABHTZ Data Library: CSD 5.42

Information Diagram 3D Visualiser

**Fragment shown**

**Measurement given**

**Click through hits**

**Number of structures**

Valence angle: 119.477°

AABHTZ

- AABHTZ
- ABAWIJ
- ABAWOP
- ABYTZL
- ACBTZC
- ACBTZD10
- ADAPIF
- ADENON
- AFESEK
- AJETAL
- AJETAL01
- AMAQAH
- AMAQUEL
- AMAQIP
- APIKUH
- AWIJIA
- AWIUG
- AWIJUM
- AXEHIU
- AXUTES
- AYEZOV
- AZITUA
- BAKYAP
- CAGNUU
- CAGPAC
- CEBLOM
- CEQZJU
- CIUGAD
- CLPTBU
- COPZOX
- DAJCAV
- DAJNUZ
- DIDMUA

196 structures

Fragment

1 2 hits

Show Parameters

View entries

CCDC

# Mogul Geometry Check – View structures

CCDC Mogul 2020.3.0: AABHTZ (P-1) - Mercury

File Searches Databases Help

Build query Results and analysis View structures

Refcode: AABHTZ Data Library: CSD 5.42

Information Diagram 3D Visualiser

Identifier	AABHTZ
Literature Reference	P.-E.Werner, <i>Cryst.Struct.Commun.</i> (1976), <b>5</b> , 873
Formula	C <sub>13</sub> H <sub>12</sub> Cl <sub>2</sub> N <sub>6</sub> O <sub>2</sub>
Compound Name	4-Acetoamido-3-(1-acetyl-2-(2,6-dichlorobenzylidene)hydrazine)-1,2,4-triazole
Synonym	
Space Group	P 1 (2)
Cell Lengths	a 11.372(9) b 10.272(5) c 7.359(9)
Cell Angles	α 108 °
Cell Volume	769.9
Z, Z'	Z: 2 Z'
R-Factor (%)	4.1
Disorder	
Polymorph	

**Customise**

Available Items (Right-click for options)

- Cell
  - Reduced Cell Lengths
  - Reduced Cell Angles
  - Reduced Cell Volume
- Experimental
  - Temperature (K)
  - Pressure
  - Density (CCDC)
  - Average Sigma (C-C)

Add >> << Remove

Selected Items (Right-click for options)

- Identifier
- Literature Reference
- Formula
- Compound Name
- Synonym
- Space Group

Up Down OK Cancel

Customise

View entries

Disorder  
comments  
can be useful  
when  
exploring  
molecular  
geometries

Click through  
the refcode list

CCDC

# Mogul Geometry Check – View structures

CCDC Mogul 2020.3.0: AABHTZ (P-1) - Mercury

File Searches Databases Help

Build query Results and analysis View structures

Refcode: ODAVIZ Data Library: CSD 5.42

Information Diagram 3D Visualiser

Reset display  Display fragments

View entries

Fragment and measurement shown

ODAVIZ

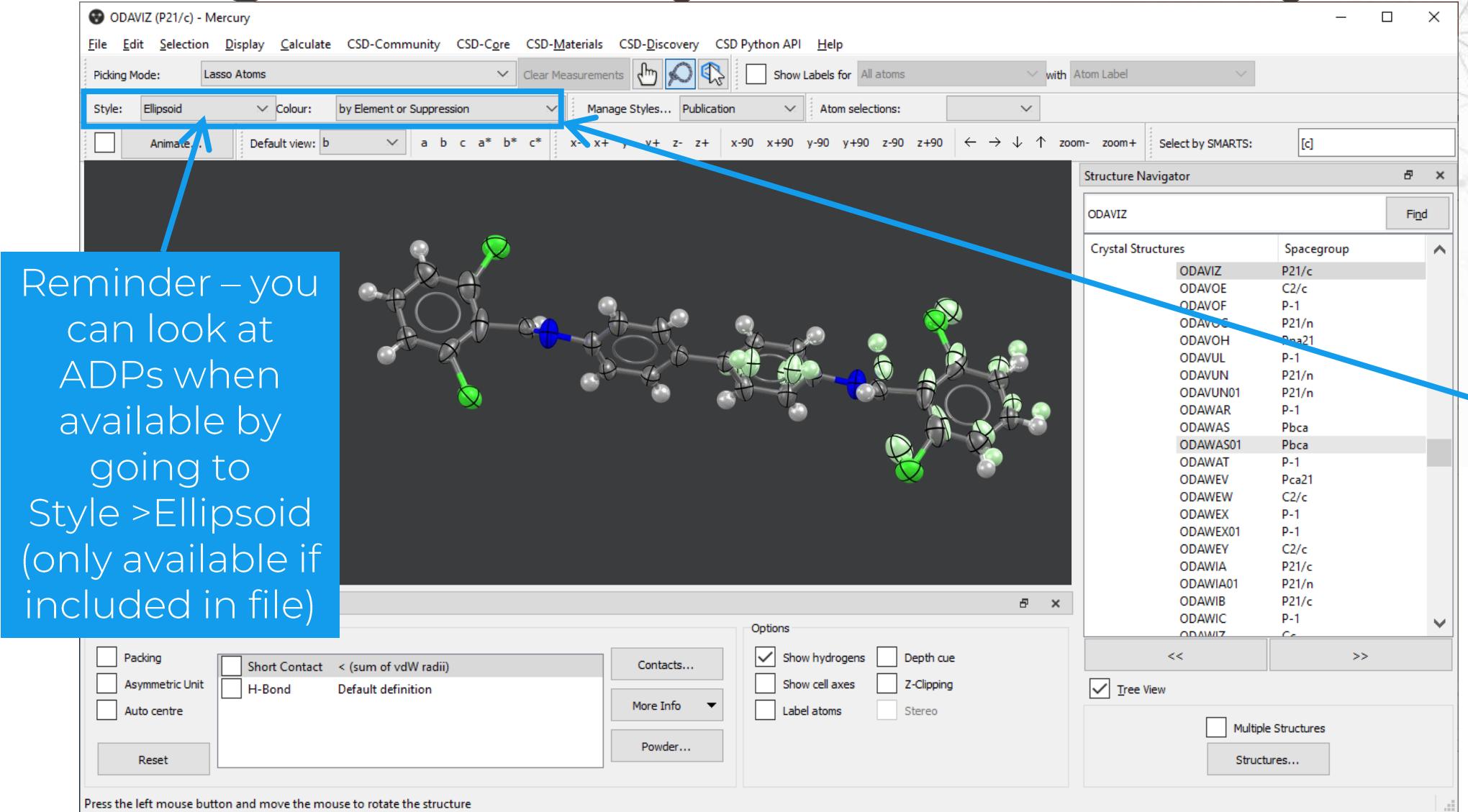
- IZAQEH
- JAXZUE
- JEYVEC
- JEYJIG
- JIYKIM
- JUFCUI
- KEKVOO
- KEKWOP
- KERVOT
- KEVQOS
- KOCZUA
- KOSKIM
- KOXWEZ
- KUSPAQ
- KUZVIL
- LAPROL
- MESSEQ
- MIFFOX
- NATWAJ
- NAVCAR
- NELYAH
- NETZIV
- OCOSEG
- ODAVIZ**
- OFOQQQ
- OKADOW
- OKIGIA
- OPUPAT
- PADKUE
- PAWWOA
- PEGZOQ
- PEHCOV
- POOHFK

196 structures

Click through the refcode list

CCDC

# Mogul Geometry Check – Mercury view



CSD Refcode  
ODAVIZ

Reminder – you can look at ADPs when available by going to Style >Ellipsoid (only available if included in file)

Reminder – disordered atoms in the CSD are often suppressed and can be coloured differently using Colour > by Element or Suppression

CCDC

# Mogul Geometry Check – Build query

CCDC Mogul 2020.3.0: AABHTZ (P-1) - Mercury

File Searches Databases Help

Build query Results and analysis View structures

Current Selection:  
C6 C5 Cl2

Search All fragments... Settings... Reset

Edit... Auto Edit... Draw... Load...

Hide hydrogens Show labels

Search progress: Stop

Current selection

Current measurement

Learn about building and editing a Mogul query in Exercise 2 of the handout.

The screenshot shows the CCDC Mogul software interface. The top menu bar includes File, Searches, Databases, Help, Build query (which is selected), Results and analysis, and View structures. The main window has a left sidebar with 'Current Selection' set to 'C6 C5 Cl2', a 'Search' button, and options for 'All fragments...', 'Settings...', and 'Reset'. Below this are buttons for 'Edit...', 'Auto Edit...', 'Draw...', and 'Load...'. At the bottom are checkboxes for 'Hide hydrogens' and 'Show labels', and a 'Search progress:' bar with a 'Stop' button. The central area displays a chemical structure with various atoms highlighted in blue and red, and a specific bond length labeled '121.43' with a green line and arrow. A callout box on the right contains the text: 'Learn about building and editing a Mogul query in Exercise 2 of the handout.'

# Mogul Settings

The screenshot shows the CSD Core interface with a molecular model of AABHTZ (P-1) - Mercury. The 'CSD-Core' menu is open, and the 'Mogul Settings...' option is highlighted with a blue arrow pointing to the 'Mogul Data Library Settings' dialog box.

**Mogul Data Library Settings**

Select which Mogul data libraries to include in new searches

Include Library	Path To Data
1 <input checked="" type="checkbox"/> CSD 5.42	C:\Program Files\CCDC\CSD_2021\Mogul V5.42
2 <input checked="" type="checkbox"/> CSD Feb21 update	C:\Program Files\CCDC\CSD_2021\Mogul V5.42\Feb21
3 <input checked="" type="checkbox"/> CSD May21 update	C:\Program Files\CCDC\CSD_2021\Mogul V5.42\May21

OK Cancel

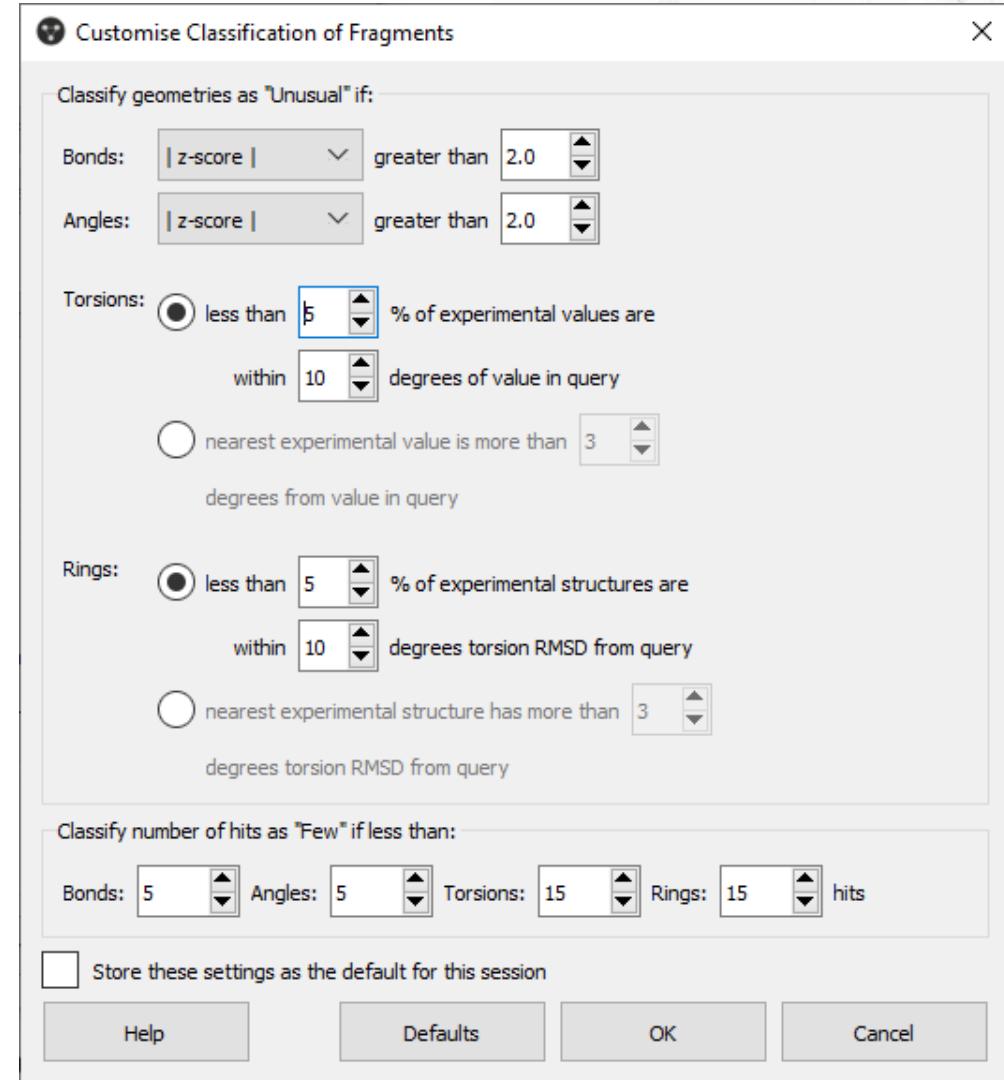
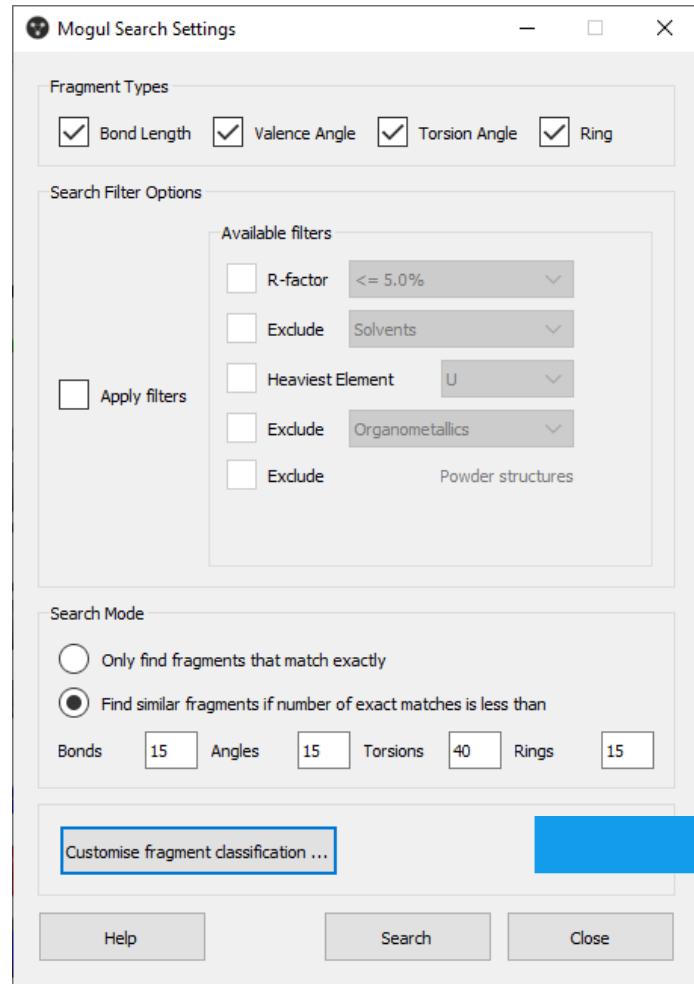
Chemical structures shown in the interface:

- AABHTZ (P-1) - Mercury molecule
- Water molecule (H2O)
- Ammonium ion (NH4+)

Databases listed in the bottom right:

- AADAMC P21/C
- AADMPC P-1
- AADMPC10 P-1
- AADRIB P21
- AAGAGG10 P2121
- AAGGAG10 P21
- AALCFE P21/c

# Customise fragment classification



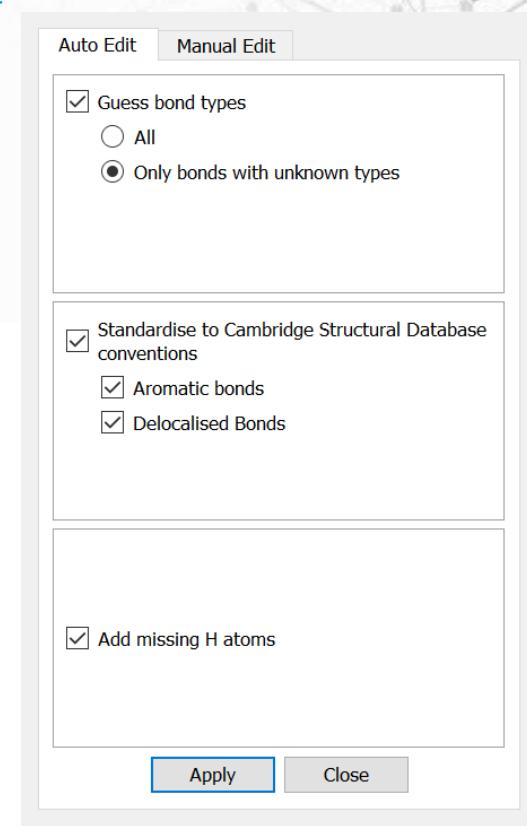
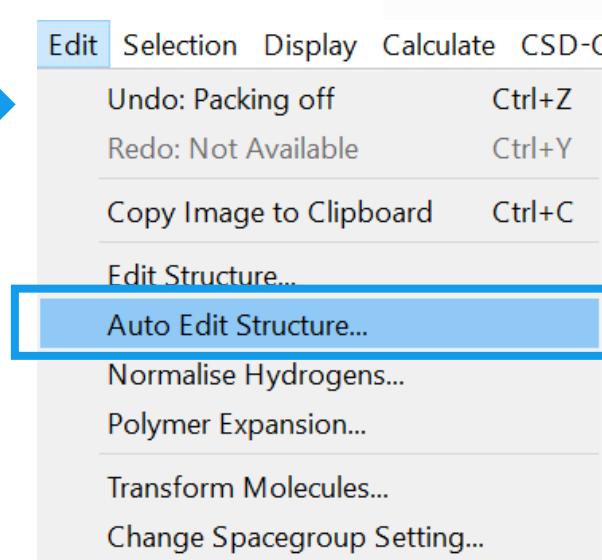
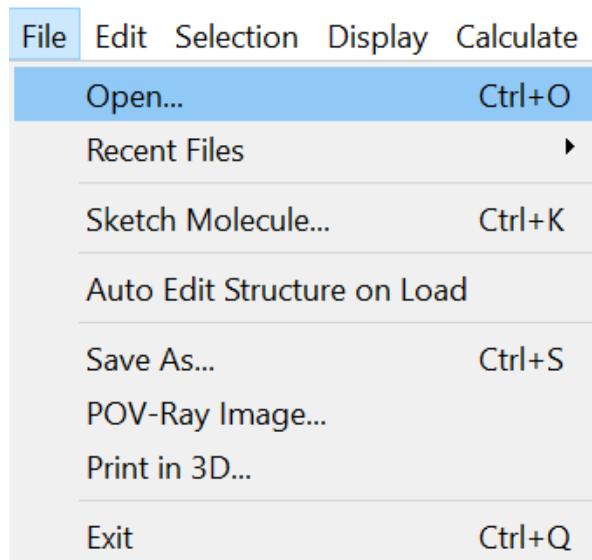
CCDC

# Using Mogul to check a new structure

## 1. Load your structure in Mercury

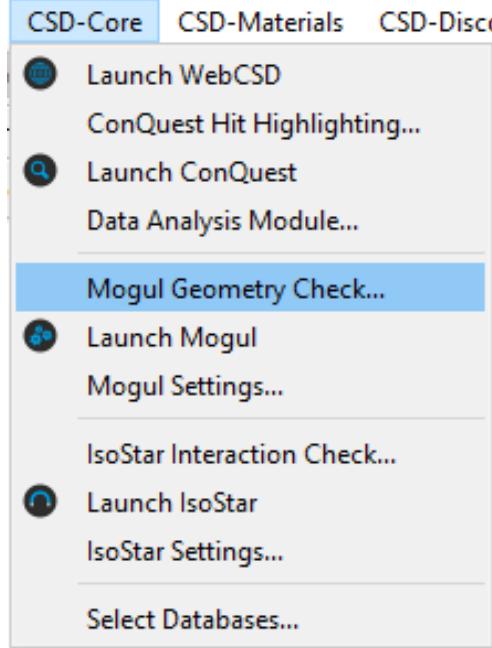
- Do you remember how to do it from the Visualisation session?
- File > Open – to open one of your own files for example a CIF
- Edit > Auto Edit Structure to assign bond types
- File > Auto Edit Structure on Load to automate

AABHTZ (P-1) - Mercury



# Using Mogul to check a new structure

2. Repeat the procedure you did before



The screenshot shows the CSD interface with the 'CSD-Core' tab selected. In the main menu, the 'Mogul Geometry Check...' option is highlighted.

**Mogul Results Viewer**

This window displays search results for a torsion angle query. The results are listed in a table:

Type	Molecule	Fragment	Classification	No. of hits	Query value	Me:
torsion	LIM_K21_901-B_pdb1hak_1	C59 O58 C1 C2	Not unusual (enough hits)	13461		
		C59 O58 C1 C6	Not unusual (enough hits)	13461		
		O23 C21 N11 C12	Not unusual (enough hits)	45		
		C24 C22 C21 N11	Not unusual (enough hits)	40		
		C22 C24 N27 C30	Not unusual (enough hits)	938		
		C22 C24 N27 C34	Not unusual (enough hits)	938		
		C45 C52 C43 C32	Not unusual (enough hits)	2357		
		C51 C52 C43 C32	Not unusual (enough hits)	2357		
		C22 C21 N11 C12	Not unusual (few hits)	3		
		O23 C21 N11 C10	Unusual (enough hits)	45		
		O23 C21 C22 C24	Unusual (enough hits)	42		
		C21 C22 C24 N27	Unusual (enough hits)	137		
		C31 C32 C43 C52	Unusual (enough hits)	131		
		C33 C32 C43 C52	Unusual (enough hits)	131		
		C22 C21 N11 C10	Unusual (few hits)	3		

**Results Navigator**

All Hits: 137  
Accepted hits: 137  
R-factor: Any Heaviest Element: Any  
Exclude: None

Relevance	Number	Contribution
> <input checked="" type="checkbox"/>	1.00	137 100.0%

**Statistics**

Total : 137  
Selected : 137  
| d(min) | : 27.484°

**Histogram display**

Number of hits vs Torsion angle /°. The histogram shows a peak around 22.055°.

**Data libraries**

- CSD 5.37
- CSD Nov15 update
- CSD Feb16 update

# Why use Mogul?

- To enable you to validate the complete geometry of a given query structure and identify any unusual features
  - Could be useful when looking at your own structure or a CSD structure or in fact ligands in the pdb
- To rapidly retrieve geometric data and gain new insights from structures in the CSD

# Explore More: More advanced uses

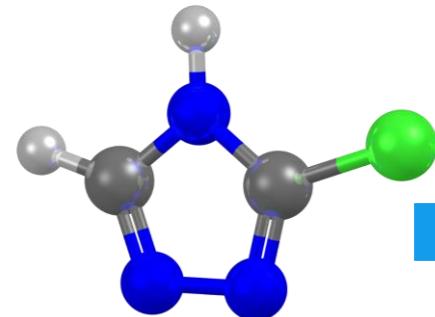
- Not enough time to explore all the insights that you can get from Mogul.
- But here are a selection of:
  - [Case studies](#) - Examples from published scientific articles that used Mogul in their work.
  - Other ways you can access Mogul data

The session has resumed recording

CCDC

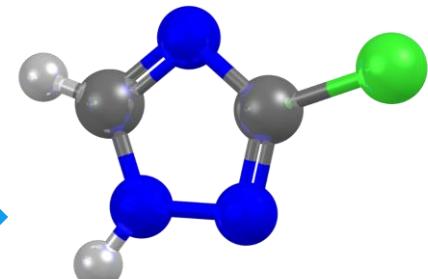
# Using Mogul to assign tautomeric forms

CSD CLTRZL

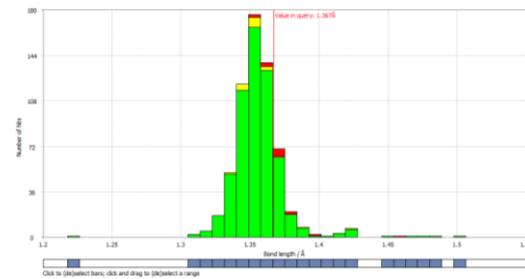
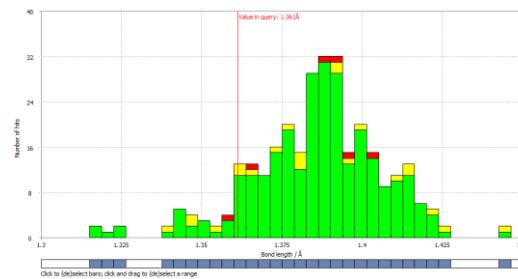


1.361 Å

CSD CLTRZL01

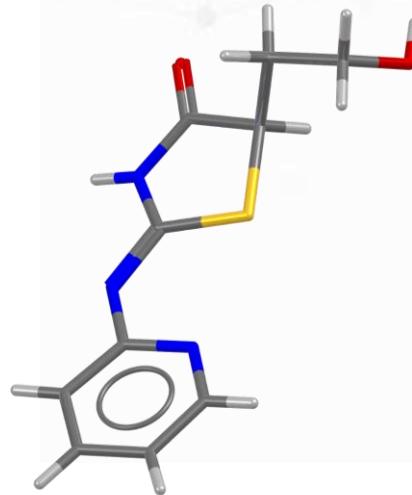


1.367 Å

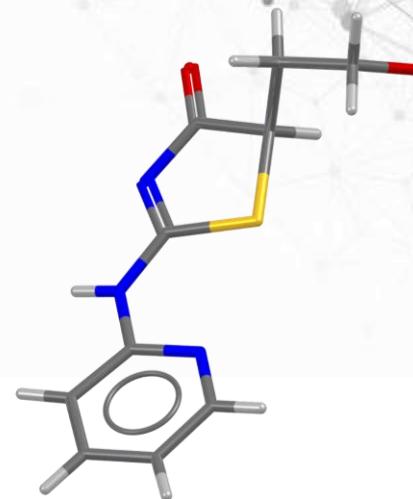


Following Mogul analysis structure re-determined in a different tautomeric form

CSD GACXOZ

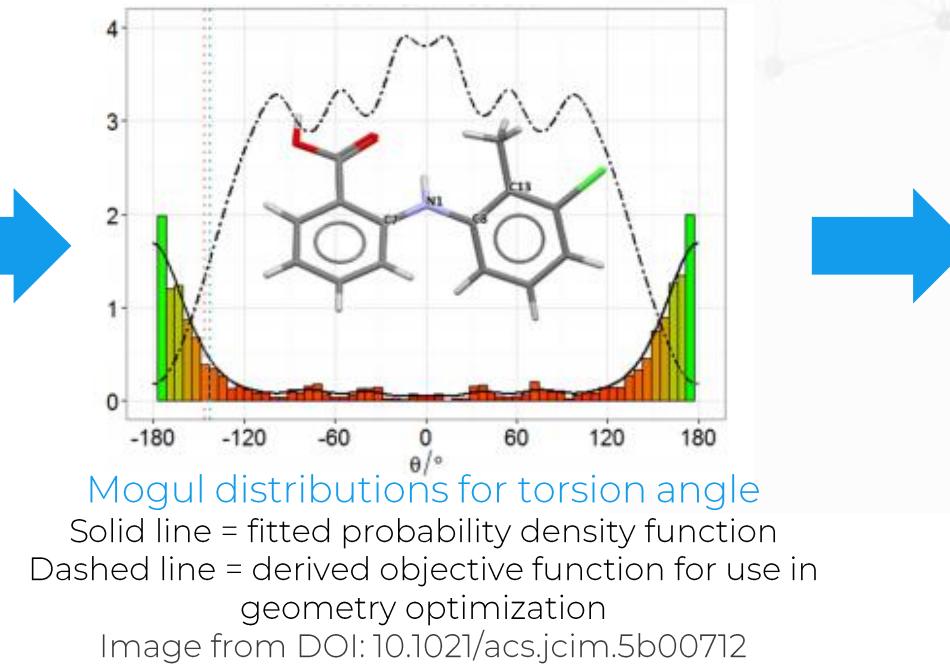
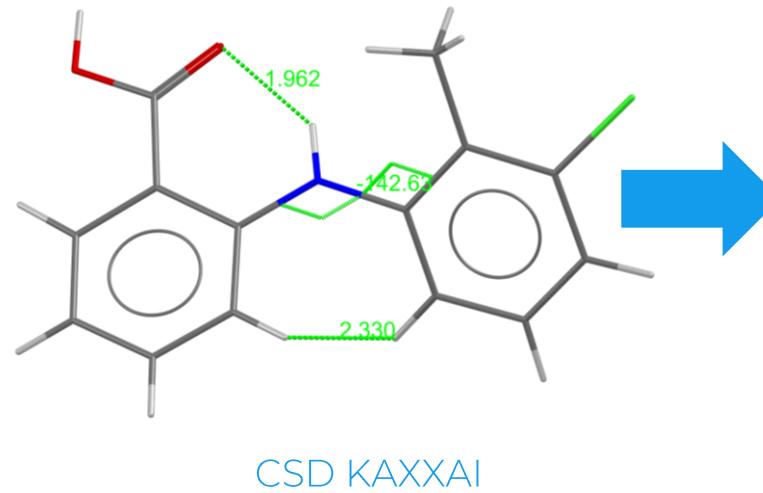


CSD LOQBIE



A comparison of the 1,3-thiazol-4-one structures in which Mogul queries the C—N bond length of the imine of GACXOZ, but finds all geometrical parameters of the amine version to be within expected limits

# Using Mogul to optimise molecular geometries



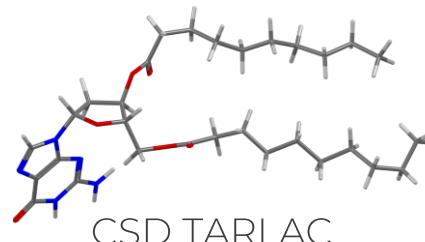
This approach provides average geometric preferences of a molecule according to data in the CSD. It is then possible to exploit this information into an object function to optimise a molecular geometry

Mogul distributions can be converted into smooth, differentiable probability density functions using kernel density estimation. This publication by McCabe *et al.* shows how this approach could be used to drive geometry optimisation of organic molecules

# Using Mogul in powder diffraction

- Powder diffraction patterns have relatively low content information
  - Overlap of reflections
  - Rapid decline in scattering with increasing diffraction angle
- How can the CSD and Mogul help?

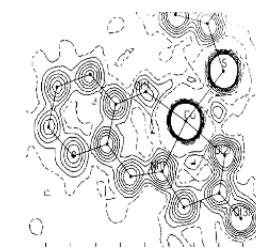
By providing sensible geometries for a starting model



To determine refinement constraints



To help validate the final solution

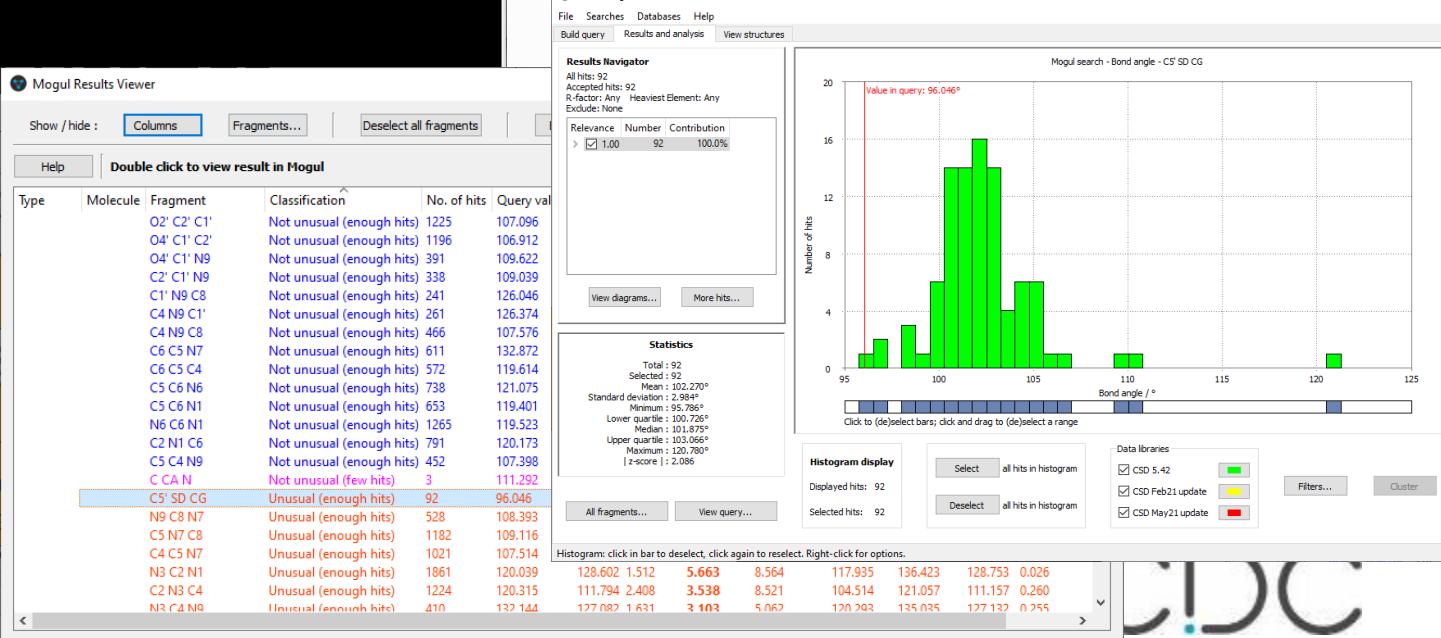
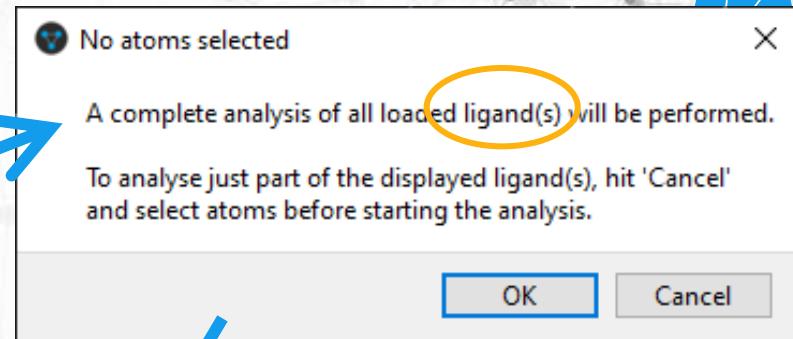
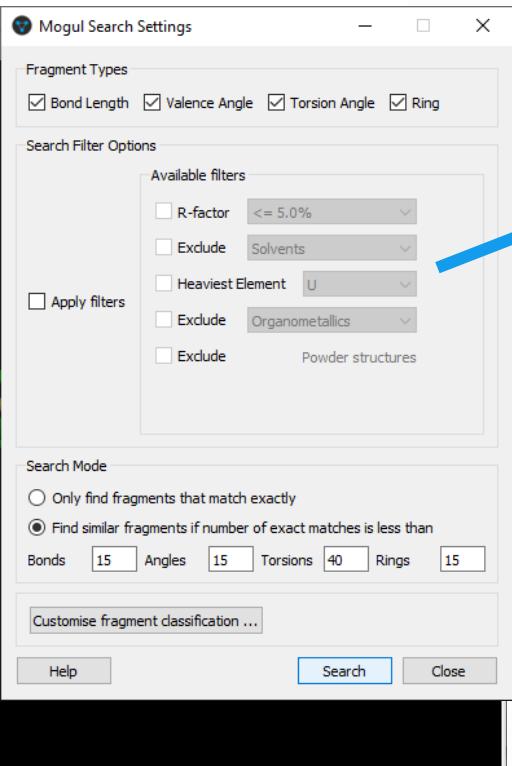
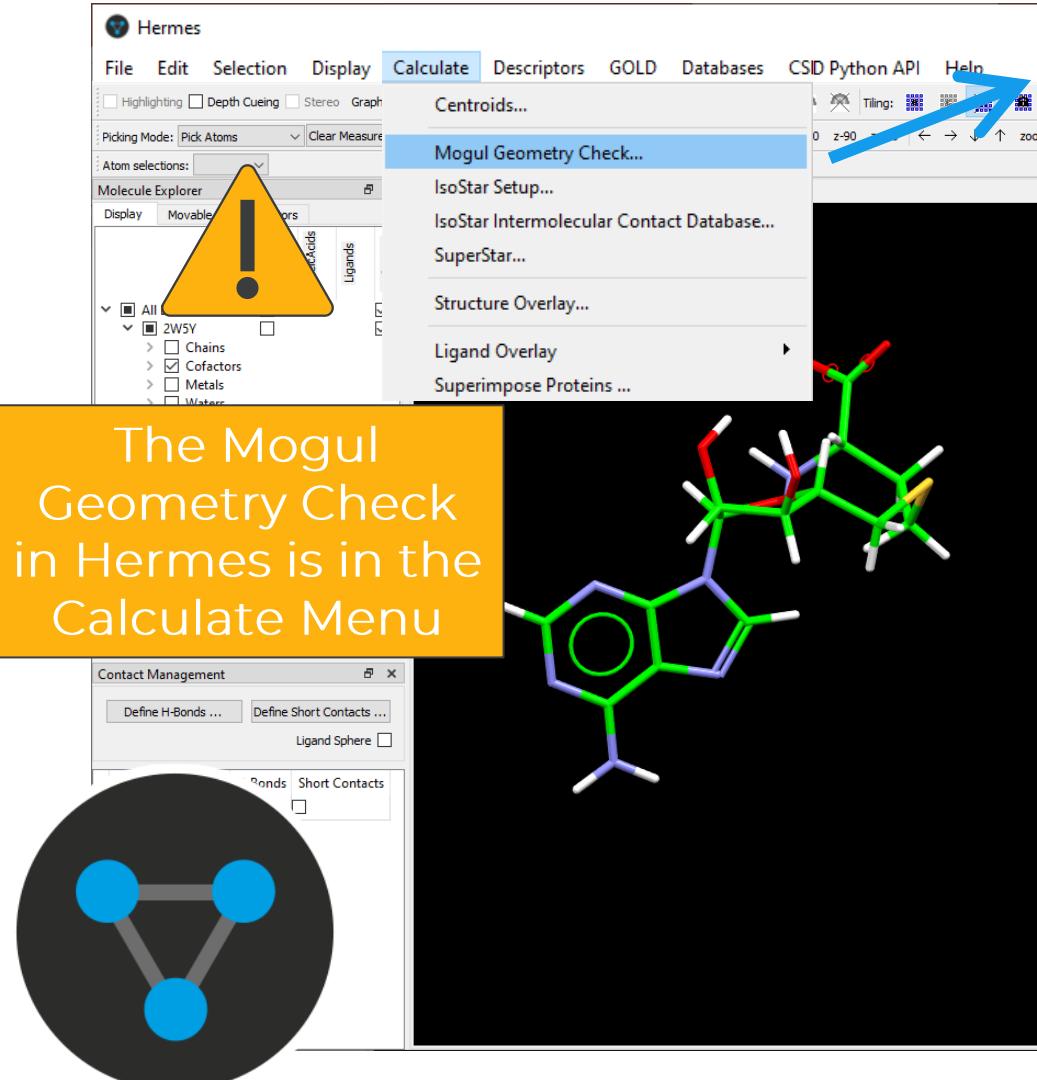


By using conformational info to reduce the search space



CCDC

# In Hermes

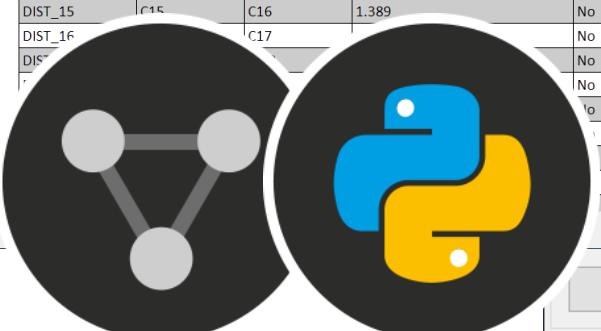


# Creating Mogul reports in Mercury

**Molecular Geometry Report for ODAVIZ**

**Bond Lengths**

Name	Atom 1	Atom 2	Bond Length (Å)	Notes
DIST_1	C13	Cl1	1.708	No
DIST_2	C17	Cl2	1.718	No
DIST_3	C1	N1	1.422	No
DIST_4	C2	C1	1.359	No
DIST_5	C3	C2	1.377	No
DIST_6	C3	C4	1.392	No
DIST_7	C5	C4	1.412	No
DIST_8	C6	C5	1.392	No
DIST_9	C6	C1	1.388	No
DIST_10	C11	N1	1.226	No
DIST_11	C12	C11	1.506	Yes
DIST_12	C12	C13	1.39	No
DIST_13	C14	C13	1.391	No
DIST_14	C15	C14	1.39	No
DIST_15	C16	C16	1.389	No
DIST_16	C17			No
DIST_17				No



Press the **Spacebar** to toggle between ball-and-stick and spacefill models.

The python interpreter is: C:\Program Files\CCDC\Python\_API\_2021\miniconda\python.exe  
The working directory is: C:/Users/ward/ODAVIZ/quick\_geometry\_check/2021\_07\_01\_10\_17\_54  
Output files will be written in: C:/Users/ward/ODAVIZ/quick\_geometry\_check/2021\_07\_01\_10\_17\_54

Script completed...  
16 seconds

**ODAVIZ (P21/c) - Mercury**

File Edit Selection Display Calculate CSD-Community CSD-Core CSD-Materials CSD-Discovery CSD Python API Help

Atoms Clear Measurements  

Colour: by Element or Suppression Manage Styles... Public

Default view: b a b c a\* b\* c\* x- x+ y- y+ z- z+

**Analysis**

**Reports**

user\_support.py welcome.py calculate\_CSD\_diversity\_score.py Options... CSD Python API Documentation CSD Python API Forum

**Spacegroup**

P21/c

ODAVUOE ODAVOF ODAVOG ODAVOH ODAVUL ODAVUN ODAVUN01 ODAWAR ODAWAS ODAWAS01 ODAWAT ODAWEV ODAWEW ODAWEX ODAWEX01 ODAWEY ODAWIA ODAWIA01 ODAWIB ODAWIC ODAWIC01

CC advancing structure

crystal\_structure\_report.py molecular\_geometry\_report.py quick\_geometry\_check.py quick\_packing\_check.py simple\_report.py

[c] Find

Number of Hits

Bond Lengths for DIST\_11 (C12\_C11) (Click to Return to Table)  
Value in query: 1.506

Number of Hits

Bond Lengths for DIST\_27 (C12B\_C11B) (Click to Return to Table)  
Value in query: 1.506

Open in Browser...

Options

Show hydrogens  Depth cue  
 Show cell axes  Z-Clipping  
 Label atoms  Stereo

<< >>

Tree View

Multiple Structures

Structures...

# In the CSD Python API

CSD Python API 3.0.4 documentation » API documentation »



## Table of Contents

### Conformer API

- Introduction
- API
  - Knowledge base version number
  - Molecule minimisation
  - Conformer generation
  - Geometry analysis

## Previous topic

Search API

## Next topic

Protein API

## Quick search

## Conformer API

### Introduction

The `ccdc.conformer` module contains classes concerned with molecular conformations.

The three main classes of the `ccdc.conformer` module are:

- `ccdc.conformer.MoleculeMinimiser`
- `ccdc.conformer.ConformerGenerator`
- `ccdc.conformer.GeometryAnalyser`

A `ccdc.conformer.MoleculeMinimiser` instance can be used to optimise the bond distances and valence angle 3D input molecule using the `ccdc.conformer.MoleculeMinimiser.minimise()` function:

```
from ccdc.conformer import MoleculeMinimiser
molecule_minimiser = MoleculeMinimiser()
minimised_mol = molecule_minimiser.minimise(mol)
```

A `ccdc.conformer.ConformerGenerator` instance can be used to generate a set of conformers for an input mol using the `ccdc.conformer.ConformerGenerator.generate()` function:

```
from ccdc.conformer import ConformerGenerator
from ccdc.io import MoleculeWriter
conformer_generator = ConformerGenerator()
conformers = conformer_generator.generate(mol)
with MoleculeWriter('conformers.mol2') as mol_writer:
    for c in conformers:
        mol_writer.write(c.molecule)
```

A `ccdc.conformer.GeometryAnalyser` instance can be used to analyse the geometry of an input molecule using knowledge-based library of intramolecular geometries based on the CSD.

The `ccdc.conformer.GeometryAnalyser` class contains nested classes:

- `ccdc.conformer.GeometryAnalyser.Settings`
- `ccdc.conformer.GeometryAnalyser.Analysis`
- `ccdc.conformer.GeometryAnalyser.AnalysisHit`

The `ccdc.conformer.GeometryAnalyser.analyse_molecule()` function can be used to validate the complete geometry of a given query structure.

```
>>> from ccdc.io import EntryReader
>>> csd_reader = EntryReader('CSD')
>>> yigpio01 = csd_reader.molecule('YIGPIO01')
>>> from ccdc.conformer import GeometryAnalyser
```



## ccdc.conformer.GeometryAnalyser()

CSD Python API 3.0.4 documentation » Descriptive documentation »



### Table of Contents

#### Molecular geometry analysis

- Introduction
- The molecular geometry analysis engine
- Geometry analysis settings
- Performing a geometry analysis on a molecule
- Analysing the results
- Molecular geometry analysis with multiple data libraries

### Previous topic

SMARTS implementation

### Next topic

Analysing molecular interactions preferences

### Quick search

[Go](#)

## Molecular geometry analysis

### Introduction

The CSD Portfolio contains a knowledge-base of intramolecular geometries. This knowledge-base provides easy and rapid access to millions of chemically classified bond lengths, valence angles, acyclic torsion angles, and ring conformations derived from the CSD.

This enables you to rapidly validate the complete geometry of a given query structure and identify any unusual features without the need to construct complex search queries, or carry out detailed data analyses.

Let us import in the `ccdc.conformer` module.

```
>>> from ccdc import conformer
```

Let us also import the `ccdc.io` module so that we can read in molecules of interest.

```
>>> from ccdc import io
```

**Note:** For more information on these knowledge-based geometry libraries please see: "Retrieval of Crystallographically-Derived Molecular Geometry Information", I. J. Bruno, J. C. Cole, M. Kessler, Jie Luo, W. D. S. Motherwell, L. H. Purkis, B. R. Smith, R. Taylor, R. I. Cooper, S. E. Harris and A. G. Orpen, *J. Chem. Inf. Comput. Sci.*, 44, 2133-2144, 2004 DOI: 10.1021/ci049780b.

**See also:** API documentation of the geometry analyser module

### The molecular geometry analysis

In order to be able to carry out a molecular geometry analysis we will use the `ccdc.conformer.GeometryAnalyser` class. Let us create an instance of this class:

```
>>> engine = conformer.GeometryAnalyser()
```

**Note:** The most important function of an instance of the `GeometryAnalyser` class is the `ccdc.conformer.GeometryAnalyser.analyse_molecule()` function which performs the geometry analysis of the molecule.

### Geometry analysis settings

The settings used in a geometry analysis are contained in the `ccdc.conformer.GeometryAnalyser.Settings` class.

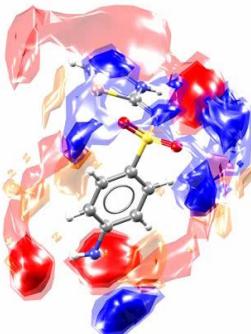
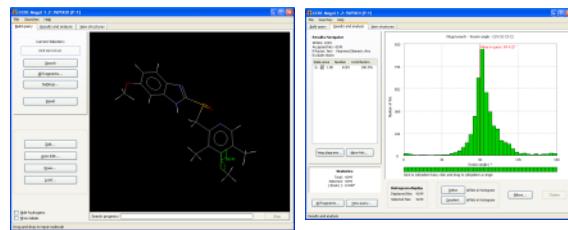
To find out what the current settings are one can make use of the `ccdc.conformer.GeometryAnalyser.Settings.summary()` function.

```
>>> print(engine.settings.summary())
Generalisation: True
Impose upper limits: False
Filter r-factors: any
```

**Find out more:**  
The basics of our CSD Python API were covered in our last series of CCDC Virtual workshops!

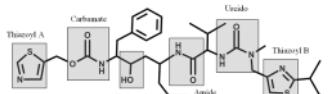
# More tools to assess your structure in CSD-Materials in Mercury

Mogul geometry library

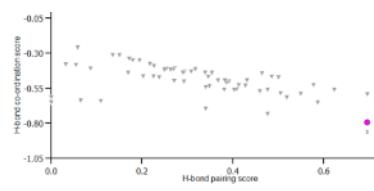


**Full Interaction Maps:**  
3D representation of a map of donor and acceptor atom available to create an interaction network.

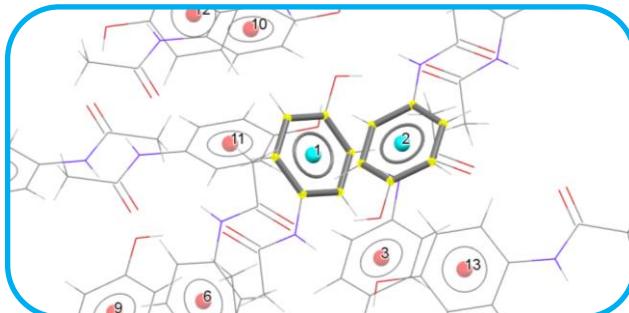
Next week's virtual workshop!



Donor group	Acceptor Group	$\pi$	$+/-  \pi $	Form I	Form II
amide	carbamate	0.618	0.094	✗	✗
amide	hydroxyl	0.551	0.052	✗	✓
carbamate	carbamate	0.538	0.090	✓	✗
amide	amide	0.501	0.055	✓	✗
carbamate	amide	0.420	0.083	✗	✓
hydroxyl	ureido	0.417	0.058	✓	✓
ureido	carbamate	0.319	0.086	✗	✓
ureido	ureido	0.224	0.044	✓	✗
hydroxyl	thiazoyl a	0.114	0.039	✓	✓



**Hydrogen Bond Propensity:**  
Evaluation of the relative likelihoods of possible H-bonding networks in any observed polymorphs of a target system.

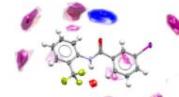


Aromatics Analyser:

Visualisation and identification of aromatic interactions within a crystal structure, and assessing their strength based on their distance and relative orientation.

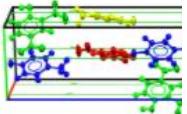
# Want to explore more?

## Training and Educational Resources



CSD-Materials

Tools to help you to understand your material's behaviours and refine its properties.



CSD-Core

Essential crystallographic and structural chemistry capabilities.



Information on the Teaching Subset

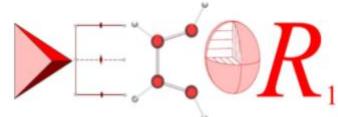
In the Cambridge Structural Database (CSD) extends far beyond a collection of crystal structures. Knowledge derived from these materials, and biology. Chemical and structural concepts are often difficult to grasp without real world, interactive examples for students to explore. We produce educational materials for use in classroom and computer lab settings, or as independent study modules. Many of these materials are freely available set of over 750 structures that can be investigated with the free version of our Mercury visualisation and analysis program. Of these entries are available for free through our Access Structures portal.

Elementary teaching materials, find out more about the Teaching Database [here](#). If you have developed your own modules using the CSD and would like to share them with the wider community, please contact us at [education@ccdc.cam.ac.uk](mailto:education@ccdc.cam.ac.uk).

Keep up to date with the latest news and developments from education and outreach at the CCDC, sign up for the Education and Outreach Newsletter [here](#).



Access a series of teaching modules for use in the classroom



DECOR: Educational Resources for Teaching Crystallography

Register for E&O newsletter

## Self-guided workshops



Download a series of self-guided workshop materials for CCDC tools and features



Watch software training and support videos



CSD University modules

On-demand modules with completion certificate



Access fun science activities for kids through the CCDC Home learning page

Group	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1																		
2																		
3																		
4																		
5																		
6																		
7																		

Explore the Periodic Table through Crystal Structures

YouTube and LabTube channels

CCDC