Intermediate Mogul – More in-depth analysis of molecular geometries

CCDC Virtual Workshop November 2021 – Session 3

Joanna Stevens, Elna Pidcock, Jeff Lengyel, Suzanna Ward, Ilaria Gimondi, Eva Myers, Yinka Olatunji-Ojo

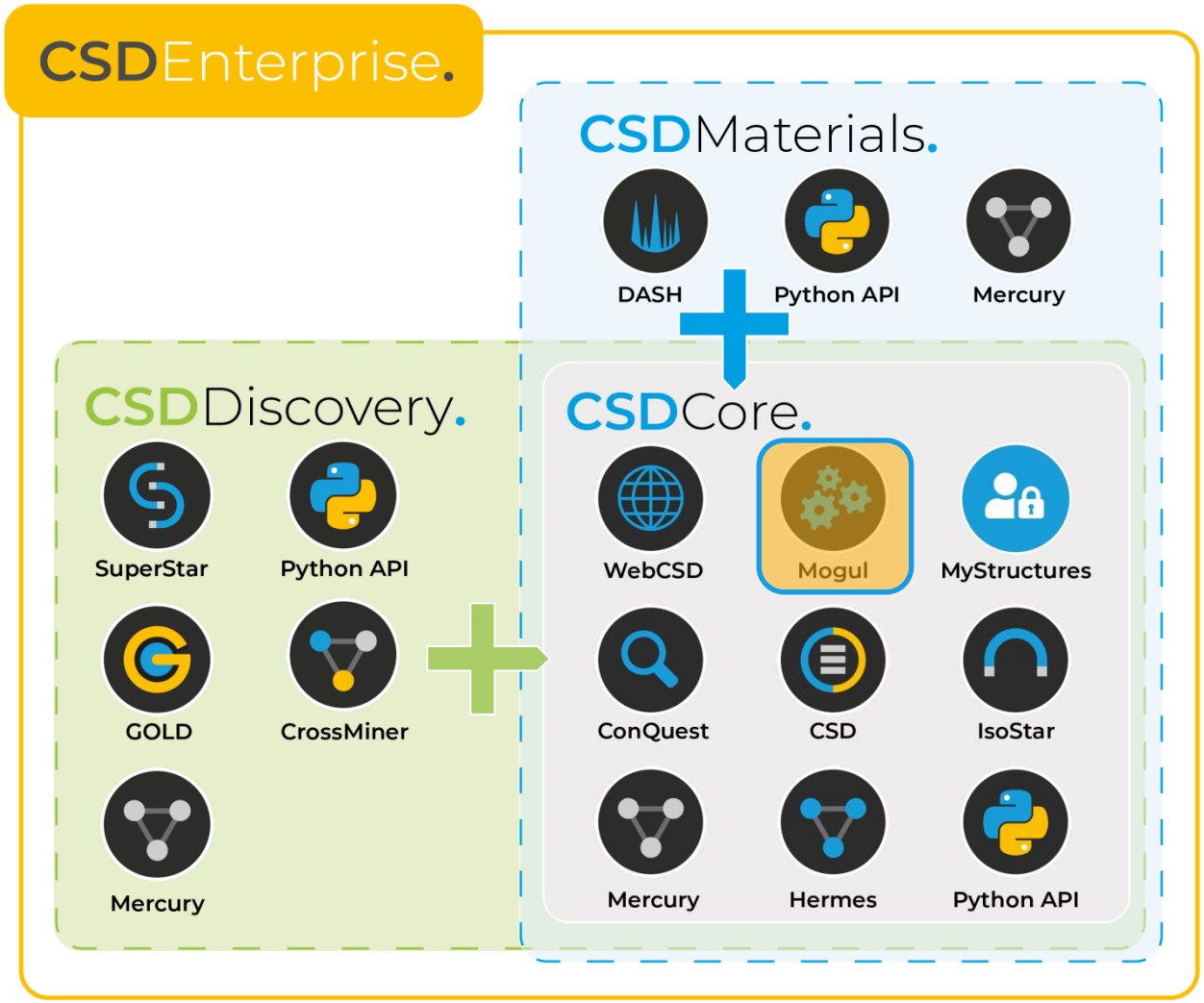
Nov 2021

CCDC
advancing structural science

Learning outcomes for today

- Reminder of some of the basics of Mogul
- How to get more insights into your crystal structures and interrogate results obtained from Mogul
  - Interpreting not unusual / unusual results given context
  - Analysis and impact of specific groups
  - Explicative case studies

The CSD software



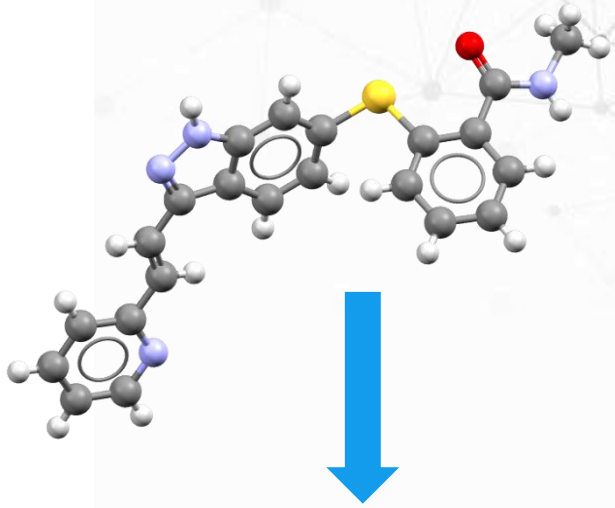
Professional Services

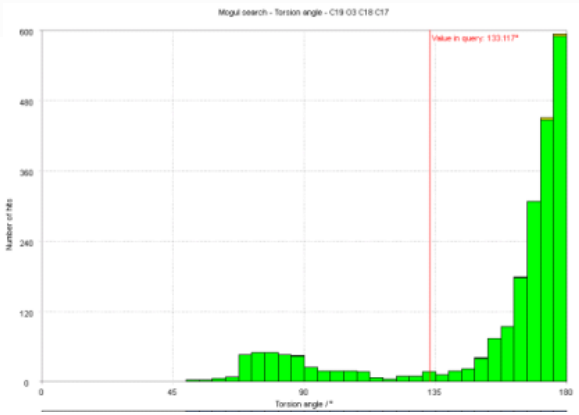
Research & Knowledge partnerships

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

Bruno *et al., J. Chem. Inf. Comput. Sci.*, 44, 2133-2144, 2004

CCDC

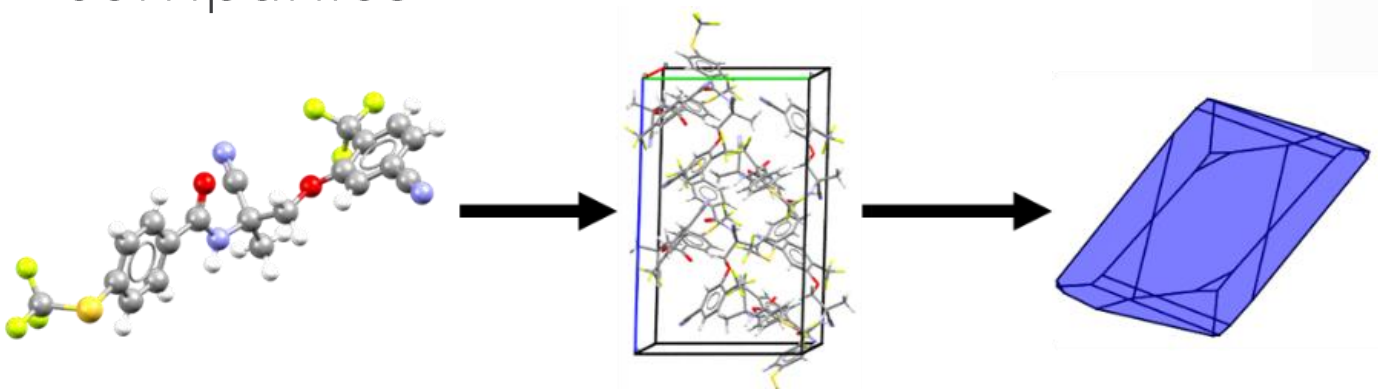
When could we use Mogul?

- 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

CCDC

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



JPP Journal of Pharmacy and Pharmacology

Research Paper

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

Neil Feeder[a], Elna Pidcock[a], Anthony M. Reilly[a], Ghazala Sadiq[a], Cheryl L. Doherty[b], Kevin R. Back[b], Paul Meenan[c] and Robert Docherty[b]

[a]The Cambridge Crystallographic Data Centre, Cambridge, [b]Pharmaceutical Science, Pfizer Global R&D, Sandwich, UK and [c]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 Struc-

CrystEngComm

PAPER

View Article Online

Knowledge-based approaches to co-crystal design†

Peter A. Wood,* Neil Feeder, Matthew Furlow, Peter T. A. Galek, Colin R. Groom and Elna Pidcock

Cite this: *CrystEngComm*, 2014, 16, 5839

Received 12th February 2014, Accepted 10th March 2014

DOI: 10.1039/c4ce00316k

www.rsc.org/crystengcomm

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.

1. Introduction

The definition of a co-crystal,[1–4] whether this is the correct term to use[5] 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 Crystal 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 structure of a synthetic β-*C*-arylglucoside published by Deshpande and co-workers in 2012.[7] This crystal structure contains water, ethanol, a zwitterionic component and a neutral API, so could be described as a co-crystal, hydrate, solvate, zwitterionic salt or all of the above. Practically, the definitions used will also vary based on the application of the resulting co-crystal and the approaches for co-crystal design. In this paper we focus mostly on systems where all the components are neutral, organic species and point out if any further assumptions or requirements are added through the use of a certain design method.

Co-crystals are of interest in both academic research and industrial application as they represent the opportunity to access new solid forms and therefore modified physical properties. This contrasts with polymorphism which, when observed, can be difficult to control and rarely represents an opportunity to

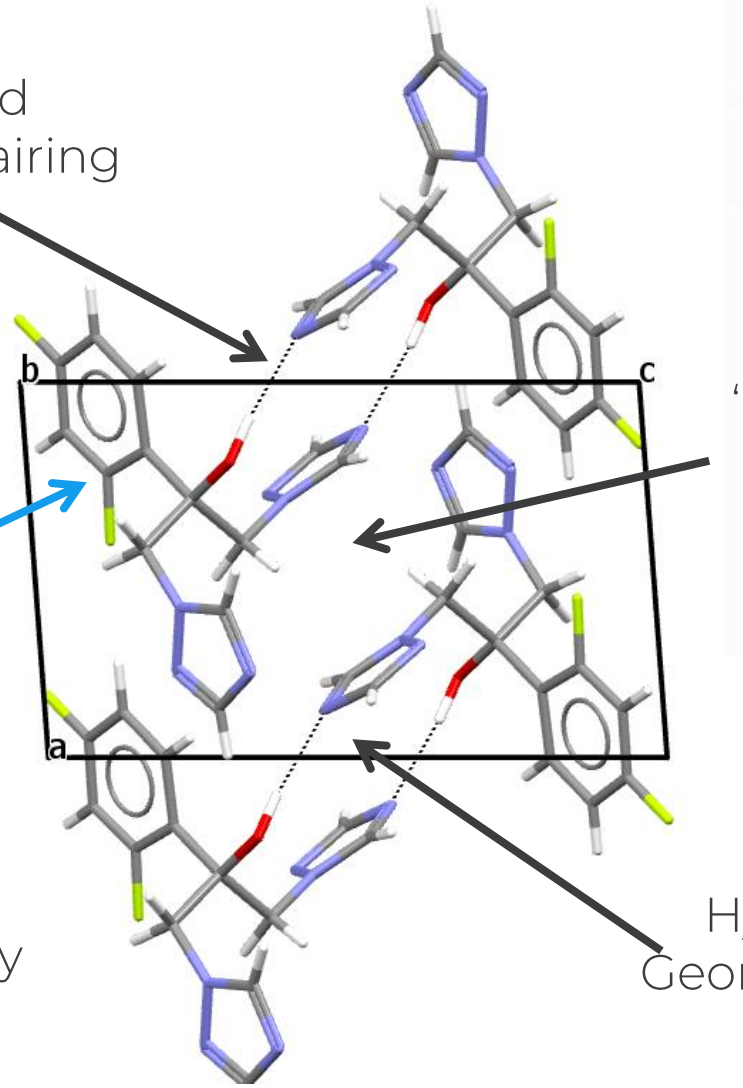
Robert Docherty *et al Journal of Pharmacy and Pharmacology*, (2015) 67, 857. doi: 10.1111/jphp.12394.
Taylor *et al. J. Chem. Inf. Model.*, (2014) 54 (9), 2500. Wood, P. A. *et al. CrystEngComm* (2013) 15, 65

CCDC

Characteristics that influence stability

20

Hydrogen Bond
Donor/Acceptor Pairing



'Non-Hydrogen Bond'
Intermolecular
Interactions

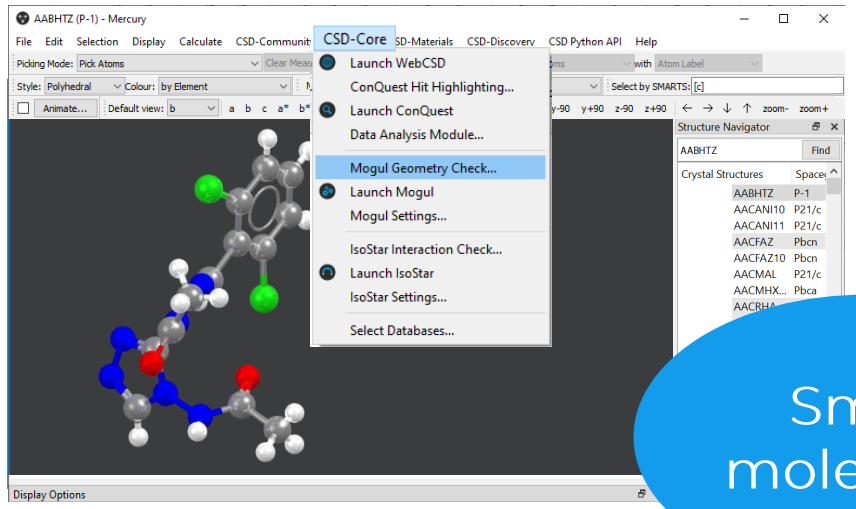
Molecular
Conformation

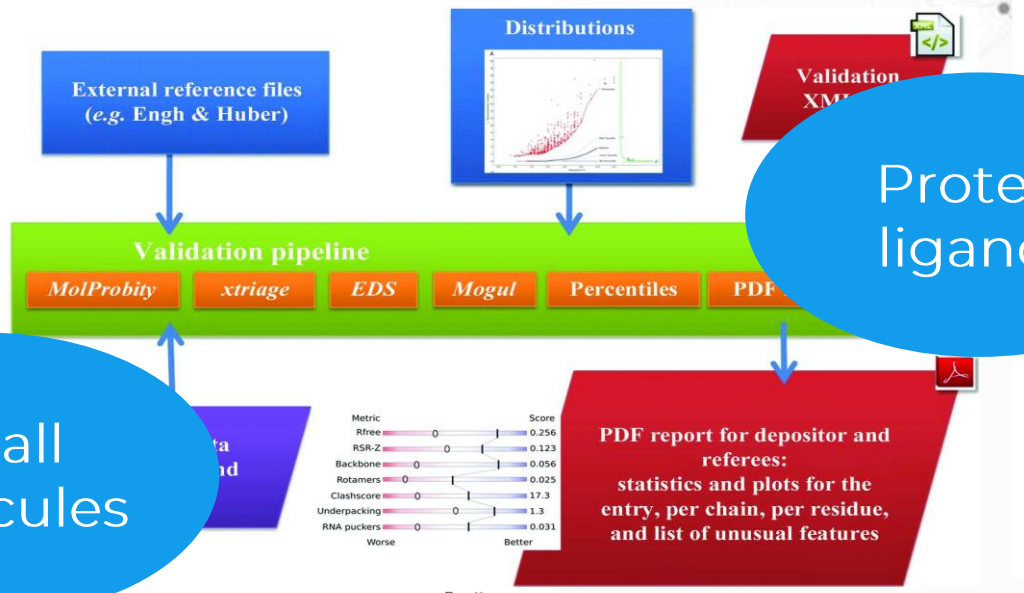
Hydrogen Bond
Geometry, Symmetry
and Motif

We can mine the CSD to identify these intra and intermolecular geometric preferences.

CCDC

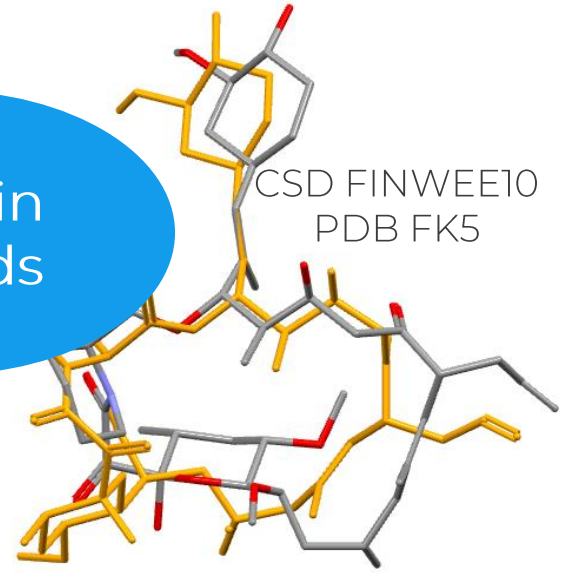
Using Mogul to check new structural data

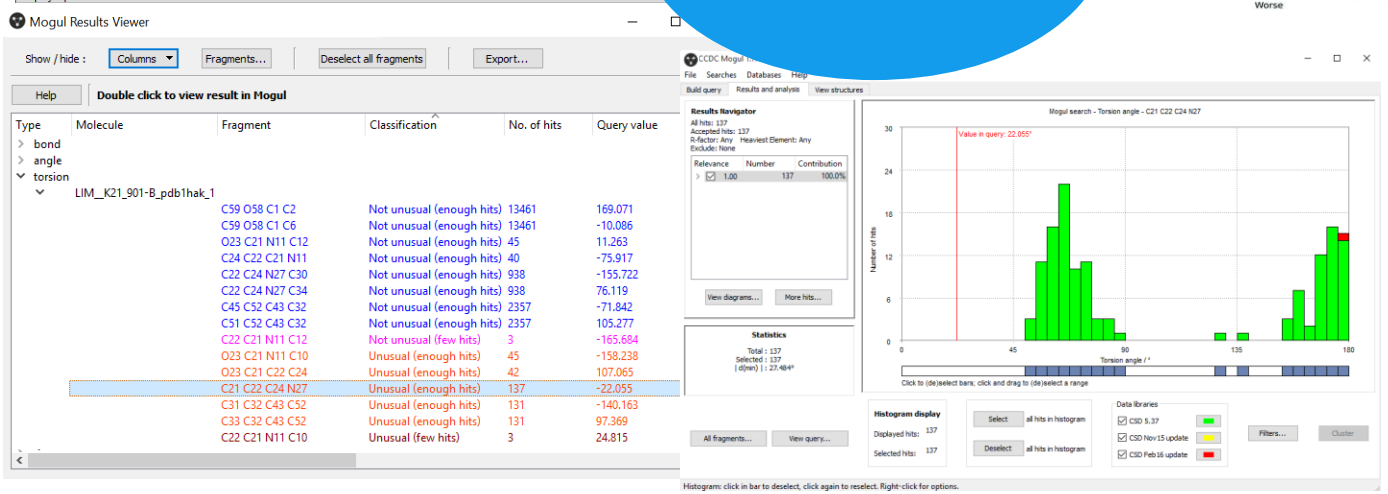


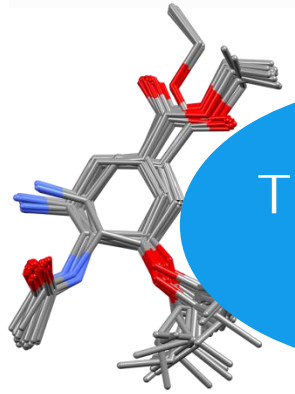


Small molecules

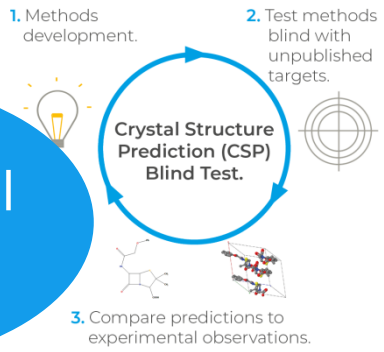
Protein ligands







Theoretical data

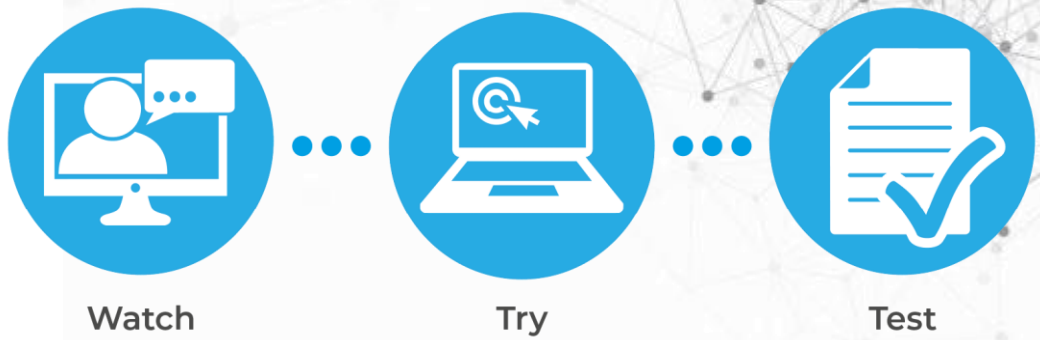


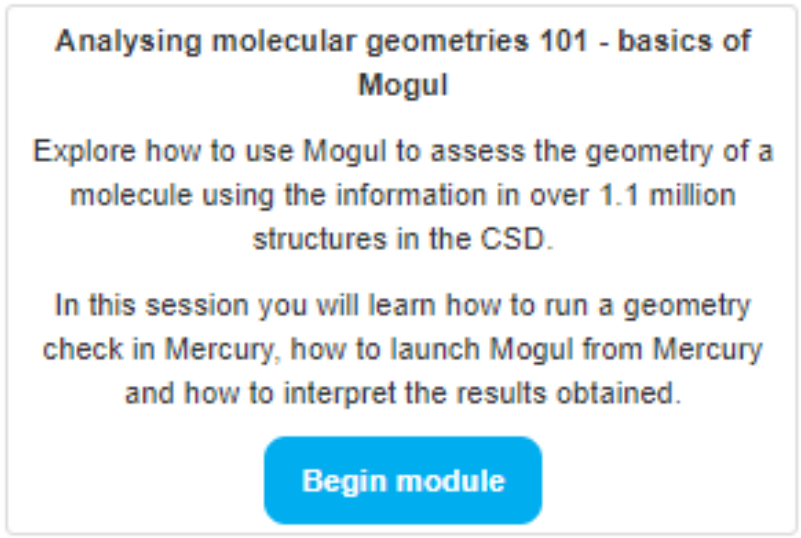
I.J.Bruno, G.P.Shields, R.Taylor, *Acta Cryst.* (2011). B67, 333-349 DOI: 10.1107/S0108768111024608

CCDC

Mogul – the basics

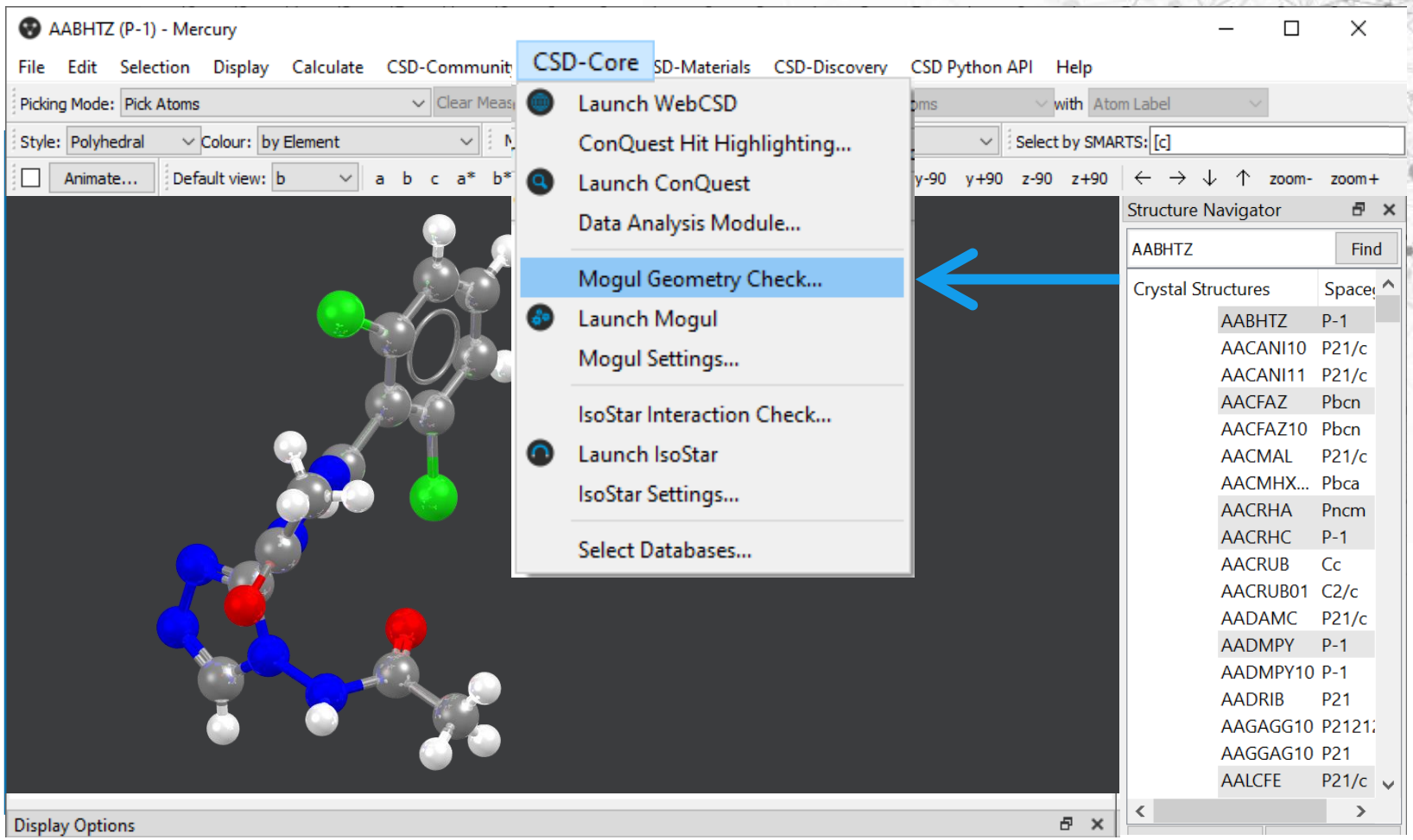
- Covered in:
  - July CCDC Virtual workshop series
  - Now available on demand in CSDU



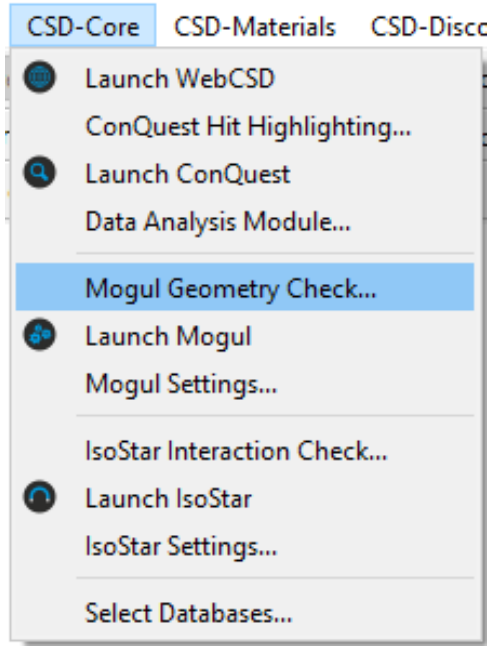


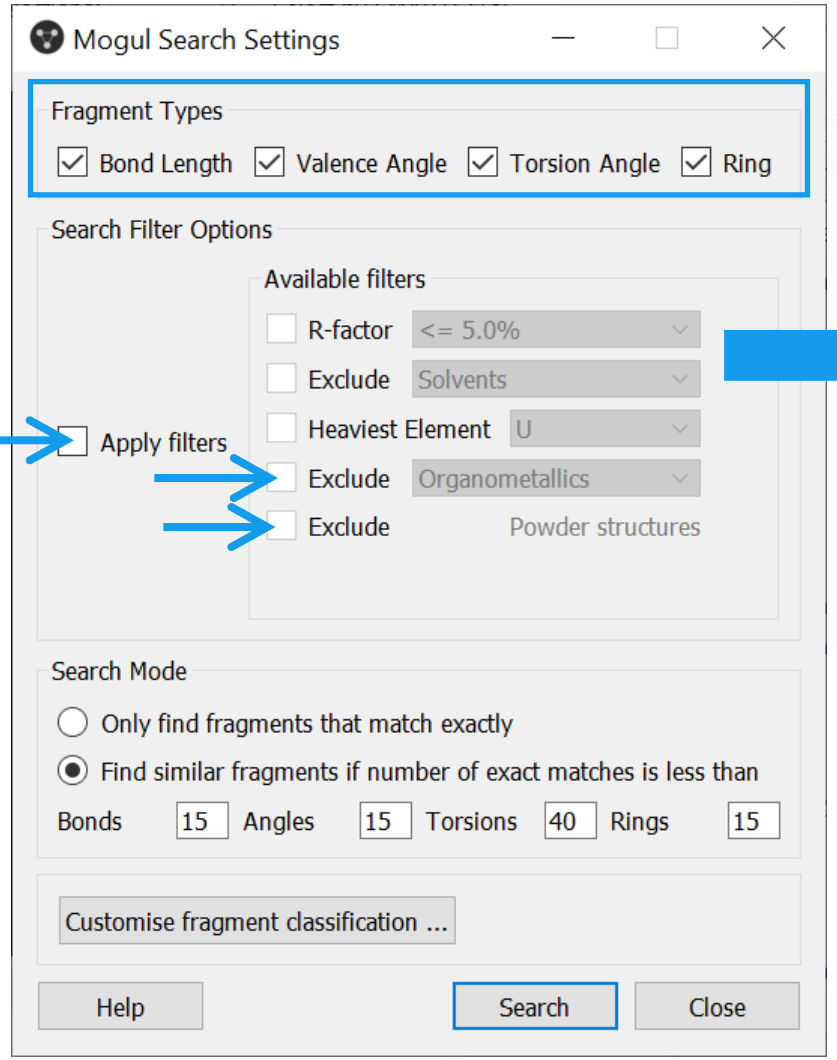
https://www.ccdc.cam.ac.uk/Community/educationalresources/CSDU/

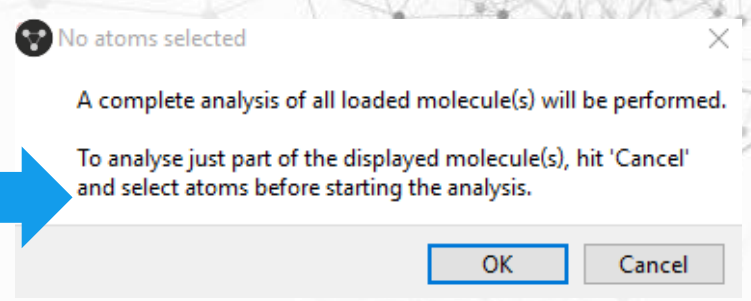
Mogul – Launch from Mercury

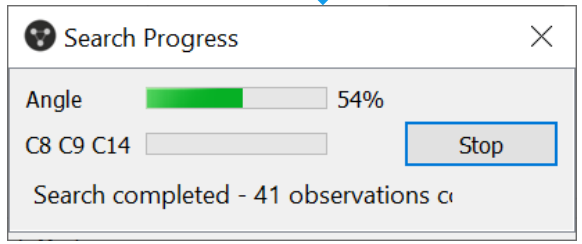


Mogul Geometry Check

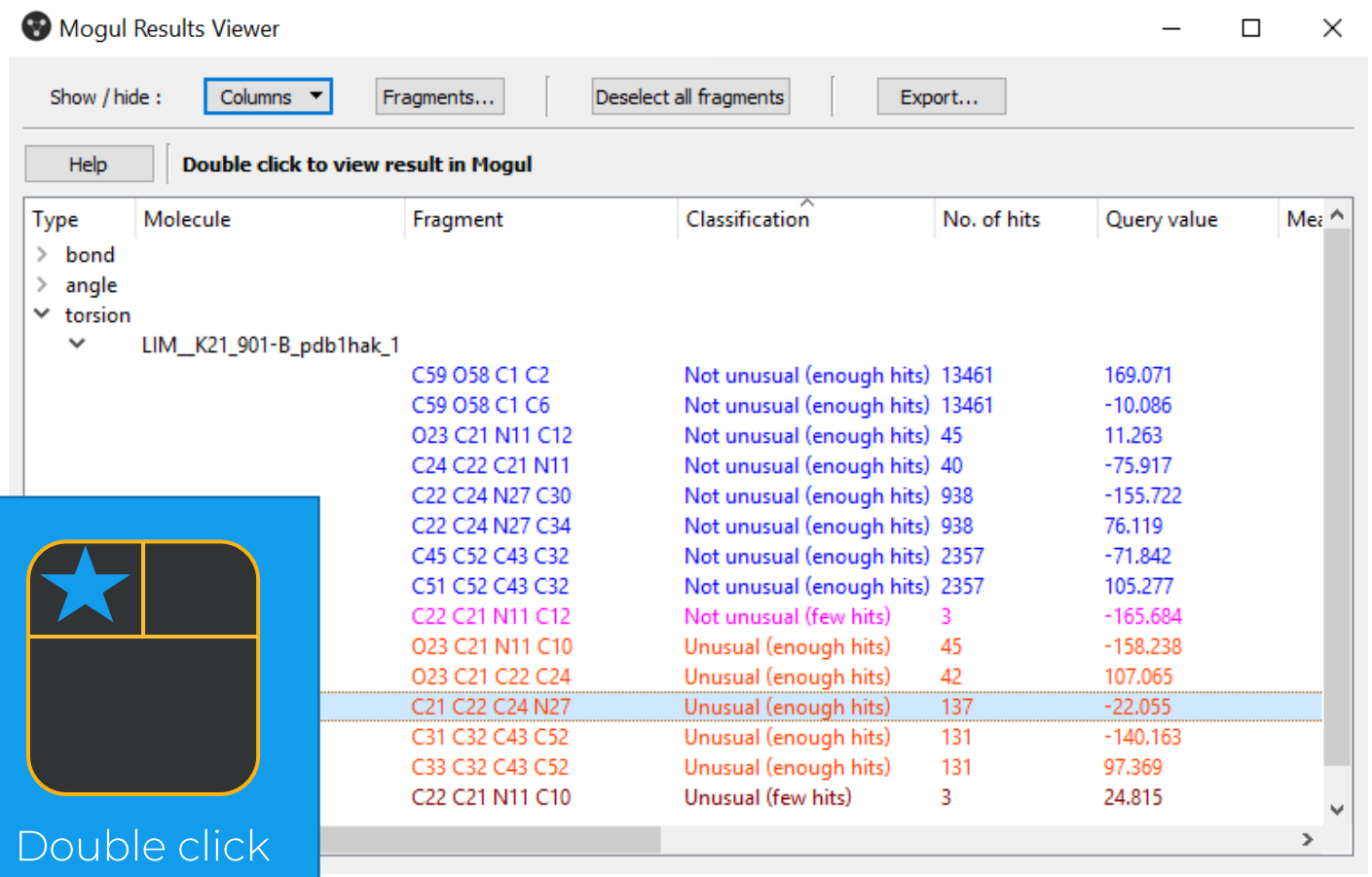








Mogul Geometry Check - Results

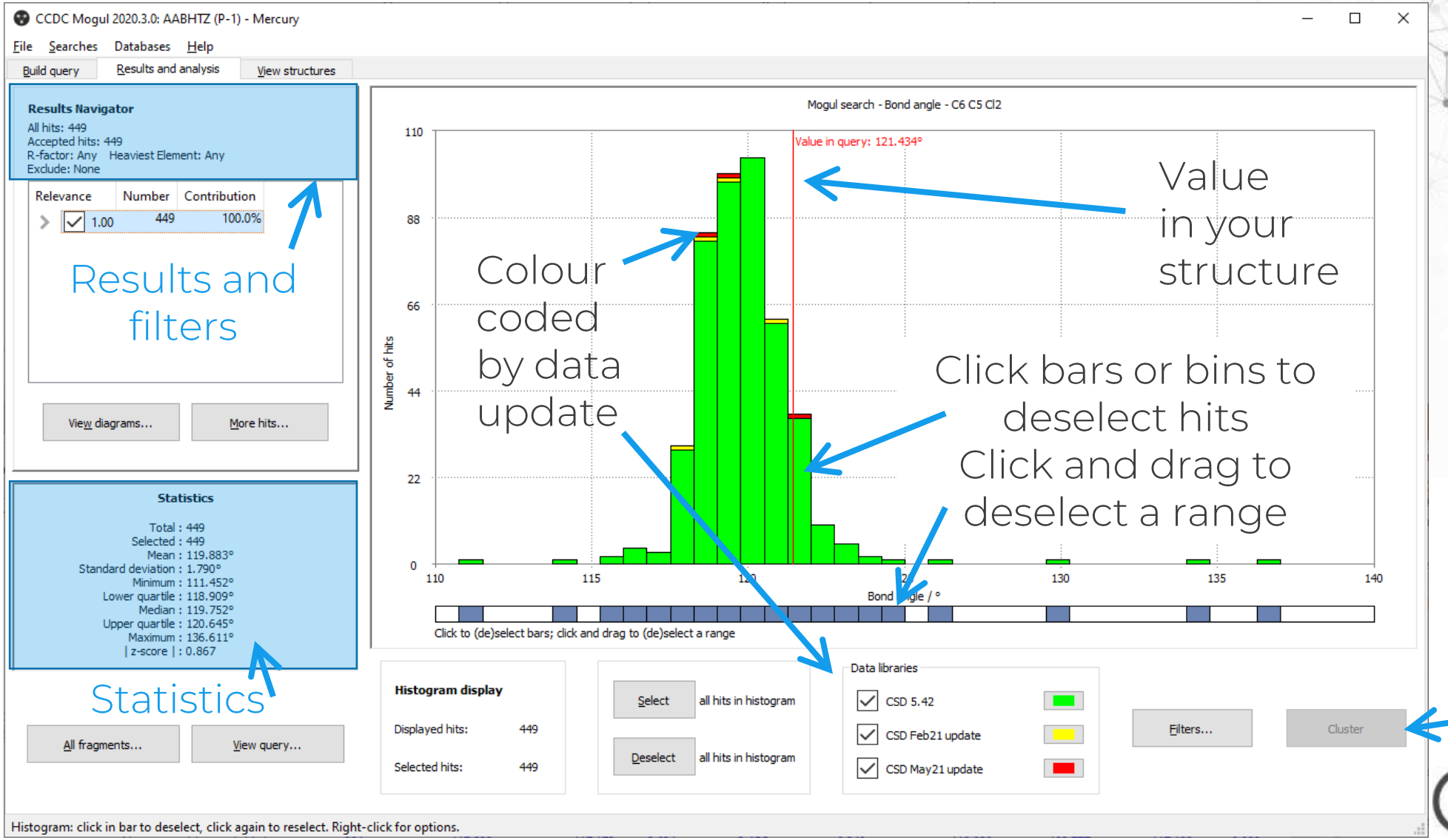


Double click on a row to view result in Mogul

Colour code:

- **Blue**: It is a **not unusual** value in the CSD and there are **enough data** in the CSD for comparison.
- **Pink**: It is a **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

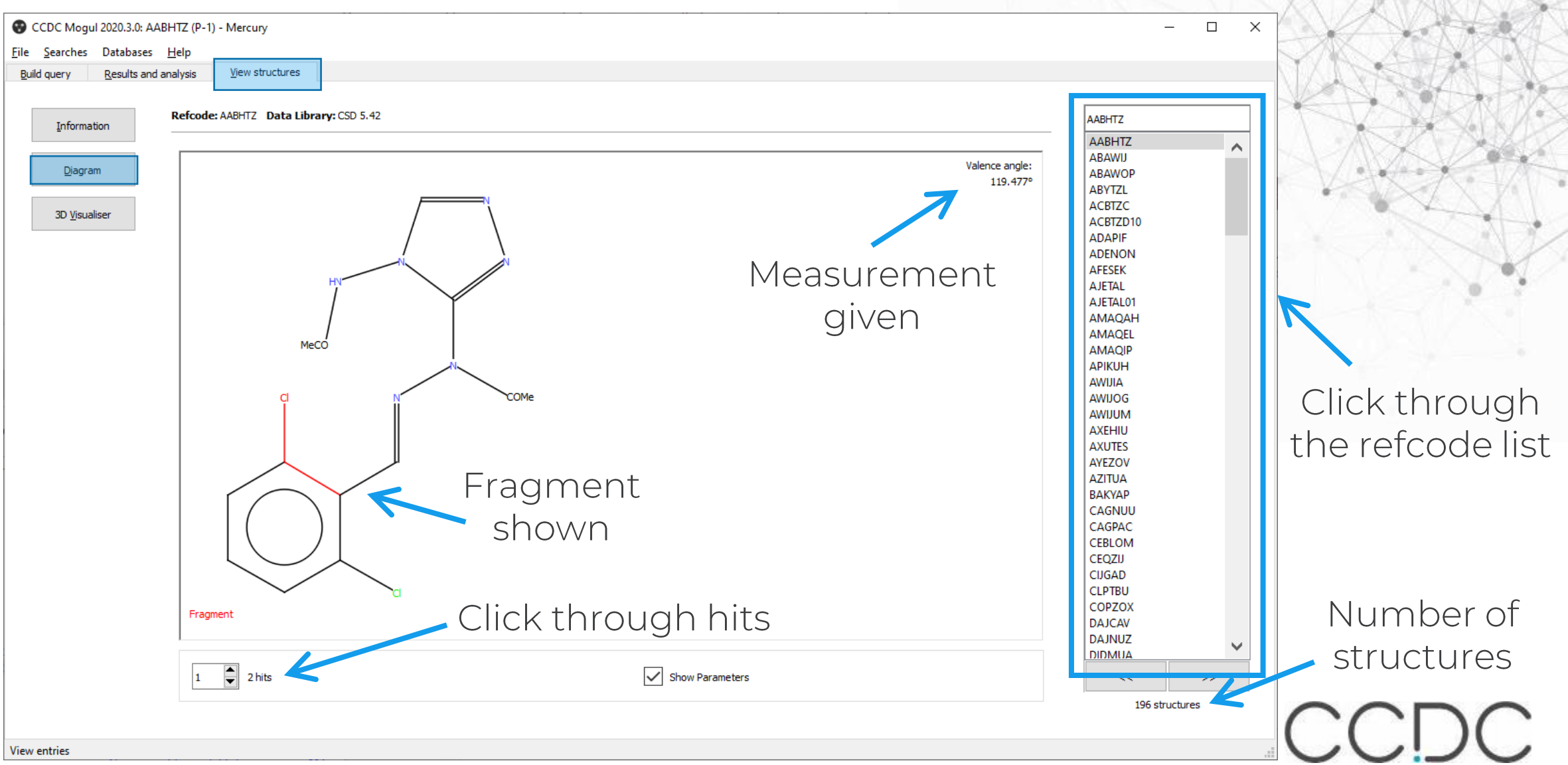


Histogram built with data from structures in the CSD

Cluster only available for rings

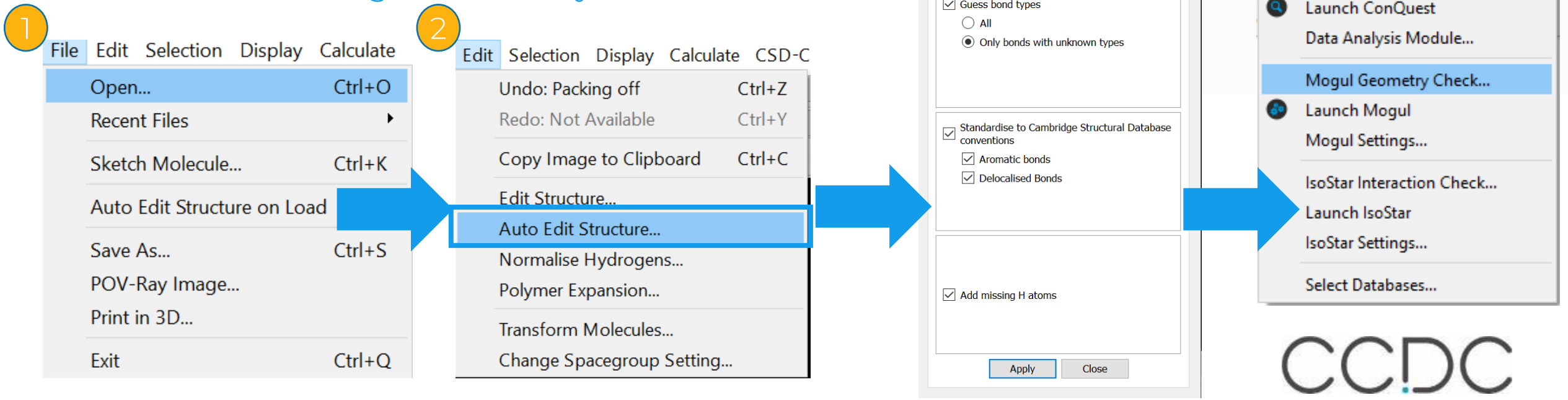
CCDC

Mogul Geometry Check – View structures



Using Mogul to check a new structure

1. Load your structure in Mercury
   - File > Open – to open one of your own files for example a CIF
2. Don't forget to assign bond types
   - Edit > Auto Edit Structure to assign bond types
   - Or File > Auto Edit Structure on Load to automate
3. Run the Mogul Geometry Check.....

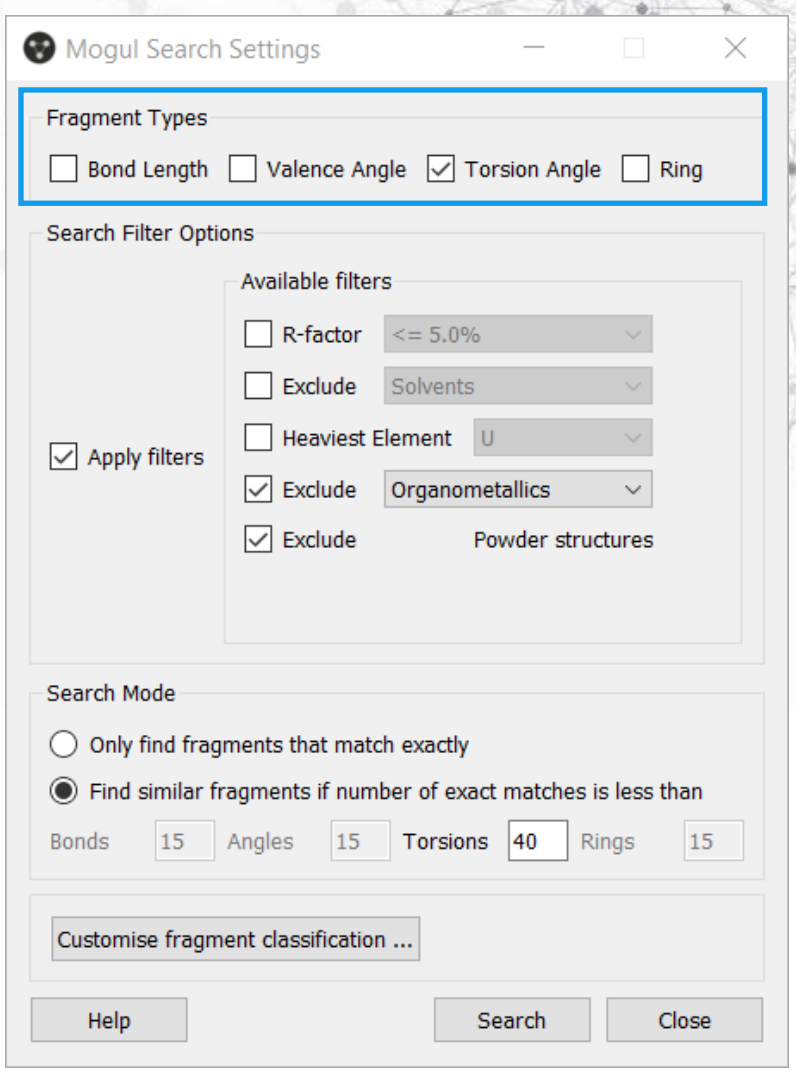


CCDC

Intermediate Mogul

- Focus on conformation
  - Torsion angles

- Local density
  - Percentage of experimental measurements that are within *x* degrees of the value in the query molecule
  - Torsion angle classified as unusual if local density < 0.05 (i.e. less than 5 % of experimental values are within 10° of the value in query molecule)



CCDC

Example #1 (COBFIK)

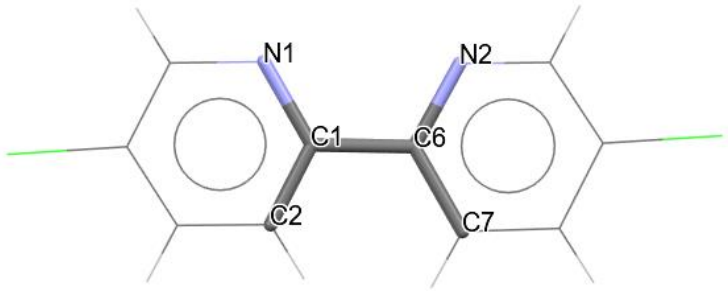
- COBFIK, a bipyridine derivative (5,5'-dichloro-2,2'-bipyridine)
  - 1 rotatable bond, with all 4 torsion angles returned as unusual

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) |	Local density
torsion														
	COBFIK													
		N1 C1 C6 N2	Unusual (enough hits)	2192	0.148								0.000	**0.024**
		C2 C1 C6 N2	Unusual (enough hits)	4023	179.113								0.000	**0.025**
		C7 C6 C1 N1	Unusual (enough hits)	4023	-179.327								0.000	**0.025**
		C2 C1 C6 C7	Unusual (enough hits)	2226	-0.362								0.000	**0.023**

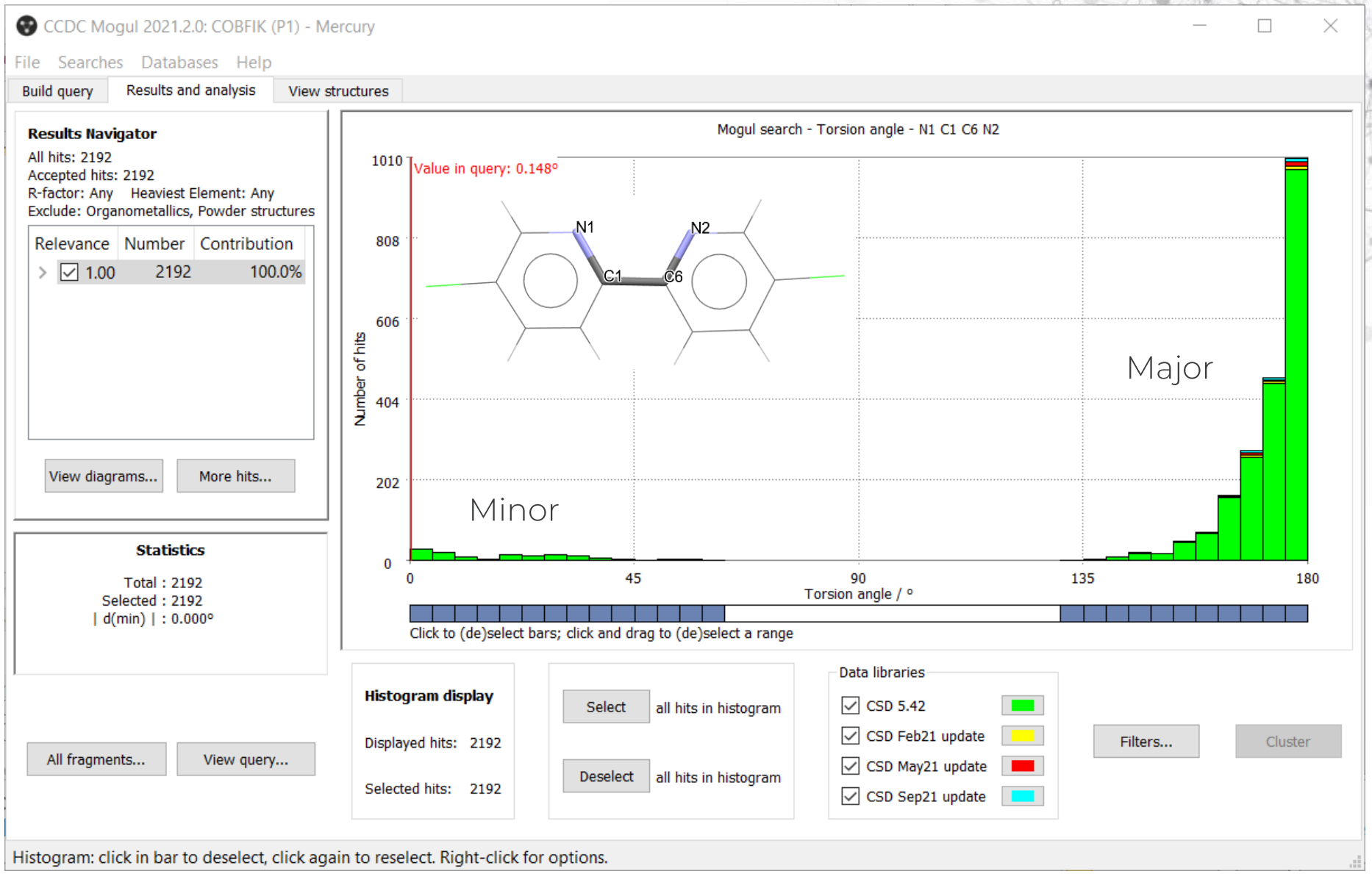


Mogul settings: No organometallics, no powder structures

CCDC

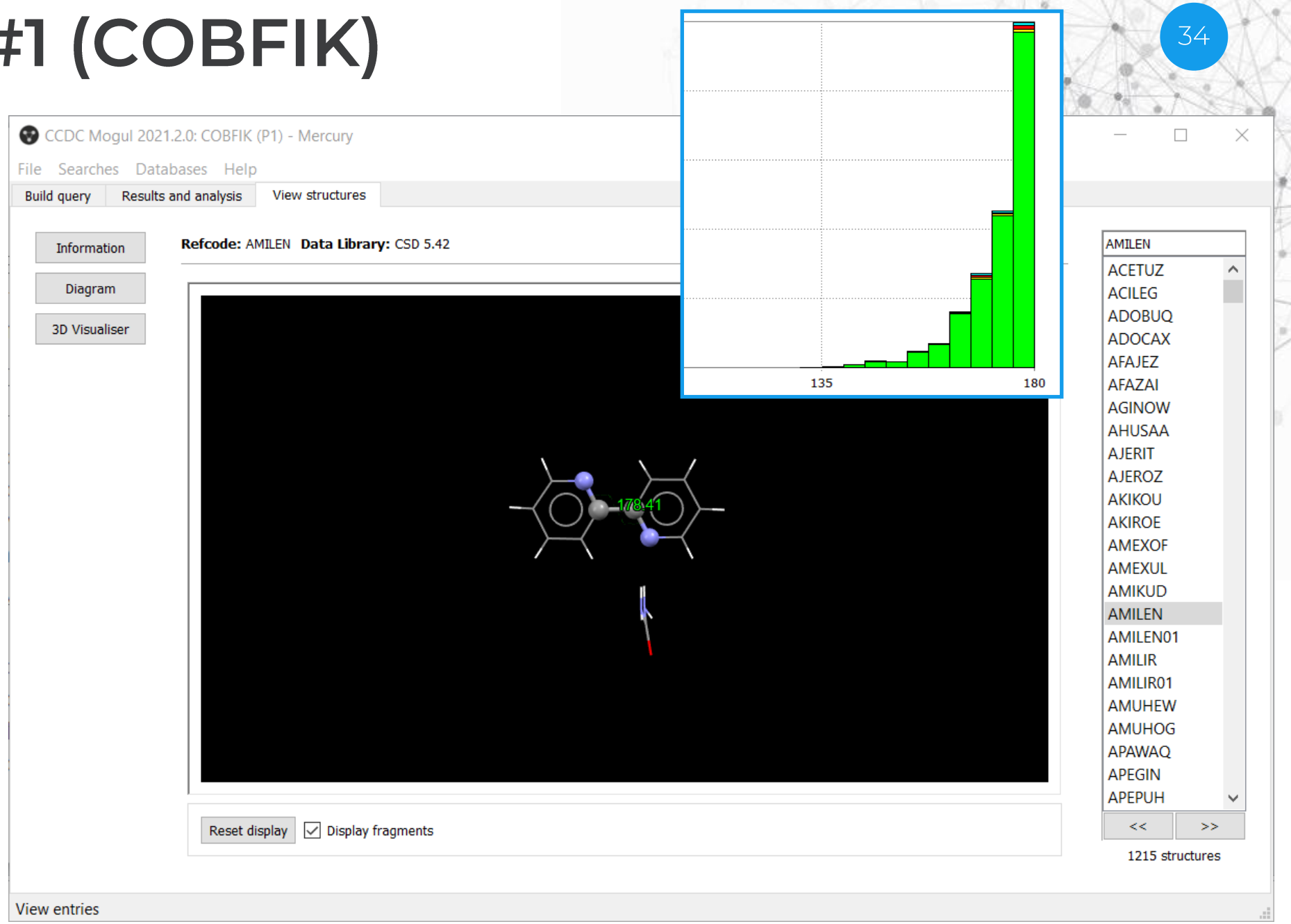
Example #1 (COBFIK)

- Bimodal distribution
- Torsion angles define relative orientation of the two nitrogen atoms
- Major population towards 180° and minority population towards 0°
- Observed torsion angle value in minor population (125/2192 hits)



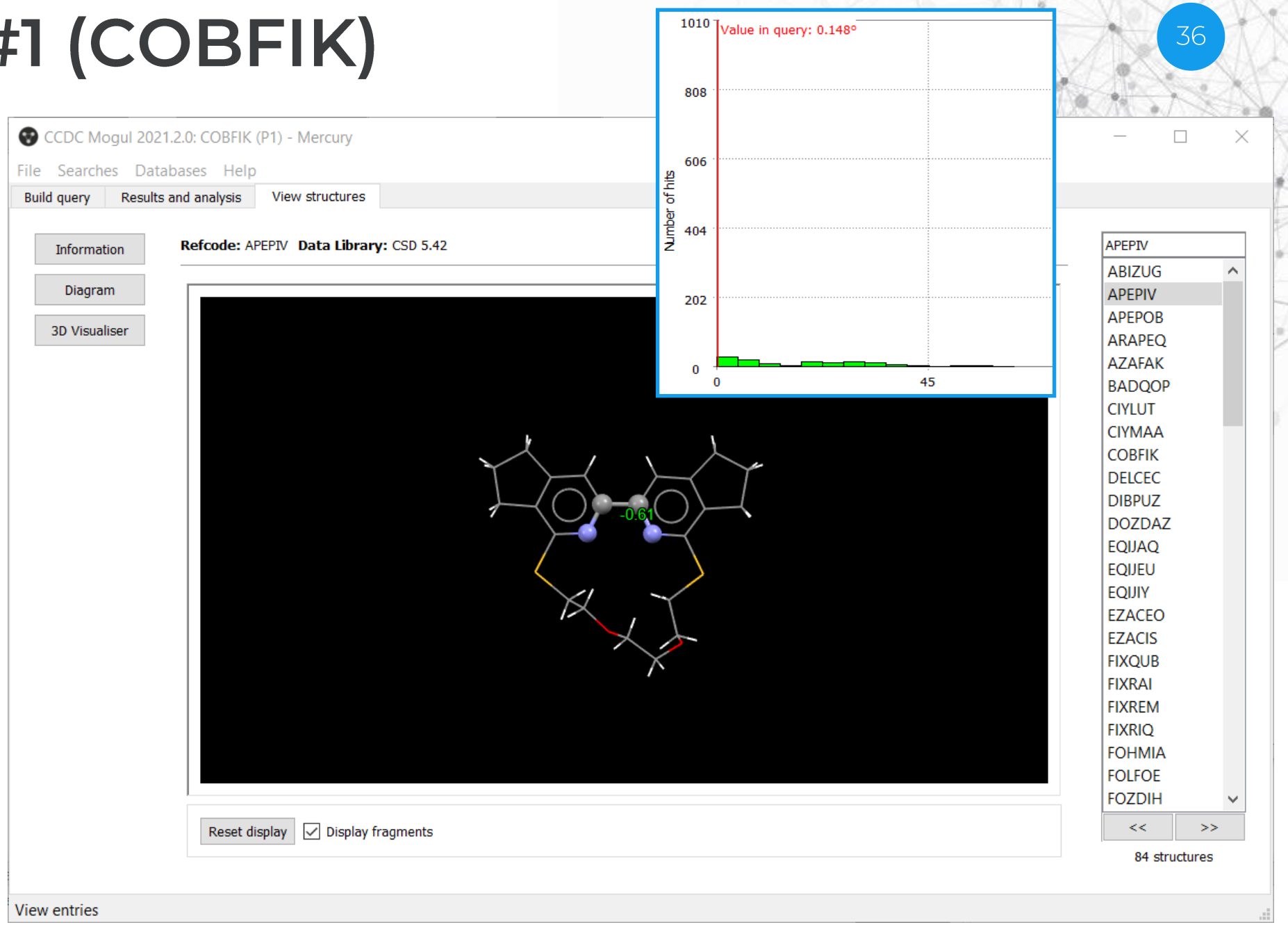
Example #1 (COBFIK)

- Exploring the structures in major population
- Over 2000 hits (2067, 94.3 % of distribution)
- All structures have bipyridine nitrogens closer to *trans* to one another (facing in opposite directions, 180°)



Example #1 (COBFIK)

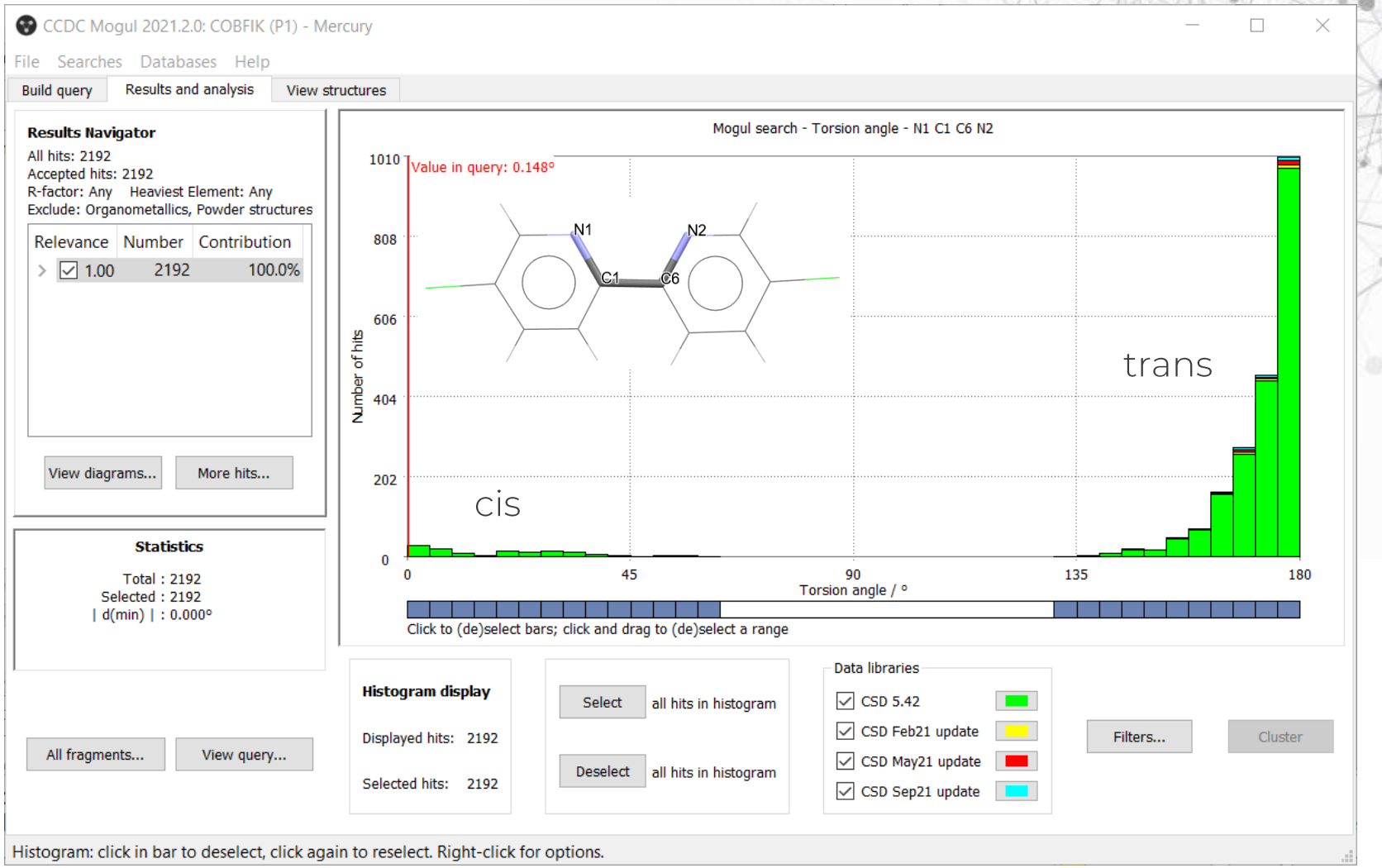
- Exploring the structures in minor population
- Only 125 hits (5.7 %)
- Bipyridine nitrogens closer to *cis* to one another (facing in same direction, 0°)
- Many (~70 %) constrained as part of a large ring (macrocycle), which may explain *cis* adoption



Example #1 (COBFIK)

Summary:

- Very unusual torsion angles with *cis* conformation
- Mogul shows a preference for *trans* arrangement of bipyridine nitrogens
- *cis* substantially less stable than *trans* for this molecule (energy calculations)*
- Unusual bond lengths and valence angles for one of the aromatic rings + ADPs look strange. Structure solution possibly incorrect?



CrystEngComm, 2020, 22, 7217–7228. DOI: 10.1039/c9ce02001b

Example #2 (BOLLUK)

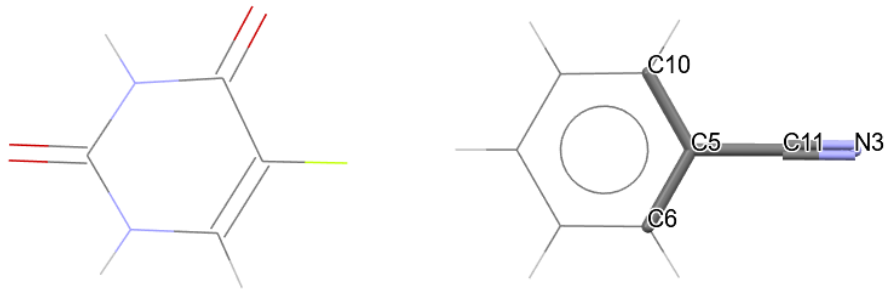
- BOLLUK (5-fluorouracil benzonitrile)
  - 2 torsion angles returned as slightly unusual

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) |	Local density
torsion														
	BOLLUK_1													
	BOLLUK_2													
		C6 C5 C11 N3	Unusual (enough hits)	187	-61.594								0.694	**0.043**
		C10 C5 C11 N3	Unusual (enough hits)	187	117.104								0.724	**0.043**

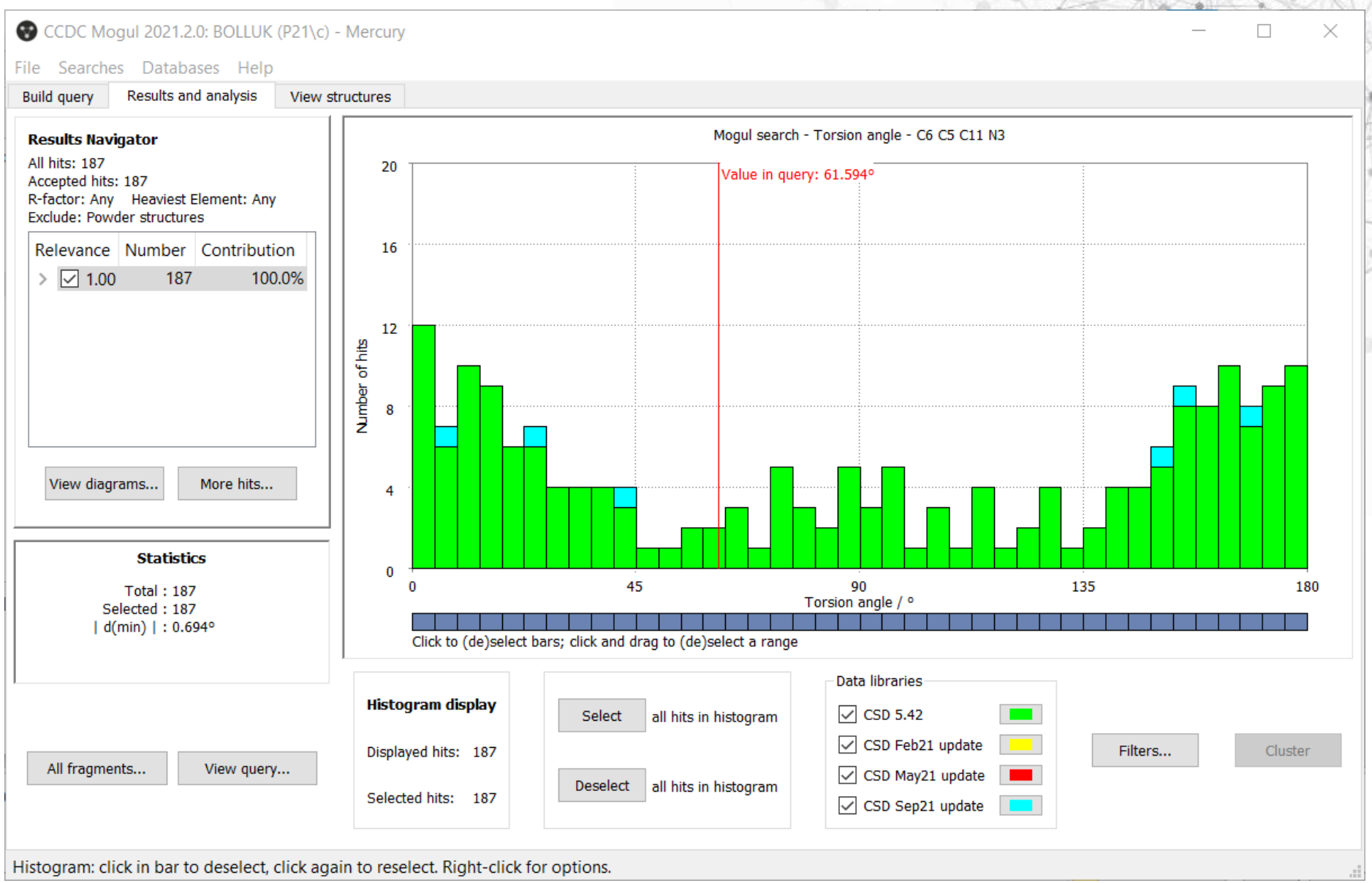


Mogul settings: No powder structures

CCDC

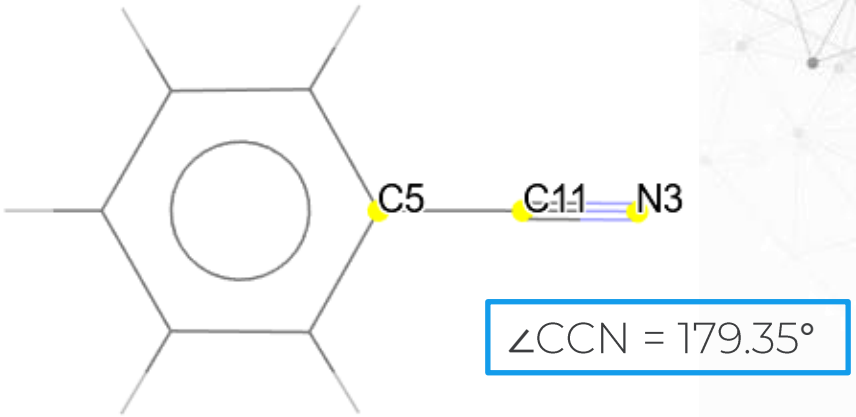
Example #2 (BOLLUK)

- Spread of torsion values across the whole degree range, but....



Example #2 (BOLLUK)

- Molecule with ‘unusual’ torsion angles has a cyano group C≡N (carbon to nitrogen triple bond)
- C-C≡N angle is typically quasi-linear (180°)
- Adjacent C-C bond is not counted as a typical rotatable bond
- ‘Not unusual/unusual’ definition is not as meaningful for nitriles (cyano-containing compound)
- Looking at the hits shows very little difference in conformation around the aromatic nitriles



Summary:

- Ignore nitrile (cyano group) results

Example #3 (RENJUR)

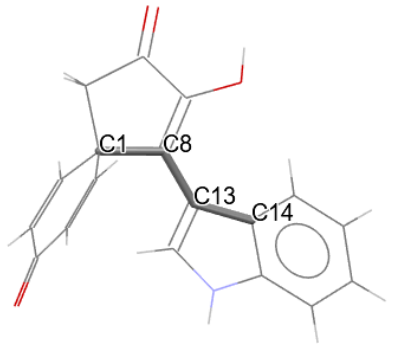
- RENJUR (Spirobacillene A)
  - 1 rotatable bond between two five-membered rings
  - 1 of 4 torsion angles returned as unusual

Mogul Results Viewer

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

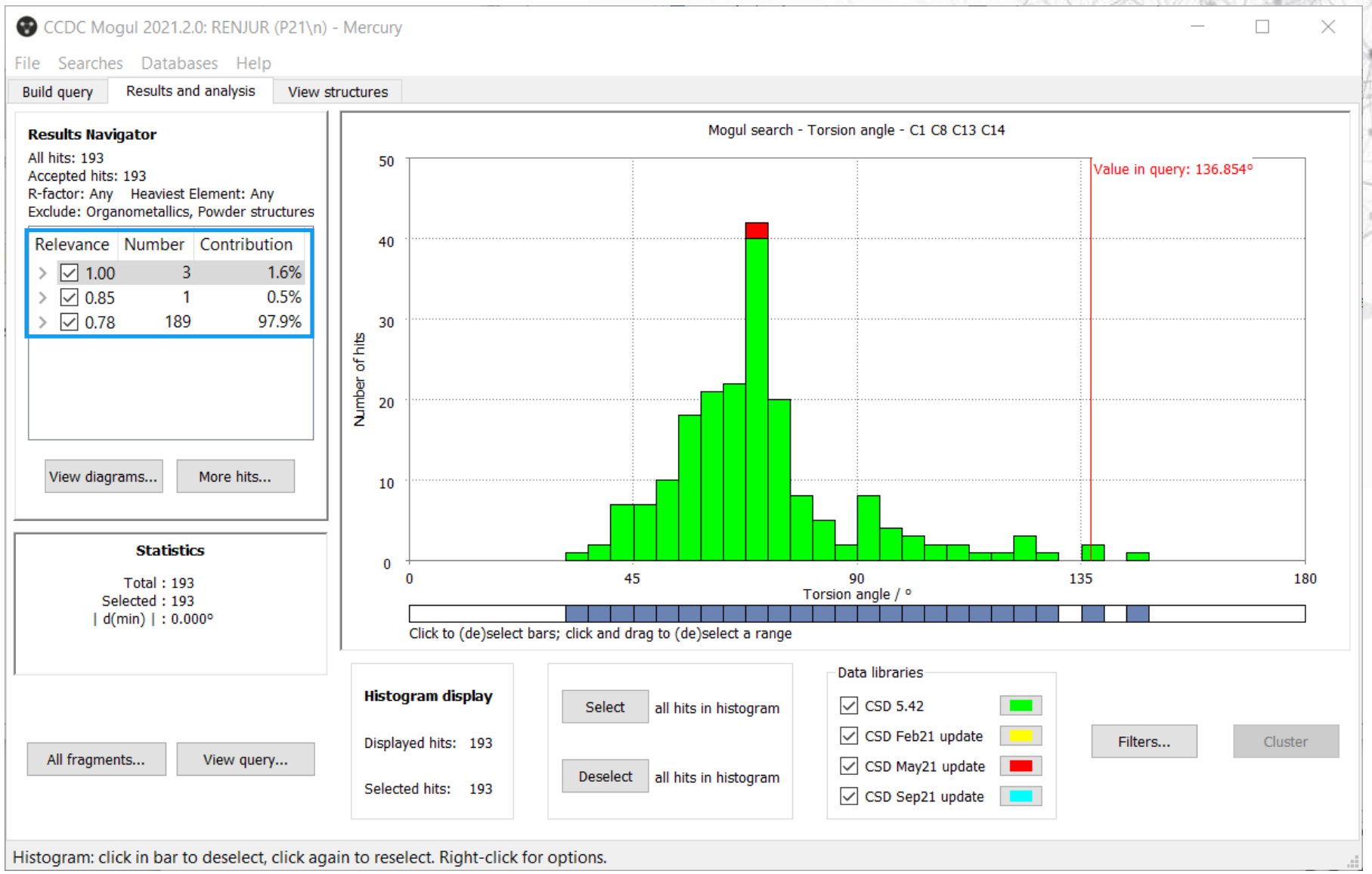
Help | **Double click to view result in Mogul**

Type	Molecule	Fragment	Classification	No. of hits	Query value	Mean	Std. dev.	| z-score |	| x - mean |	Minimum	Maximum	Median	| d(min) |	Local density
torsion														
	RENJUR													
		C11 C13 C8 C7	Not unusual (enough hits)	70	-146.184								0.000	**0.386**
		C11 C13 C8 C1	Not unusual (enough hits)	36	40.278								0.000	**0.278**
		C14 C13 C8 C7	Not unusual (enough hits)	81	36.684								0.000	**0.333**
		C1 C8 C13 C14	Unusual (enough hits)	193	-136.854								0.000	**0.016**



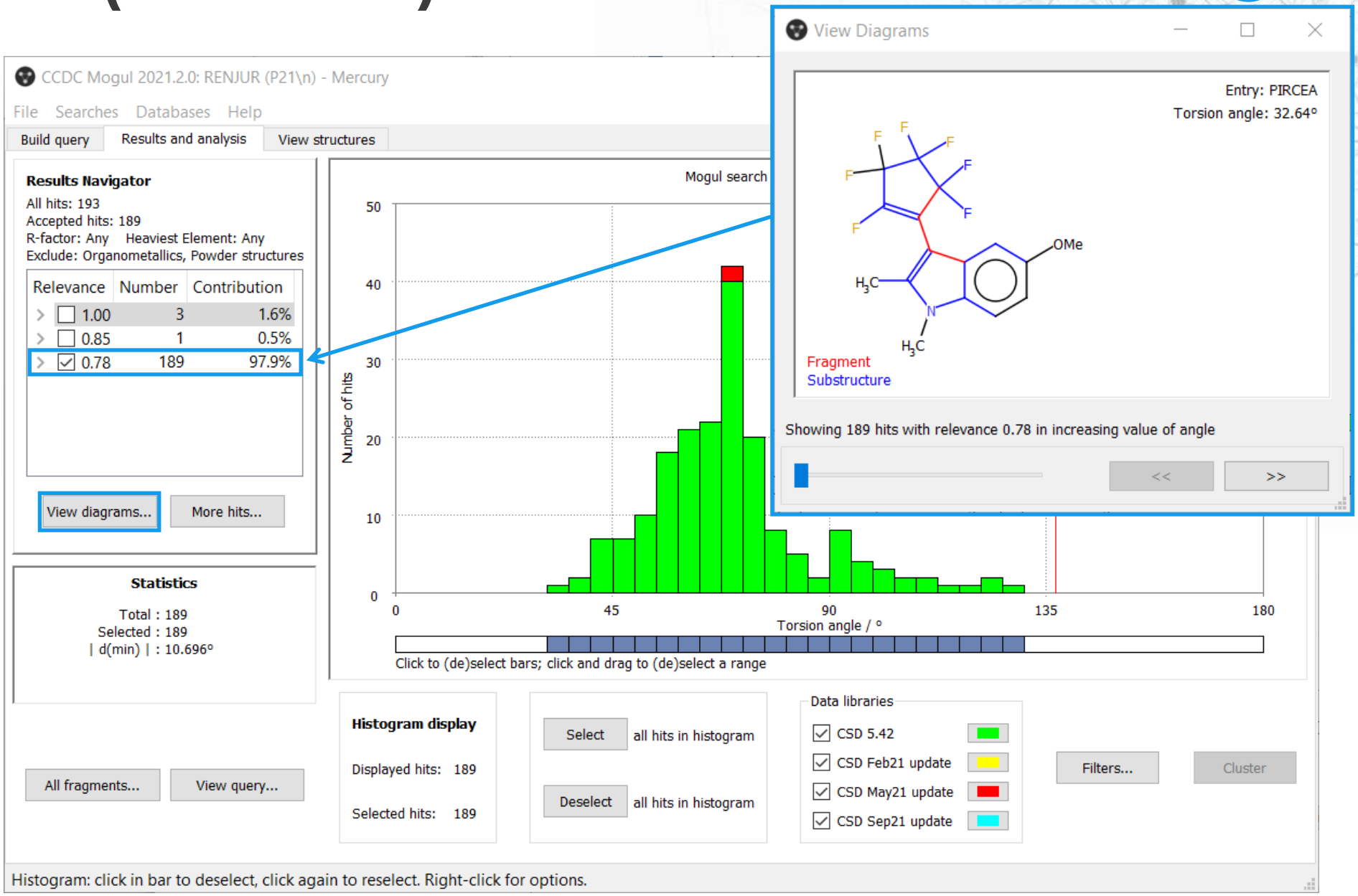
Example #3 (RENJUR)

- Observed value appears unusual, lying in extreme tail of distribution
- Structures with lower relevance are included with default settings (find ≥ 40 hits)



Example #3 (RENJUR)

- Exploring lower relevance structures
- Majority are fluorinated around 5-membered ring
- Few more are constrained as part of macrocycle
- Fluorinated compounds known to show different conformational preferences to non-fluorinated compounds

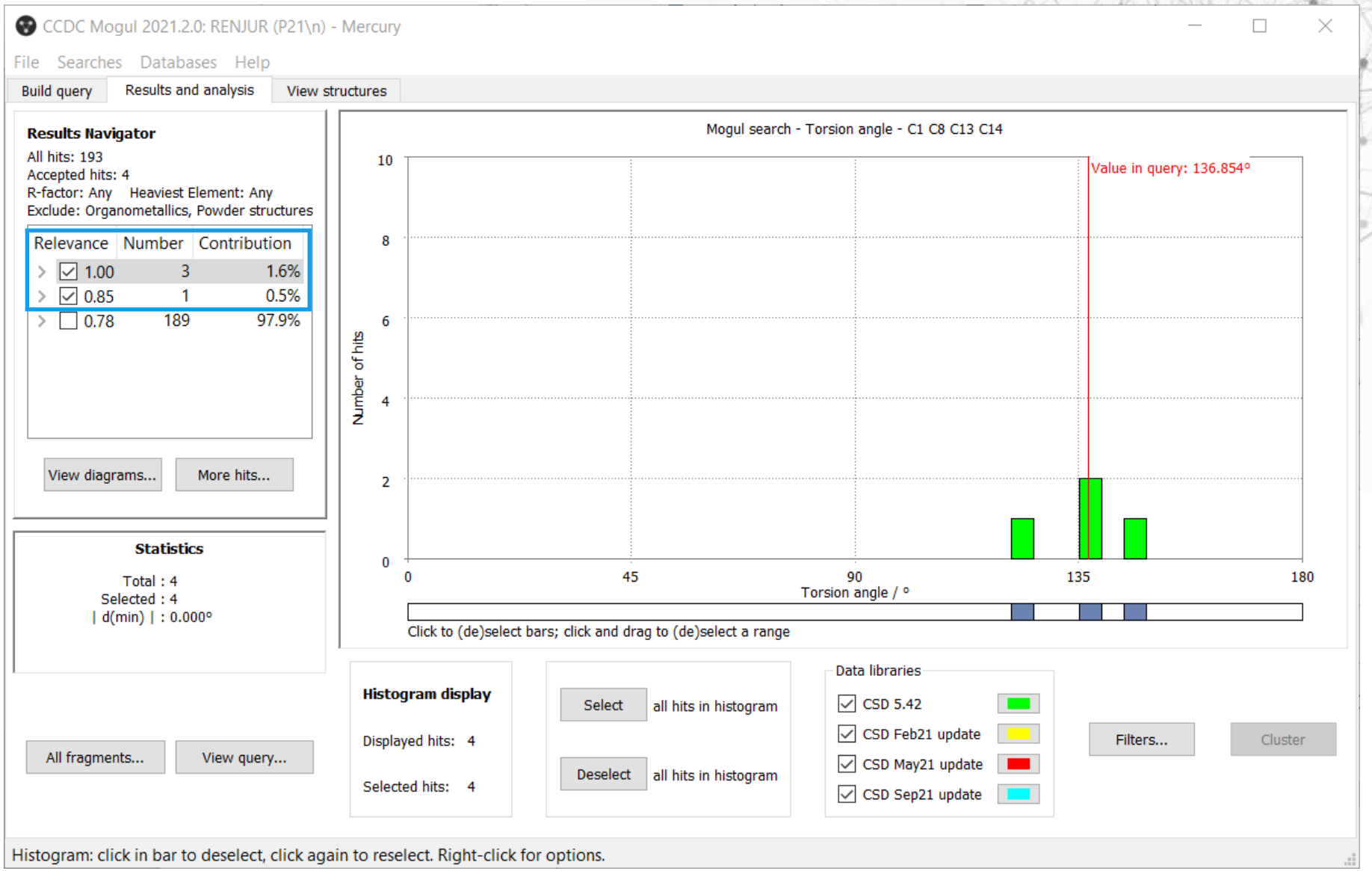


Example #3 (RENJUR)

- Only considering high relevance excludes fluorinated structures
- Hits localised around observed torsion value

Summary:

- Torsion angle in RENJUR not unusual given context (though only low number of structures to compare with)



Example #4 (ADELUR)

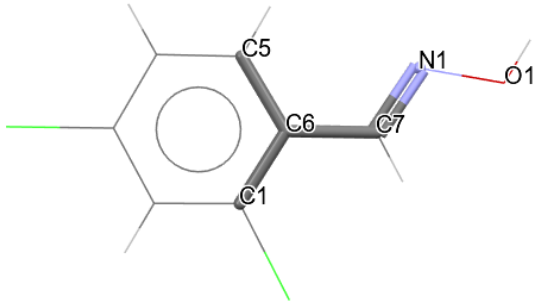
- ADELUR ((E)-2,4-Dichlorobenzaldehyde oxime)
  - No unusual torsion angles

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) |	Local density
torsion														
	ADELUR													
		C6 C7 N1 O1	Not unusual (enough hits)	330	179.286								0.000	**0.967**
		C1 C6 C7 N1	Not unusual (enough hits)	170	162.701								0.000	**0.529**
		C5 C6 C7 N1	Not unusual (enough hits)	250	-18.062								0.000	**0.116**



2 torsion angle definitions for same rotatable bond

Both returned as ‘not unusual’, but….

Example #4 (ADELUR)

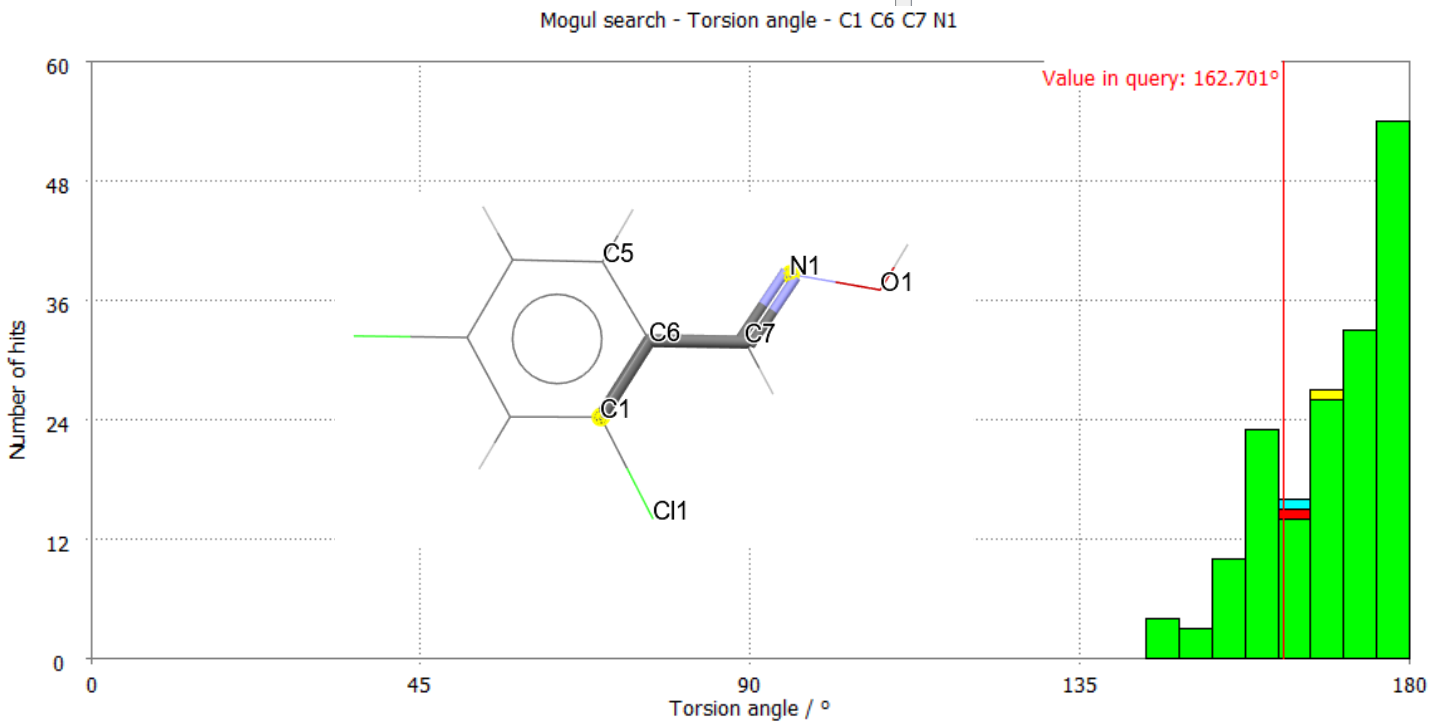
- Torsion angle from N to C with Cl substituent (N1 to C1)
- Single population distribution

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) |	Local density
torsion														
	ADELUR													
		C6 C7 N1 O1	Not unusual (enough hits)	330	179.286								0.000	**0.967**
		C1 C6 C7 N1	Not unusual (enough hits)	170	162.701								0.000	**0.529**
		C5 C6 C7 N1	Not unusual (enough hits)	250	-18.062								0.000	**0.116**



- All hits for this torsion have Cl substituent on the carbon atom
- Always *trans* arrangement between this C and the N, never *cis*

CCDC

Example #4 (ADELUR)

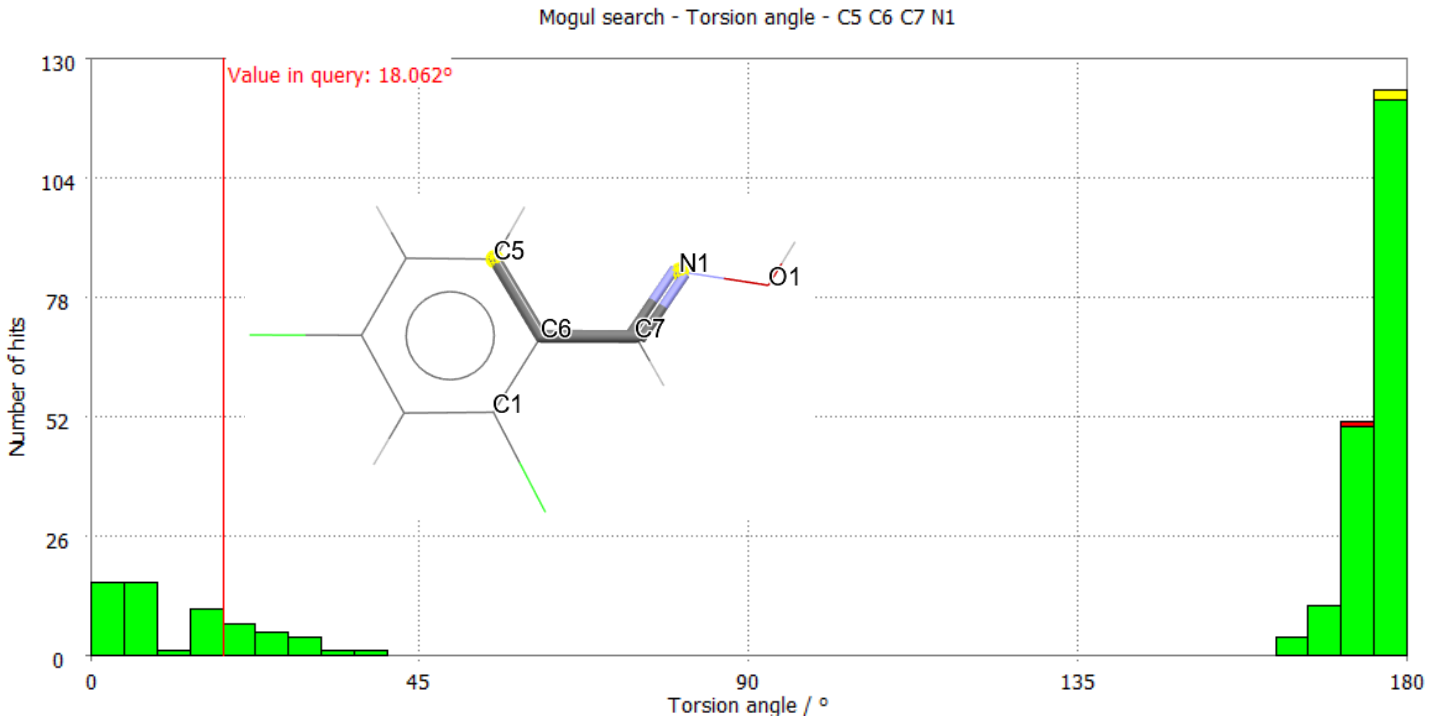
- Torsion angle from N to the other C (N1 to C5)
- Major and minor populations in distribution

Mogul Results Viewer

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

Help | **Double click to view result in Mogul**

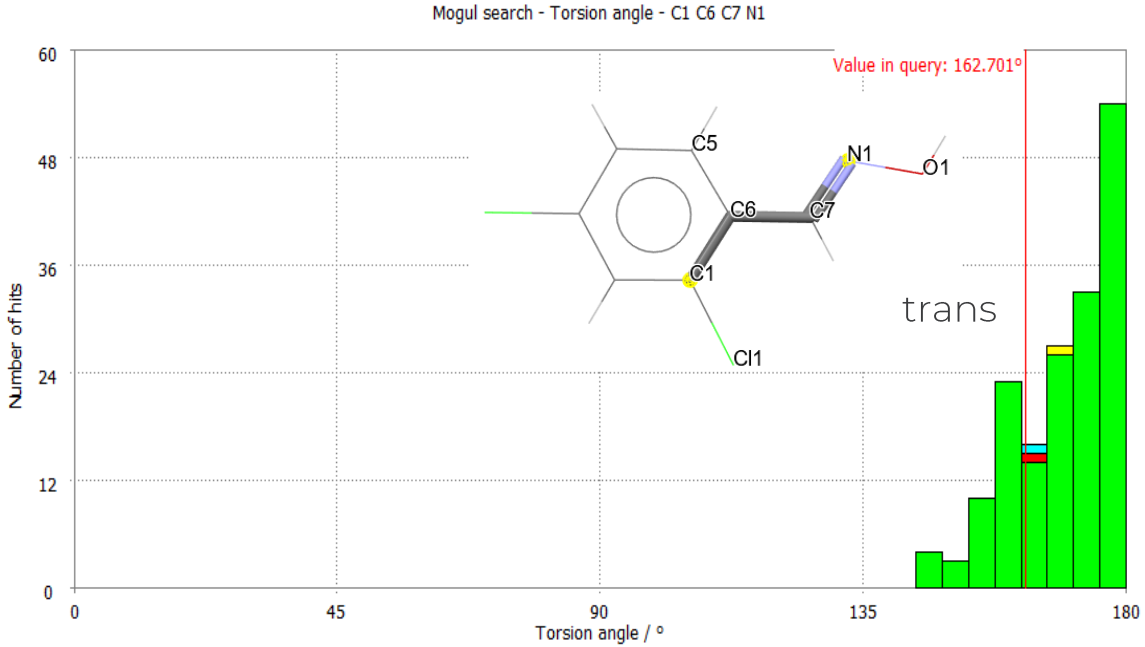
Type	Molecule	Fragment	Classification	No. of hits	Query value	Mean	Std. dev.	| z-score |	| x - mean |	Minimum	Maximum	Median	| d(min) |	Local density
torsion														
	ADELUR													
		C6 C7 N1 O1	Not unusual (enough hits)	330	179.286								0.000	**0.967**
		C1 C6 C7 N1	Not unusual (enough hits)	170	162.701								0.000	**0.529**
		C5 C6 C7 N1	Not unusual (enough hits)	250	-18.062								0.000	**0.116**

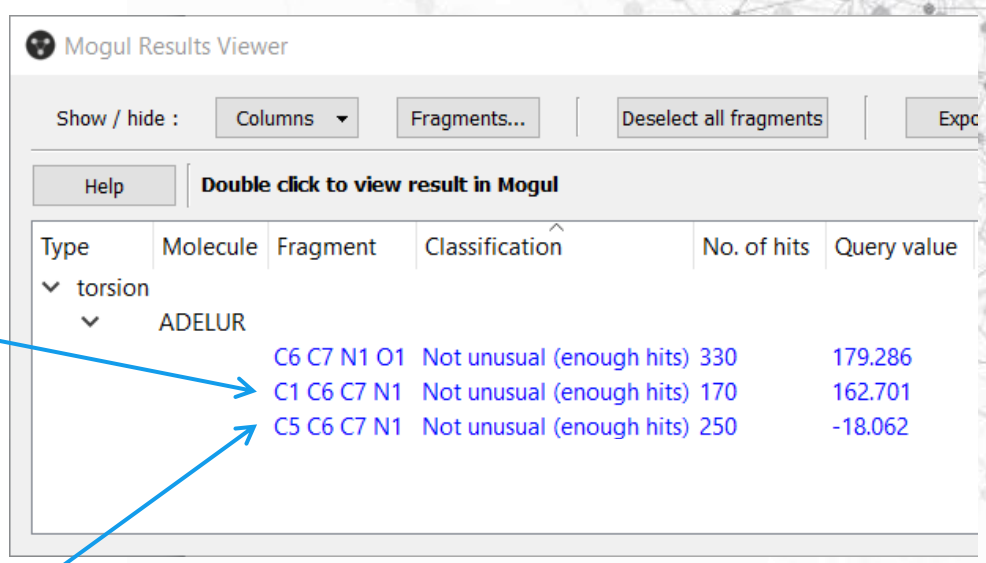


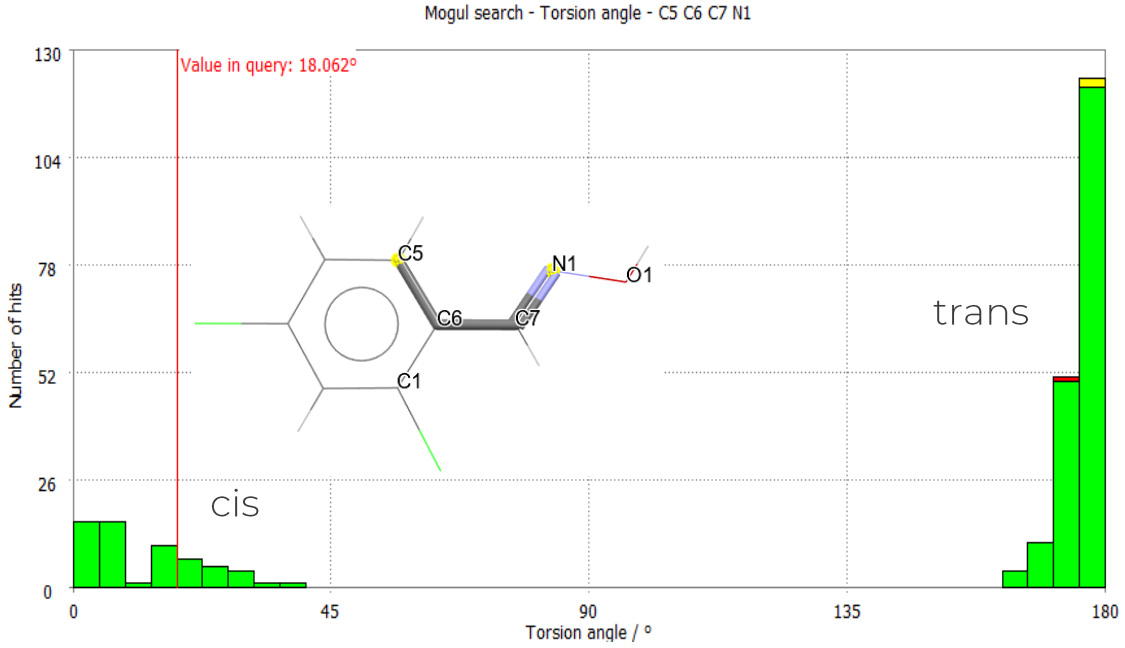
- Observed torsion in minority population (*cis*) – higher energy? *Trans* more usual (major population)
- Always check the distributions, even when reported 'Not unusual'

CCDC

Example #4 (ADELUR)



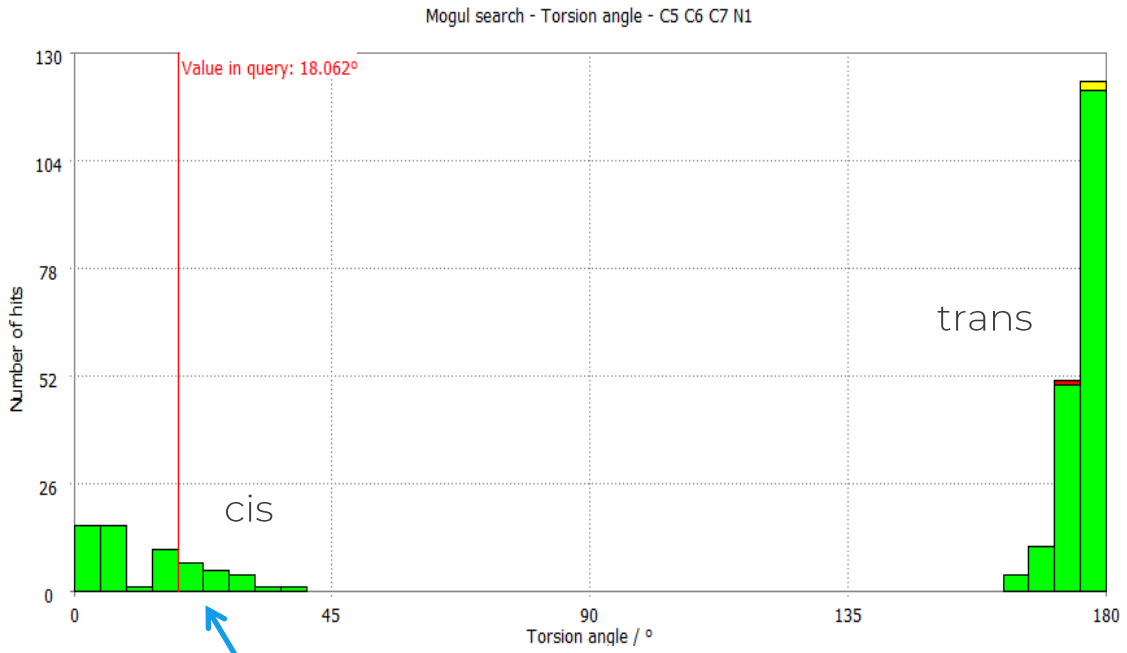




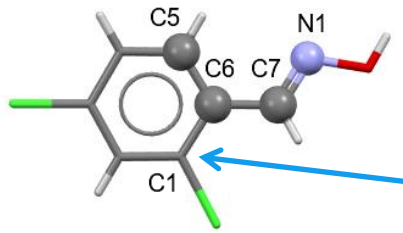
- Histograms for complementary torsions of same rotatable bond suggest that both prefer a *trans* arrangement – not possible
- Only one possibility for N1 to C1 torsion, two possibilities for N1 to C5 torsion….

CCDC

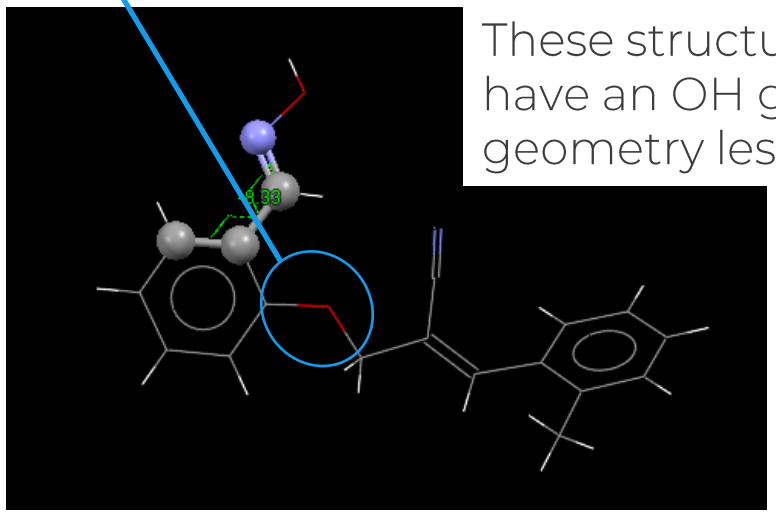
Example #4 (ADELUR)



- Exploring contributing structures

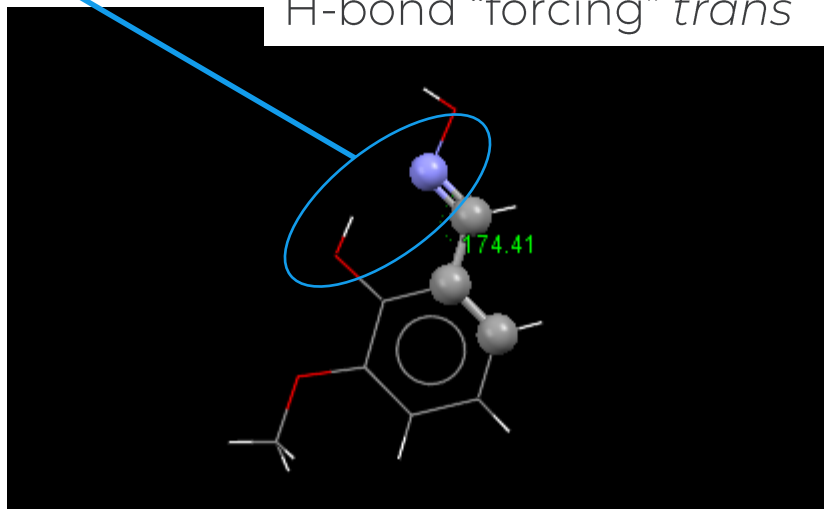


Mogul "requires" non-H substitution at this point adjacent to the target torsion but does not have to be chlorine



These structures don't often have an OH group hence geometry less constrained

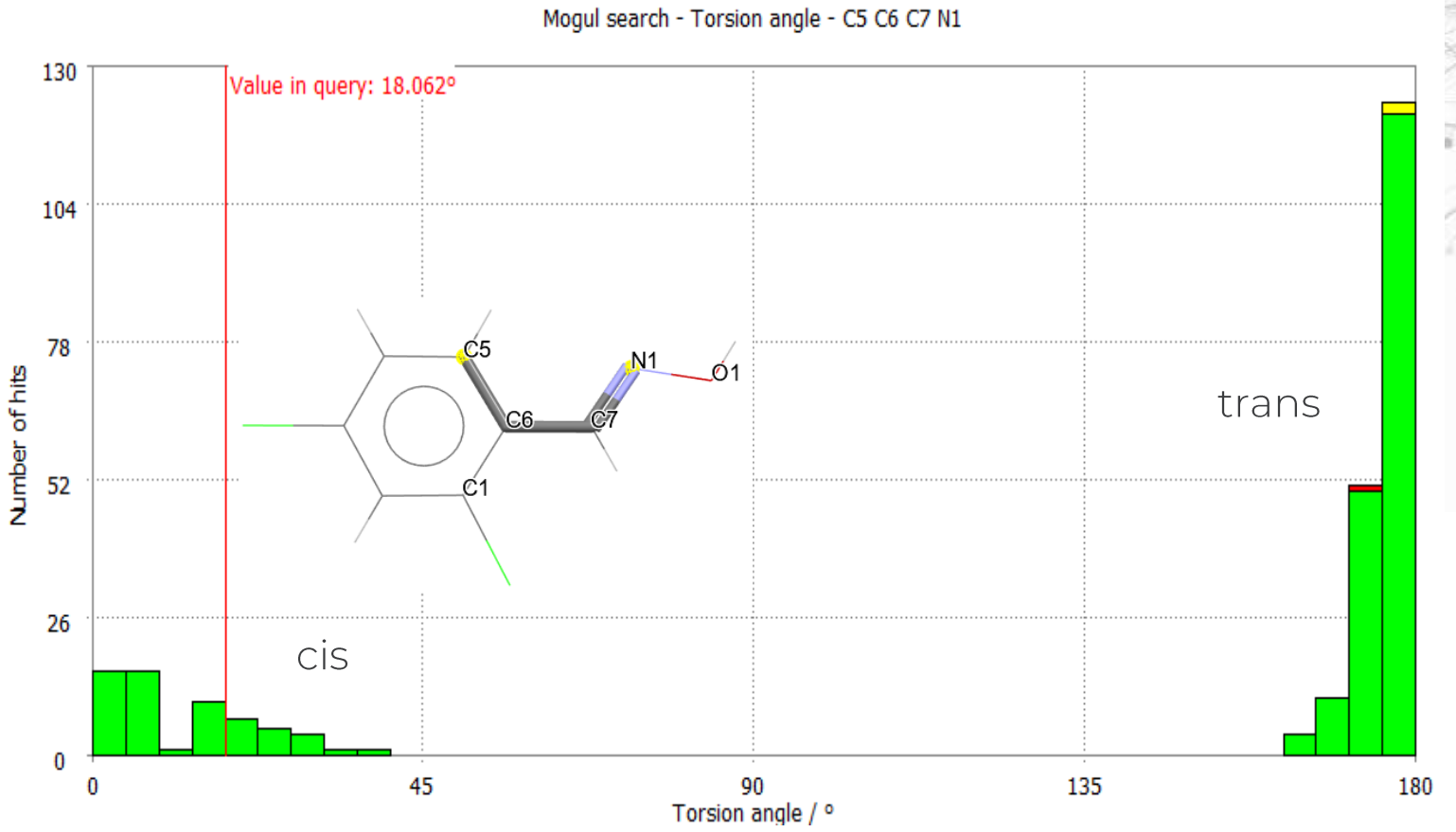
Lots of these structures have OH group and intramolecular H-bond "forcing" *trans*



Example #4 (ADELUR)

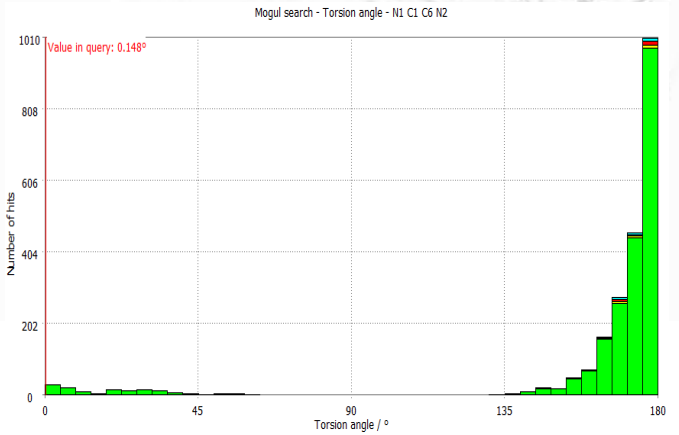
Summary:

- One torsion angle in minor population, reflecting *cis* geometry of N and non-substituted C
- Intramolecular H-bond with OH on substituted C "forces" *trans* geometry (major population)
- Torsion angle in minor population not unusual, especially when context examined



Exploring Mogul results

- If unusual torsion angle(s), investigate how unusual
  - Are there extra things to learn?
  - Are less relevant structures included in the hits?
  - Could the torsion be “usual” given context?
- Consider multiple torsion angles for same rotatable bond
  - Are they all unusual? If not, investigate further
- Check all torsion angles, not just those returned as unusual
  - Observation in minority population → could be a risk (higher energy?)



- Bond lengths and valence angles more relevant for highlighting poor structure refinement, difficulties with accurately locating disordered atom positions, or some strain in the structure (usually in conjunction with unusual torsion angles)

CCDC

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 examples from published scientific articles.

CCDC

Prevalence of unusual torsions

- Occurrence of unusual torsions in the CSD investigated with Python API
  - Organic molecules
  - Best R-factor subset (no redeterminations)
  - 3D coordinates, no disorder, no errors, no polymers
- Around 3% of the total torsions are unusual
- Around 15% of organic crystal structures have at least one "unusual" torsion

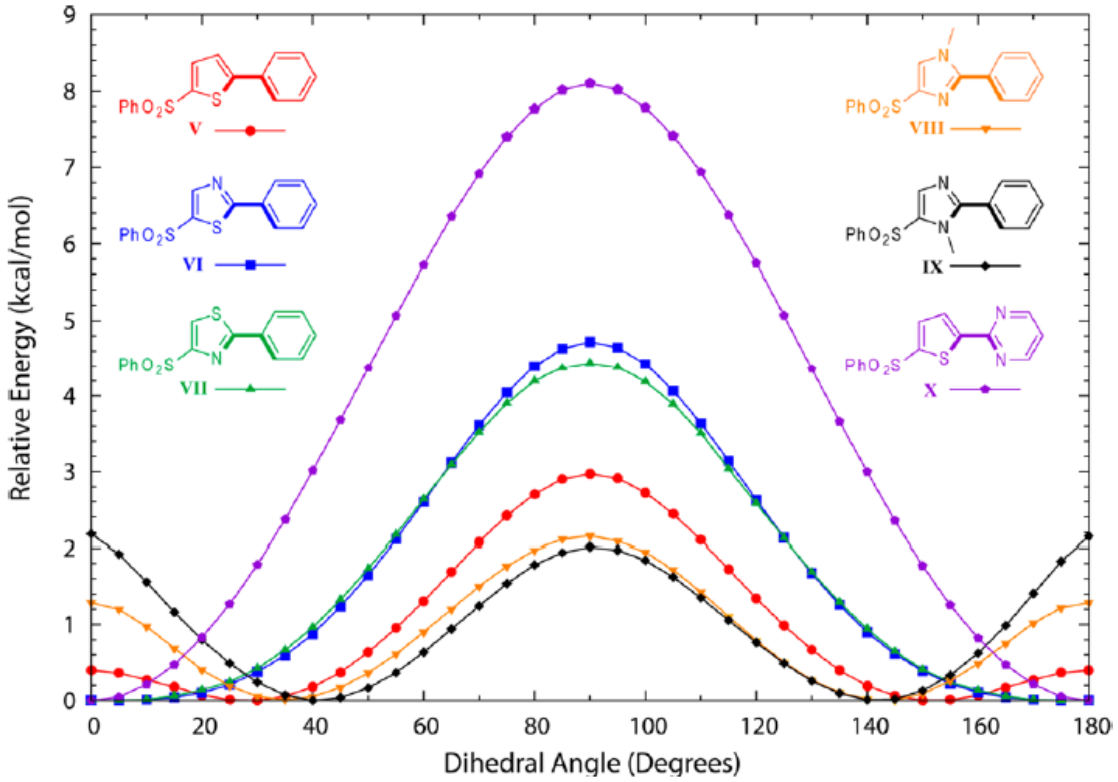
Solid form type	Torsions	Unusual torsions (%)	Crystal structures	Structures with unusual torsions (%)
Single component	948 839	3%	172 524	15%
Solvates	72 038	3%	9837	17%
Hydrates	45 858	3%	7232	15%
Cocrystals	32 928	3%	6191	12%
Salts	76 944	4%	15 559	14%

CrystEngComm, 2020, 22, 7217–7228. DOI: 10.1039/c9ce02001b

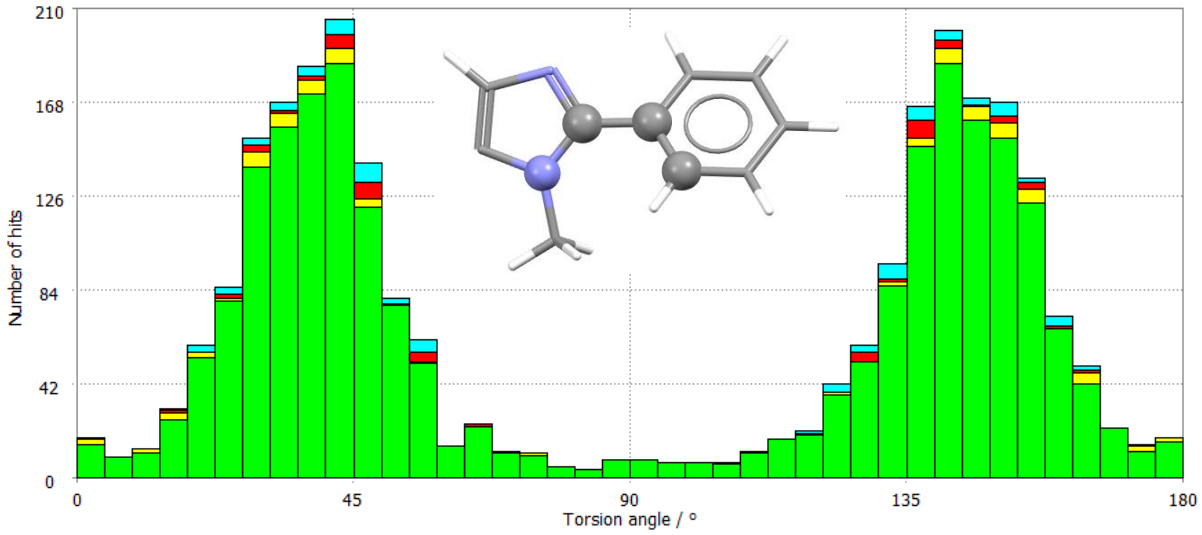
CCDC

Energy distributions

- DFT energy scan across torsion angle range for protein complex inhibitors



- Correlation with Mogul histogram distribution for fragment search (for black/orange molecules)
  - Mogul > Build Query > Draw > Search
  - No organometallics / powder structures

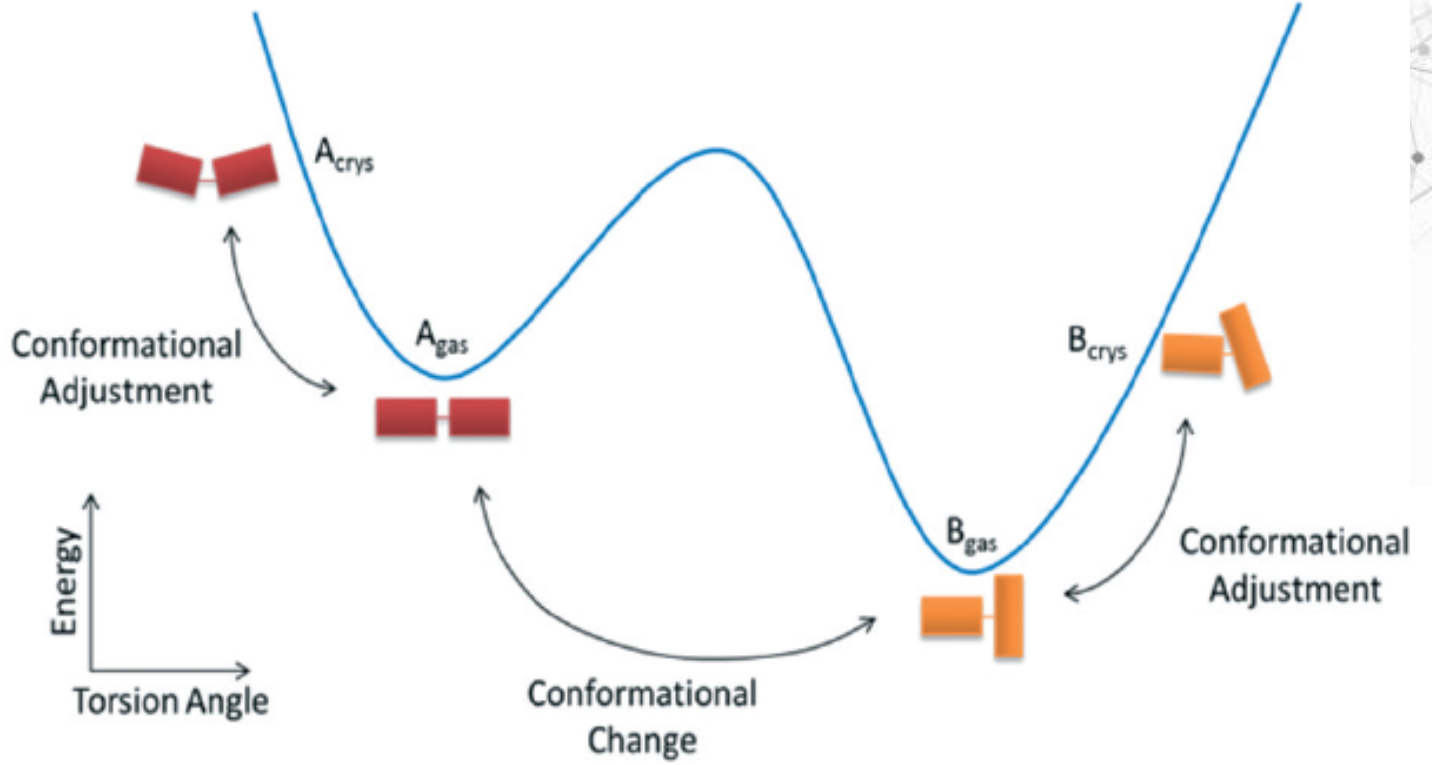


J. Med. Chem., 2015, 58, 4462-4482. DOI:10.1021/jm5018175

CCDC

Conformational adjustment & conformational change

- Relationship between different conformers based on (gas-phase) energy minima for torsion angle with rotation around bond
- Unusual conformational adjustment more common than unusual conformational change

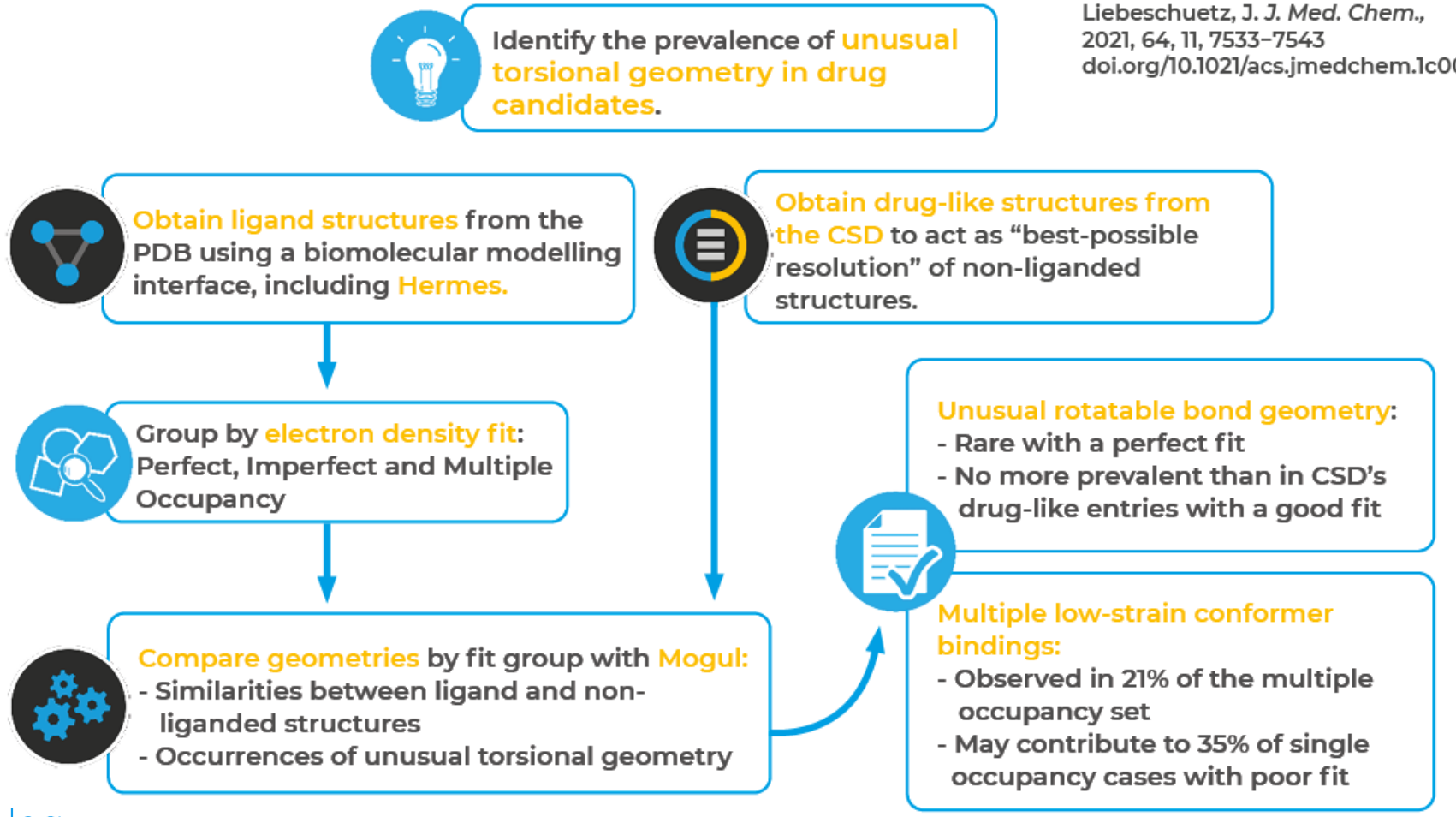


CrystEngComm, 2020, 22, 7217–7228. DOI: 10.1039/c9ce02001b

CCDC

Protein-ligand complex analysis to identify unusual torsional geometry in drug candidates

Liebeschuetz, J. *J. Med. Chem.*, 2021, 64, 11, 7533–7543
doi.org/10.1021/acs.jmedchem.1c00228



Read the blog
https://www.ccdc.cam.ac.uk/Community/blog/Protein-ligand-complex-analysis-to-identify-unusual-torsional-geometry-in-drug-candidates/

CCDC