Insights from CSD crystallographic data applied to drug discovery: How can we do more?

Neysa Nevins

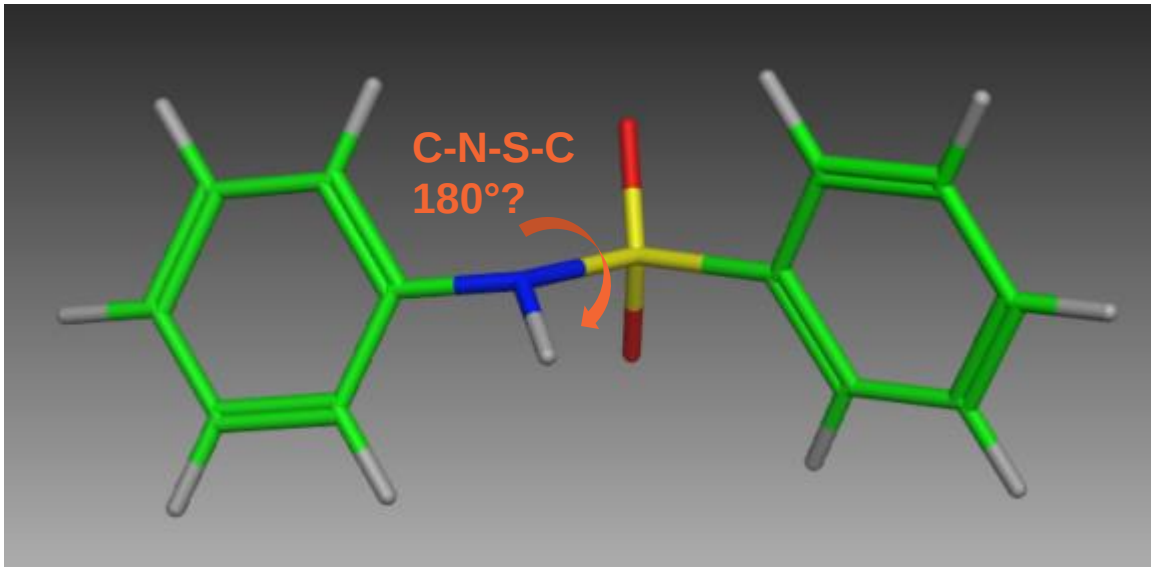
CCDC US RP meeting
Philadelphia, PA
October 17, 2019

My history with CSD software

- 1995-1999: checking torsions during post-doc at Emory University
- 1999-2001: teaching molecular structure at Elizabethtown College
- 2002-present: I became CSD point of contact for GSK US computational chemistry group
  - Not clear who else in GSK used it (*e.g.* folks in Product Development)
  - Our UK comp chem group used more extensively, Colin Edge & Roy Copley
  - I'm still mostly using CSD to check torsions

Rookie mistake: sulfonamide torsion

Circa 2002



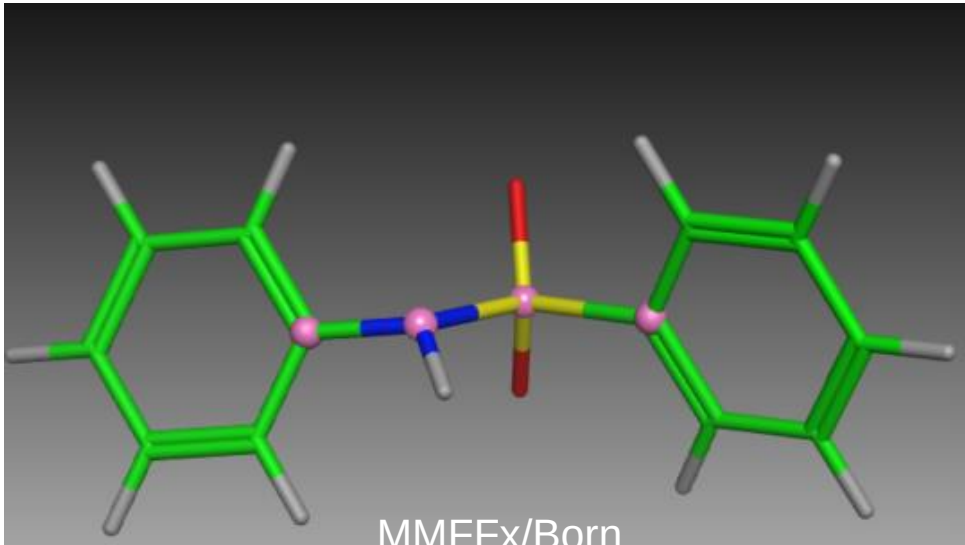
- Accept docking pose with NH eclipsing phenyl?
- Every time torsion question arose needed to fire-up Conquest and search
- Often waited too long and allowed poor torsion to carry through in my proposed docking poses

My history with CSD software

- 1995-1999: checking torsions during post-doc at Emory University
- 1999-2001: teaching about molecular structure at Elizabethtown College
- 2002-present: became CSD point of contact for GSK US computational chemistry group
  - Not clear who else in GSK used it (e.g. folks in Product Development)
  - Our UK comp chem group used more extensively, Colin Edge & Roy Copley
  - I still mostly used CSD to check torsions
- 2013: Met JW Feng at CCG UGM who shared Mogul analysis tool integrated into MOE

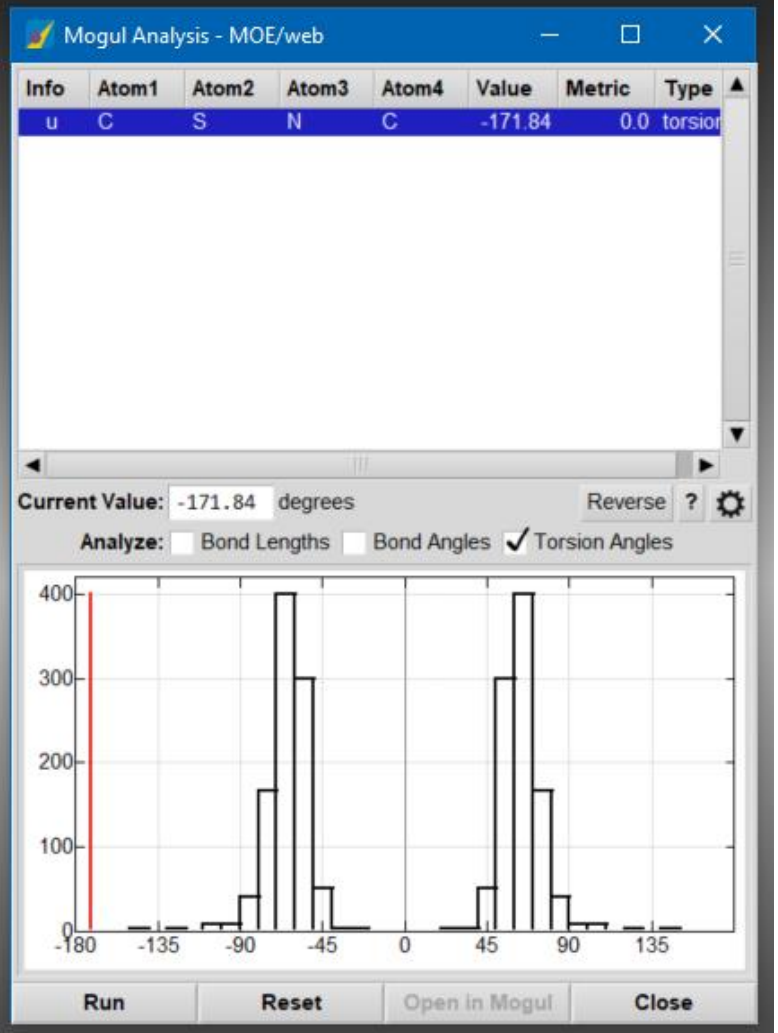
Since ~2015: Mogul Analysis integrated in MOE & Maestro

Soon?: Ligand pose “Mogul Analysis” integrated into docking workflows



MMFFx/Born
ΔE = 3.3 kcal/mol

NH from sulfonamide never eclipses [c,C]



My history with CSD software

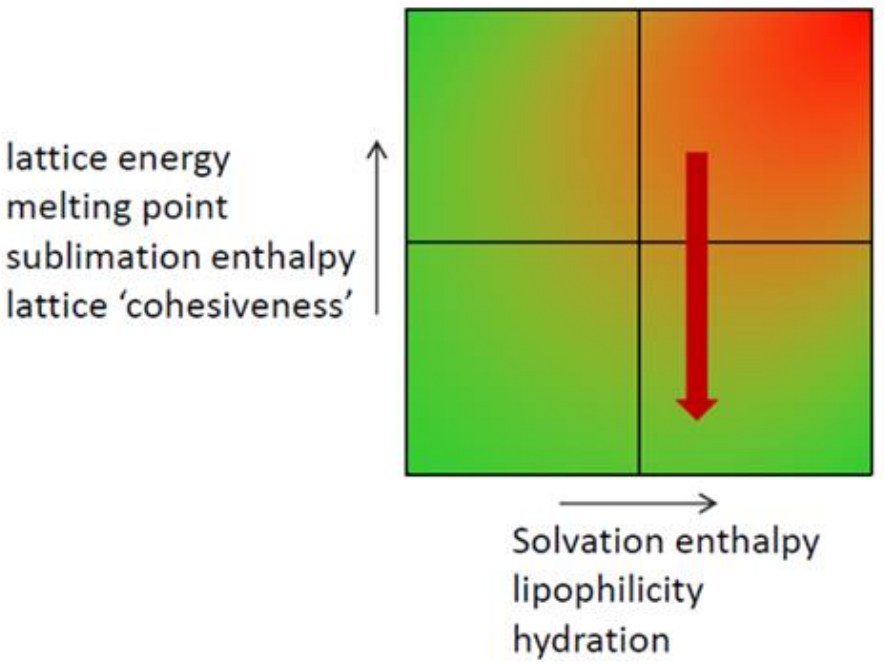
- 1995-1999: checking torsions during post-doc at Emory University
- 1999-2001: teaching about molecular structure at Elizabethtown College
- 2002-present: became CSD point of contact for GSK US computational chemistry group
  - Not clear who else in GSK used it (e.g. folks in Product Development)
  - Our UK comp chem group used more extensively, Colin Edge & Roy Copley
  - I still mostly used CSD to check torsions
- 2013: Met JW Feng at CCG UGM who shared Mogul analysis tool integrated into MOE
- 2014: Attended first US Research Partner meeting at Rutgers
  - Colin Groom presented interesting examples
  - Downloaded his presentations from following Spring 2015 UK RP meeting
- 2015: Attended second US Research Partner meeting at Rutgers

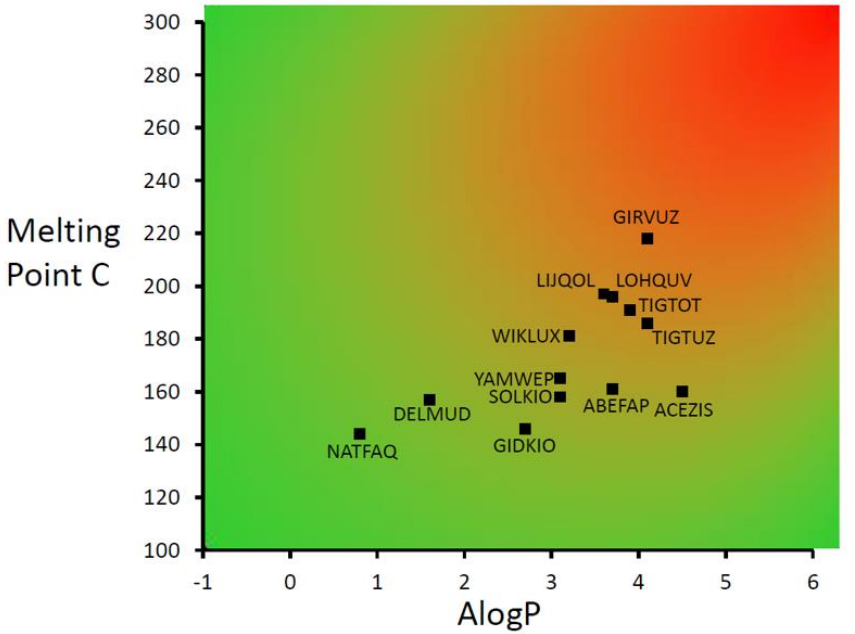
Solubility governed by log P and Melting Point

Groom, UK RP meeting 3/2015

Compounds may be less water soluble if

- They are lipophilic
- They form cohesive, stable crystalline lattices…





$\log S_0 = 0.5 - \log P - 0.01\,(MP - 25)$

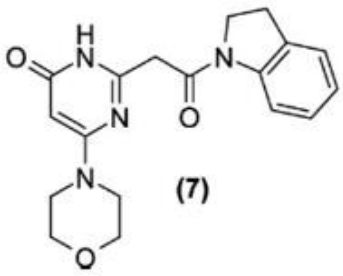
Jain, N.; Yalkowsky, S.H. J. Pharm. Sci. 2001, 90(2) 234-252

Wassvik et al., J. Med. Chem. 2008, 51 (10), pp 3035–3039 DOI: 10.1021/jm701587d

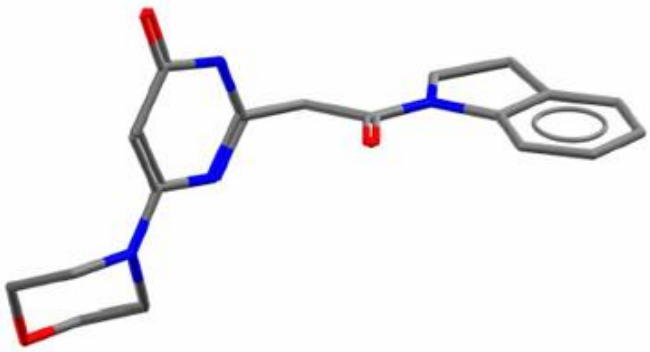
S_0 = intrinsic solubility

Well-placed methyl group improves solubility

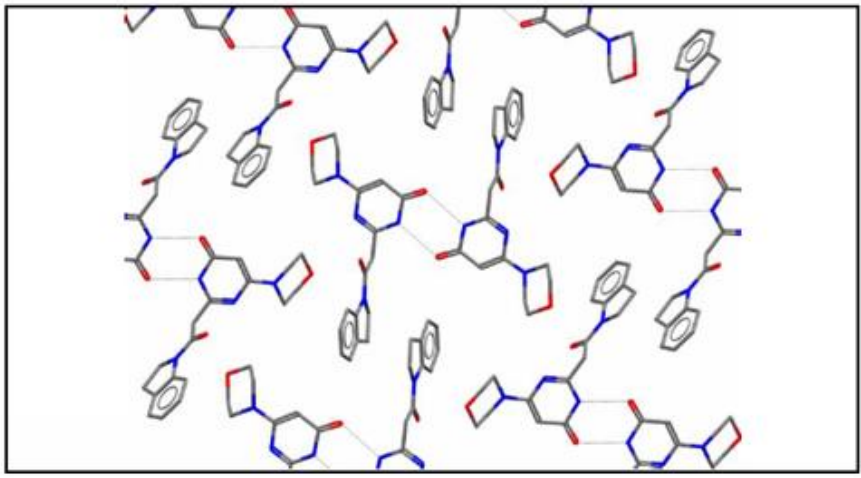
Groom, US RP meeting 9/2014

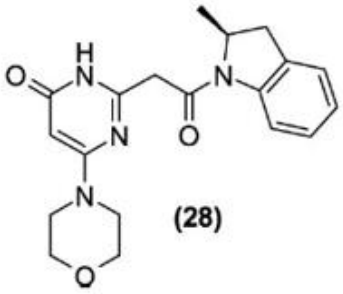


PI3KB IC_{50} 4nM
Solubility pH 7.4 12µM

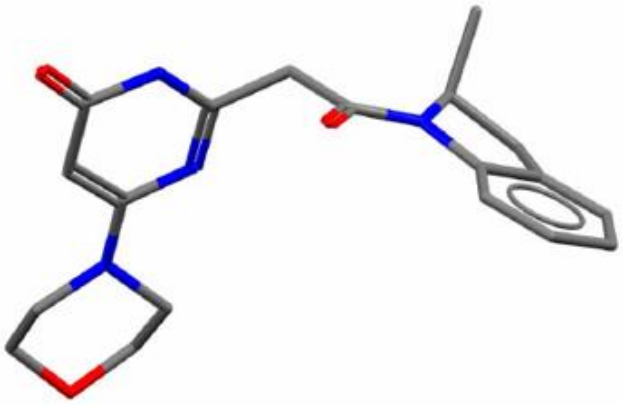


TIZLAR

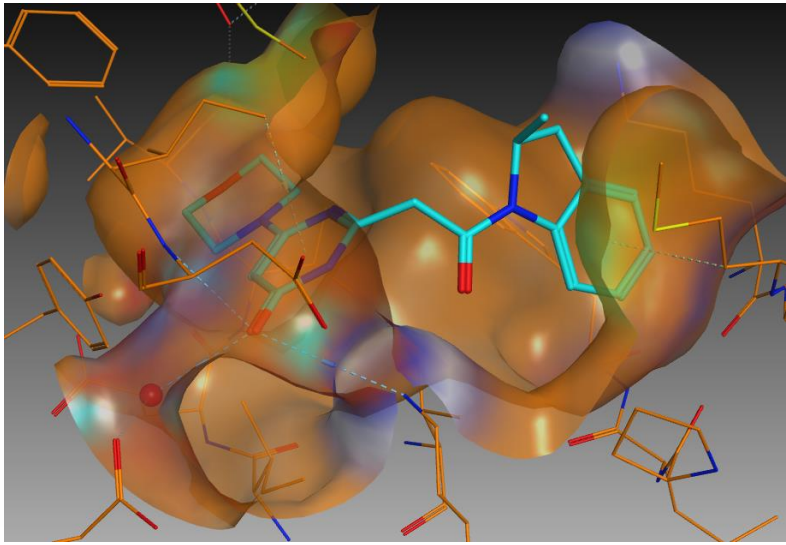




PI3KB IC_{50} 23nM
Solubility pH 7.4 928µM



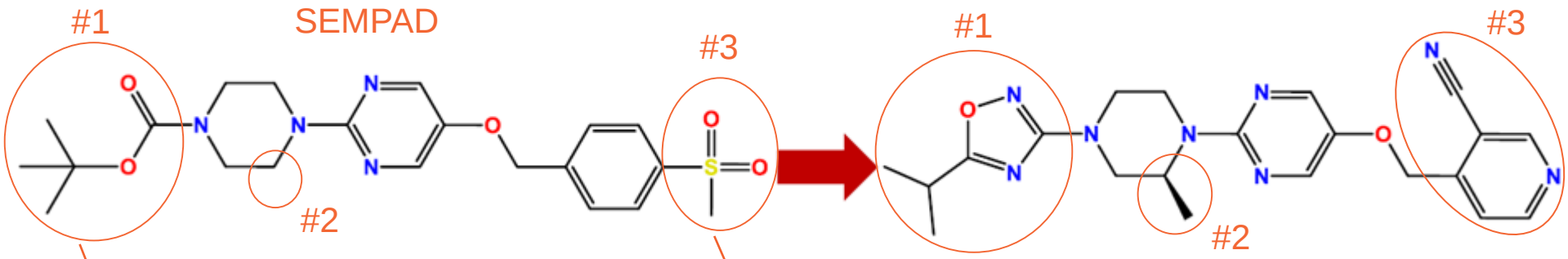
4BFR



Certal *et al. J. Med. Chem.* 2014 57 (3) pp 903-920 DOI: 10.1021/jm401642q (Sanofi, PI3K inhibitors)

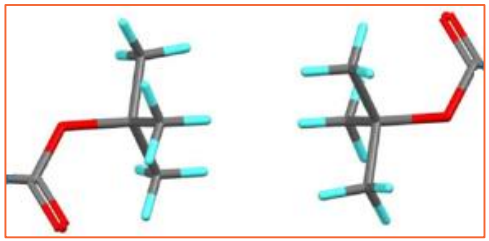
Solubility by design with three replacements

Groom, UK RP meeting 3/2015



GPR119 EC50	65 nM
Solubility	0.03 µM
$LogD_{7.4}$	3.2

GPR119 EC50	6 nM
Solubility	6 µM
$LogD_{7.4}$	3.3



GPR 119 agonists

Scott *et al.*, *J. Med. Chem.*, 2012, *55* (11), pp 5361-5379 DOI: 10.1021/jm300310c
Scott *et al.*, *J. Med. Chem.*, 2014, *57* (21), pp 8984–8998 DOI: 10.1021/jm5011012

Interaction propensity

Groom, US RP meeting 9/2014

Base type	Contact type	R_F	Rank
Br(-)	H(polar)	9.43	1
Cl(-)	H(polar)	7.73	2
I	N(acc)	5.68	3
O	H(polar)	5.13	4
H(polar)	Cl(-)	5.11	5
N(acc)	H(polar)	4.96	6
H(polar)	Br(-)	4.86	7
H(polar)	N(acc)	4.36	8
H(polar)	O	3.83	9
I	O	2.80	10

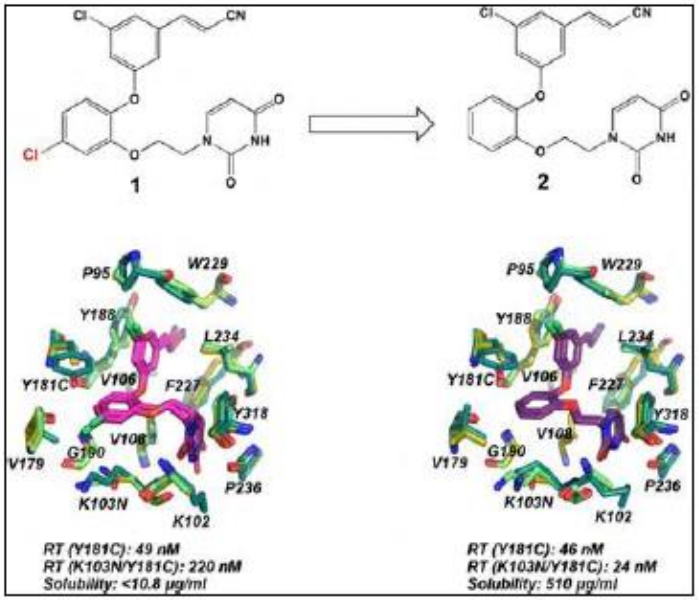
Interaction	R_F
F...H(C)	1.84
F...F	1.11
F...Cl	1.00
F...Br	0.86
F...H(polar)	0.63
F...C(sat)	0.59
F...C(unsat)	0.55
F...I	0.52
F...O	0.41
F...N(nonacc)	0.38
F...N(acc)	0.23

Ratio (R_F) of observed frequency of occurrence to the frequency expected at random

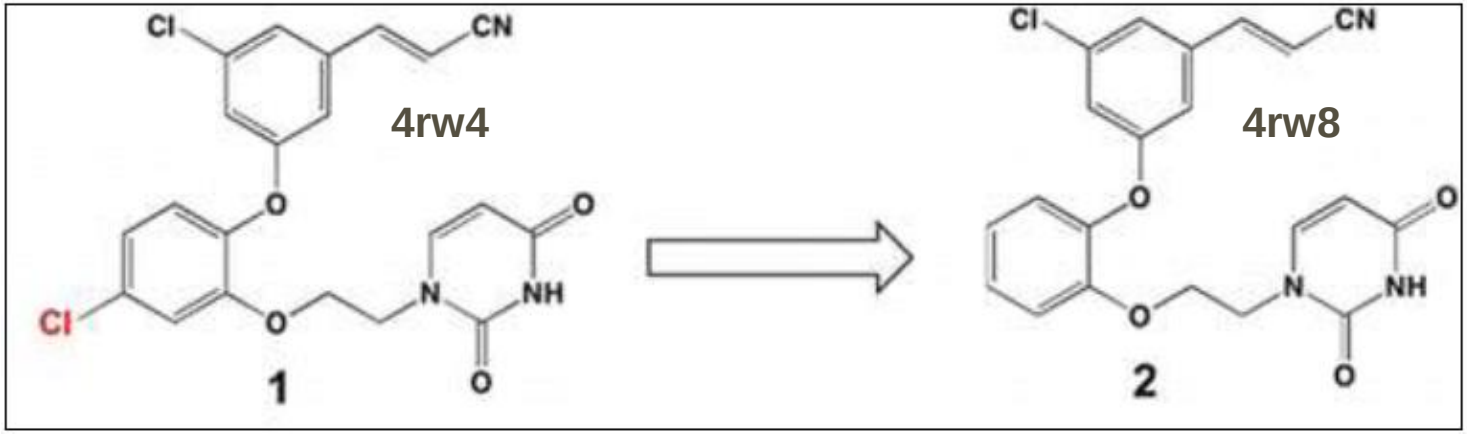
Taylor, CrystEngComm, 16, 6852-6865, 2014 DOI: 10.1039/C4CE00452C

Solubility (Cl -> H)

Groom, US RP meeting 10/2015



Frey et al., J. Med. Chem., 2015, 58 (6), pp 2737–2745
DOI: 10.1021/jm501908a



Solubility < 10.8 µg/ml

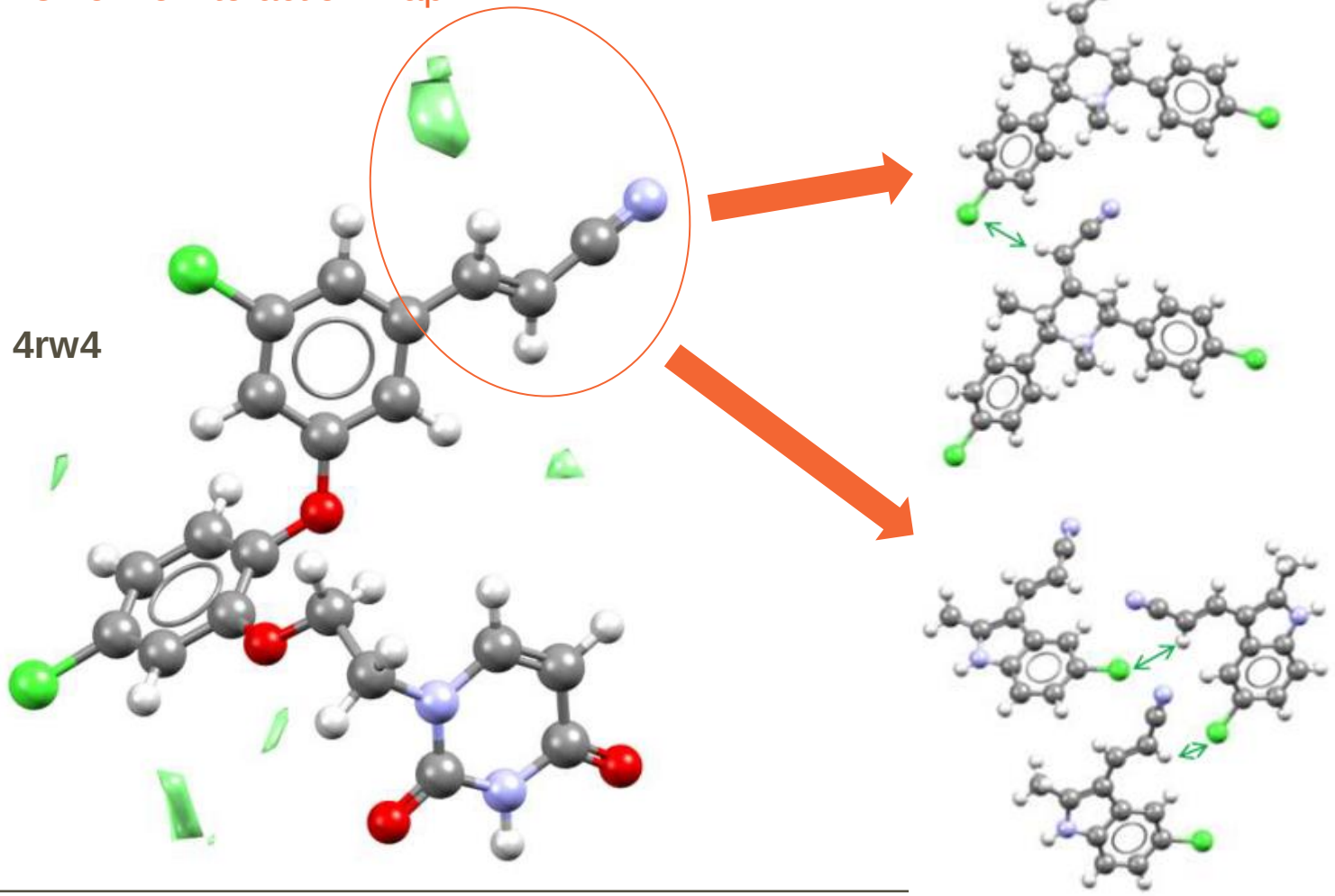
Solubility = 510 µg/ml

Full interaction maps in Mercury

Groom, US RP meeting 10/2015

Olefinic C-H…Cl interaction examples

Chlorine interaction map

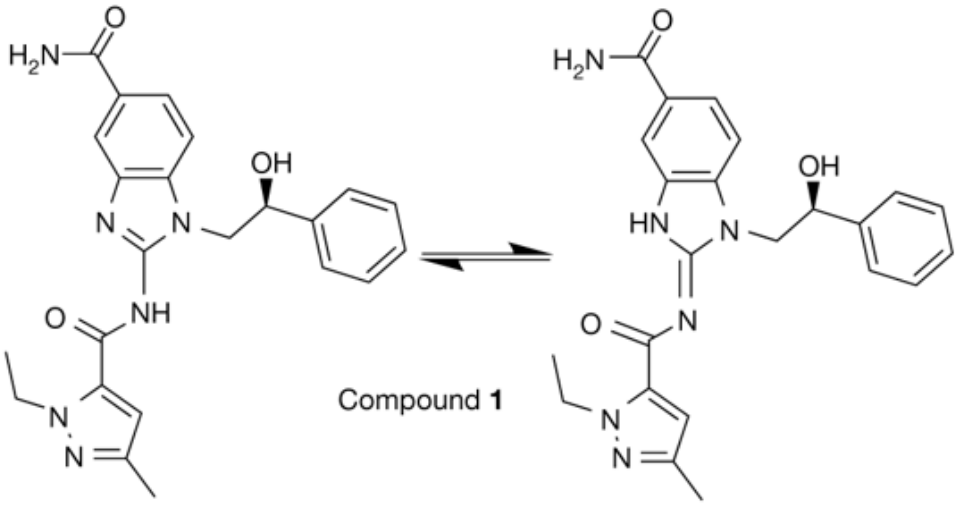


My history with CSD software

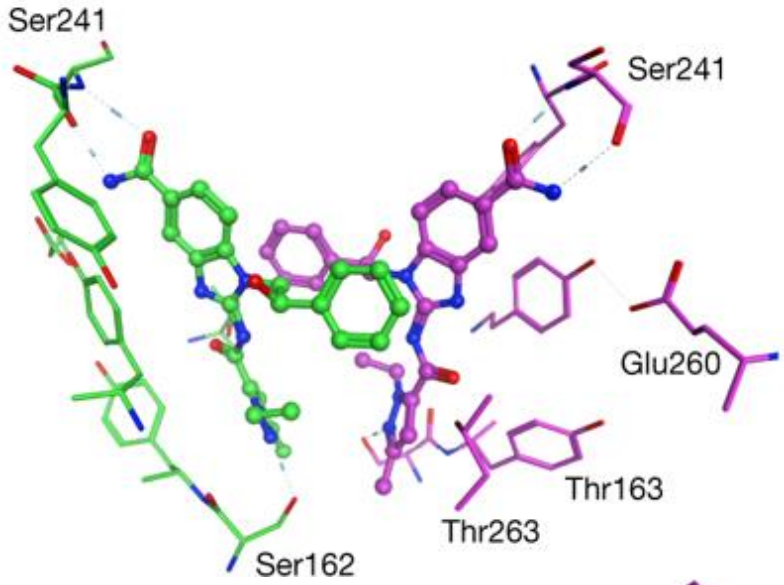
- 1995-1999: checking torsions during post-doc at Emory University
- 1999-2001: teaching about molecular structure at Elizabethtown College
- 2002-present: became CSD point of contact for GSK US computational chemistry group
  - Not clear who else in GSK used it (e.g. folks in Product Development)
  - Our UK comp chem group used more extensively, Colin Edge & Roy Copley
  - I still mostly used CSD to check torsions
- 2013: Met JW Feng at CCG UGM who shared Mogul analysis tool integrated into MOE
- 2014: Attended first US Research Partner meeting at Rutgers
- 2015: Attended second US Research Partner meeting at Rutgers
- 2016-2018: More regular visits from CCDC scientists to GSK to highlight capabilities
- Recent CSD insights for project support:
  - Aminobenzimidazole tautomer
  - Triazole tautomer

Aminobenzimidazole (ABZI) hit binds to STING

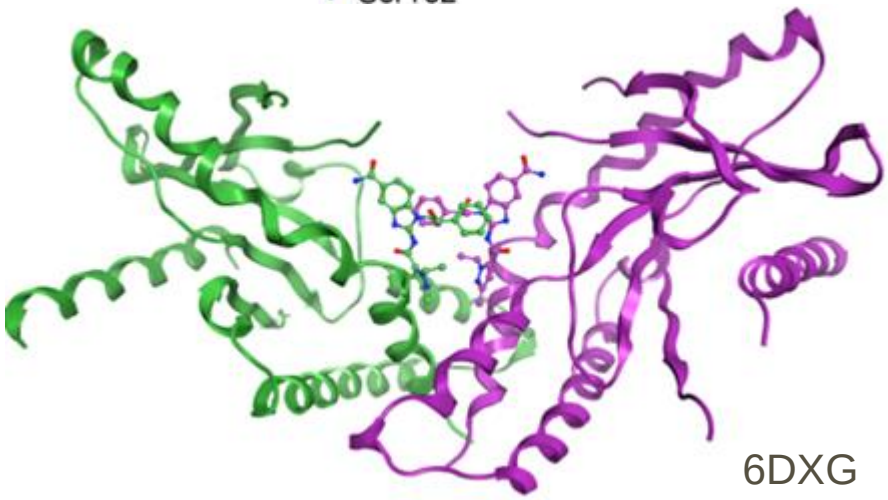
Ramanjulu *et. al.*, *Nature* **564**, 439-443 (2018)



tautomer #1 tautomer #2



How likely is tautomer #2 according to CSD data?

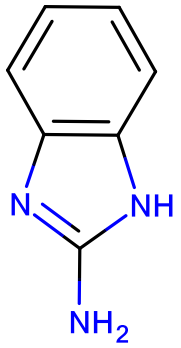


6DXG

Aminobenzimidazole (ABZI) tautomer insights

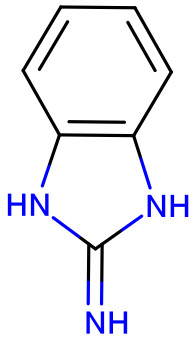
CSD search

– Two possibilities for proton location



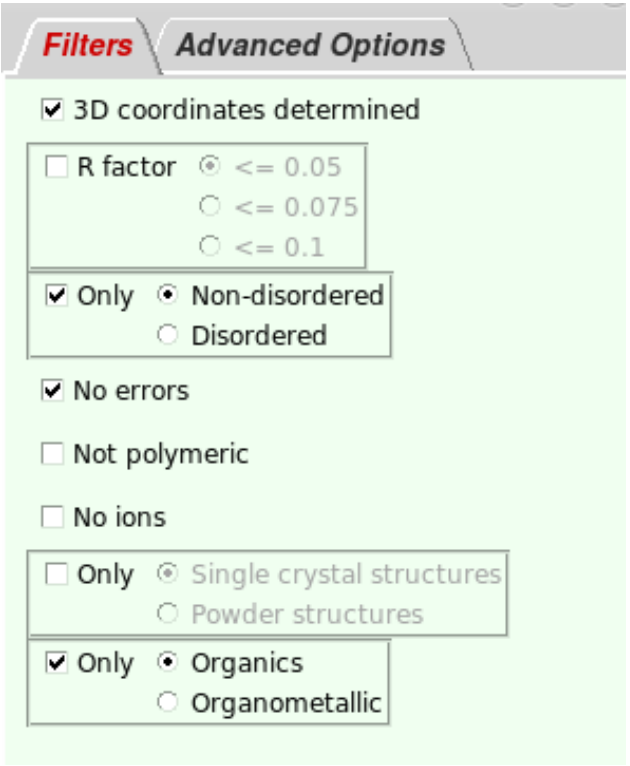
tautomer #1

91 CSD hits

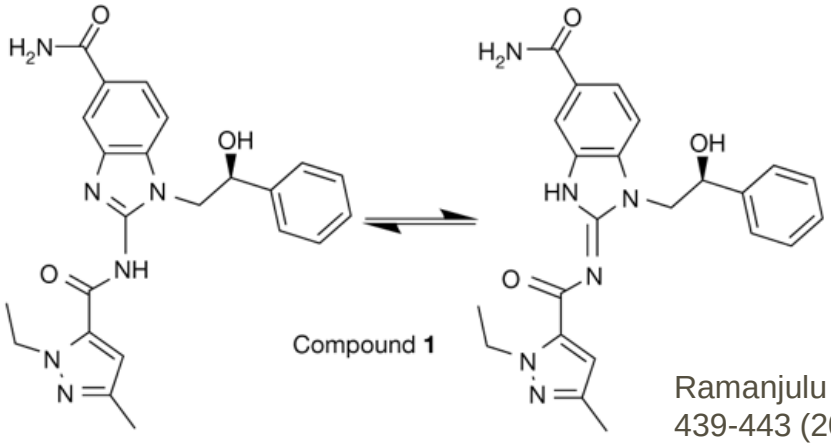


tautomer #2

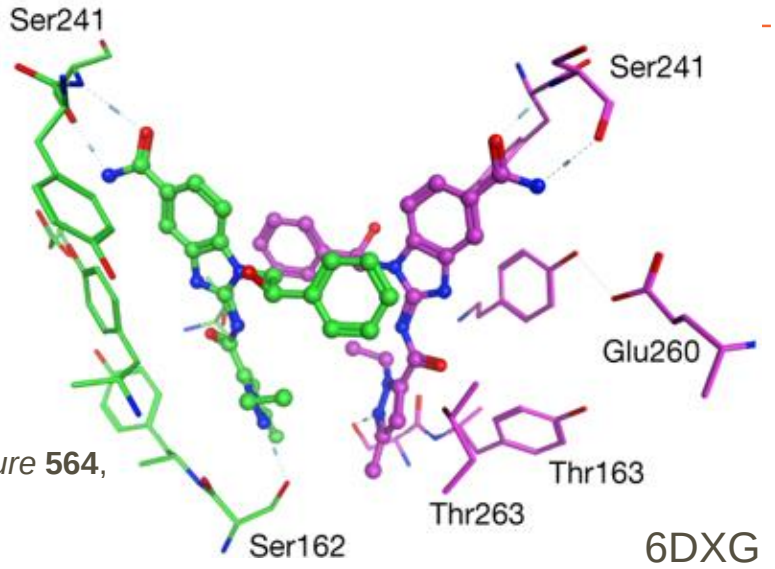
49 CSD hits

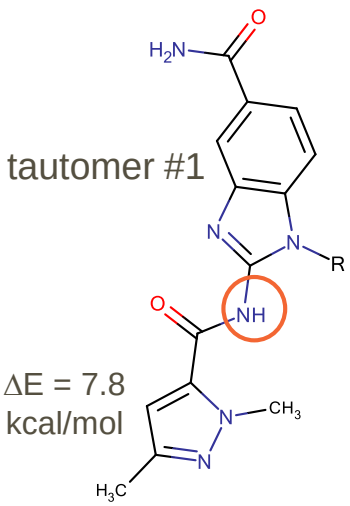


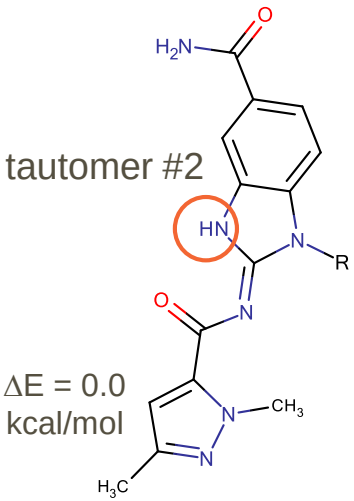
Quantum mechanics calculations favor tautomer #2

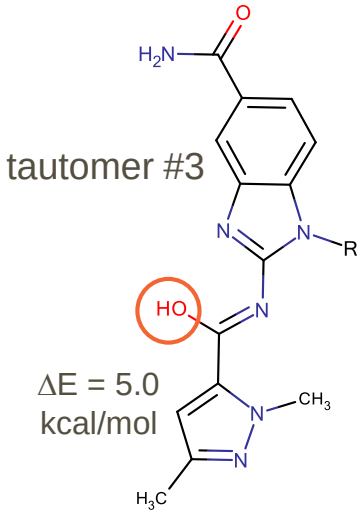


Ramanjulu *et. al.*, *Nature* **564**, 439-443 (2018)







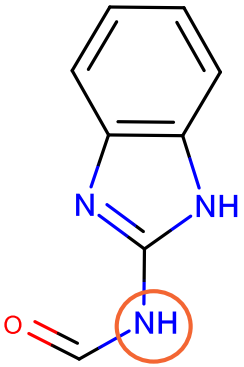


QM tautomer energies (R=CH_3): B3LYP/6-311G**(PCM=water)//B3LYP/6-311G**

Aminobenzimidazole (ABZI) tautomer insights

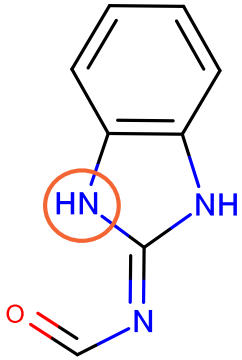
CSD search with amide

– Three possibilities for proton location



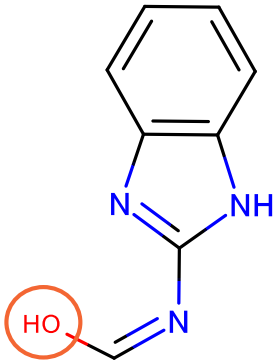
tautomer #1

23 CSD hits
(raw 39 hits)



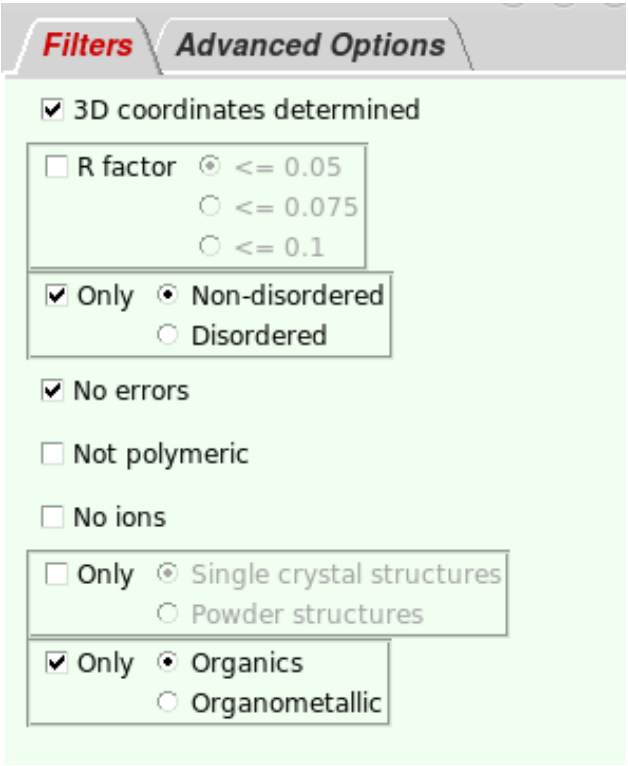
tautomer #2

6 CSD hits
(raw 9 hits)



tautomer #3

0 CSD hits



RIP1 inhibitor with triazole

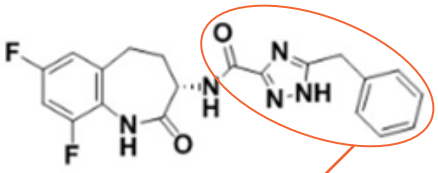
Harris *et. al.*, *ACS Med Chem Lett* **10**, 857-862 (2019)

c1nnc[nH]1
tautomer #1

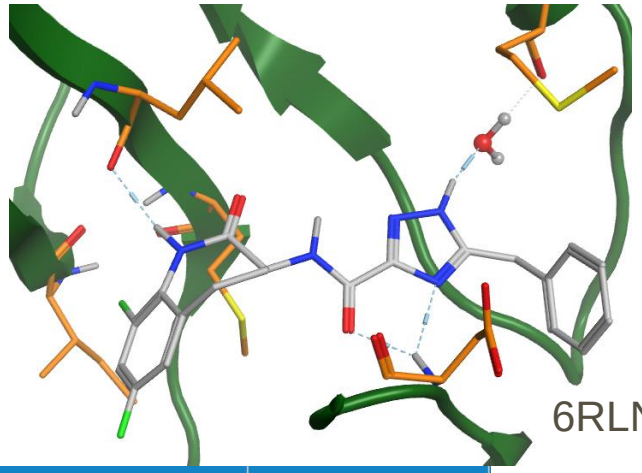
c1n[nH]cn1
tautomer #2

ACD	23K	30K
eMols	20K	42K

RIP1 inhibitor clinical lead for pancreatic cancer

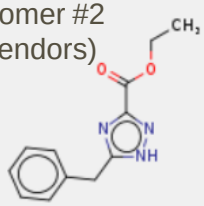
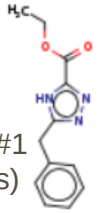


GSK3145095



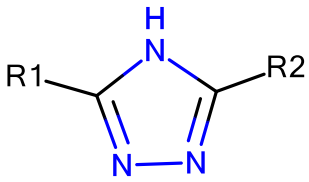
6RLN

eMolecules (building blocks)

Structure	Supplier	Catalog Number	CAS Number
tautomer #2 (2 vendors)  View Compound Info	Aurum Pharmatech	AR36152	648430-85-7
	Enamine BB - EU Stock	EN300-211066	648430-85-7
tautomer #1 (6 vendors)  View Compound Info	AA Blocks LLC	AA00EL0U	648430-85-7
	Aldlab Chemicals	KH-56473	648430-85-7
	Angene Chemical	AG00EL3M	648430-85-7
	AstaTech	C10548	648430-85-7
	ChemScene	CS-B1347	648430-85-7
	Key Organics/BIONET - Building Blocks	CS-13160	648430-85-7

Ligands with triazoles in PDB

Downloaded via PSILO



Tautomer #1
c1nnc[nH]1

33 ligand exemplars with this PDB depiction

1V97	3KQB	5B5P	5OWZ	5XDS
1WBG	3RTP	5CNM	5OX1	5XYY
2GZ8	4CC5	5LRC	5P93	6DGT
3CQU	4CL6	5LSC	5QK1	6F3J
3DB6	4KLB	5N4N	5TX5	6QEJ
3E7O	4NI1	5NR8	5UVC	6RLN
3IU9	5B5O	5O54		

Tautomer #2
c1n[nH]cn1

14 ligand exemplars with this PDB depiction

2J94	4BTW
2OAZ	4IJ9
3AM9	4NI0
3G1L	4Q2F
3PTG	4TZ2
3SMI	5EI4
4B0C	5JCB

Upon visual inspection with receptor/water environment

Depicted	#1 by visual inspection	#2 by visual inspection	Can't assign
Tautomer #1	6	12	15
Tautomer #2	0	8	6

```
                Tautomer #1   Tautomer #2
PDB depicted         33            14
PDB receptor/h2o      6            20
```

Triazole tautomer

Vendor and CSD database searches

– Two triazole proton locations

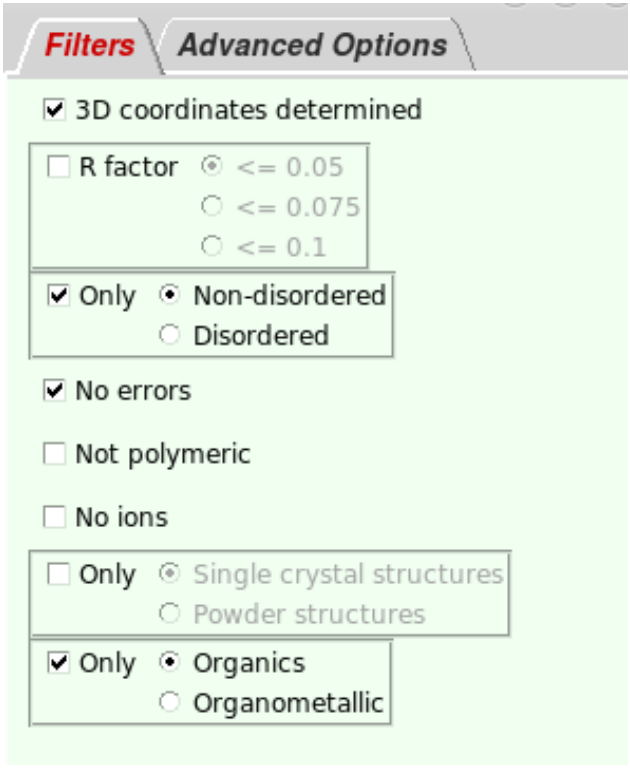
c1nnc[nH]1

tautomer #1

c1n[nH]cn1

tautomer #2

ACD	23K	30K	1:1.3
eMolecules	20K	42K	1:2.1
PDB depicted	33	14	2.4:1
PDB rec/h20	6	20	1:3.3
CSD	11	190	1:17



Triazole not charged (positive or negative)

My history with CSD software

- 1995-1999: checking torsions during post-doc at Emory University
- 1999-2001: teaching about molecular structure at Elizabethtown College
- 2002-present: became CSD point of contact for GSK US computational chemistry group
  - Not clear who else in GSK used it (e.g. folks in Product Development)
  - Our UK comp chem group used more extensively, Colin Edge & Roy Copley
  - I still mostly used CSD to check torsions
- 2013: Met JW Feng at CCG UGM who shared Mogul analysis tool integrated into MOE
- 2014: Attended first US Research Partner meeting at Rutgers
- 2015: Attended second US Research Partner meeting at Rutgers
- 2016-2018: More regular visits from CCDC scientists to GSK to highlight capabilities
- Recent CSD insights for project support:
  - Aminobenzimidazole tautomer
  - Triazole tautomer
- 2018: GSK internal version of WebCSD (experimentalists more likely to access)
  - Fall visit to GSK (PA) by CCDC CEO (Jürgen Harter) and Dir of Business Dev (Jonathan Betts)
- 2019: Informal CCDC US open house at co-working space in Princeton, NJ
  - Discussion with Ian Bruno & Paul Labute to integrate Full Interaction Maps (FIMs) into MOE
- ACS Fall 2019 - One Million Crystal Structures Symposium

Wish list

More tools to prospectively leverage CSD data

- **Integration of CSD data into existing vendor software (***e.g.* Maestro, MOE) for solubility prediction, conformer analysis√, tautomer prediction, interaction hotspots, etc
- **Integration into workflow platforms**, *e.g.* LiveDesign, Orion, Pipeline Pilot
- **User community to discuss priorities** academic and/or industrial, in Slack or MS Teams

Neysa.2.Nevins@gsk.com
Twitter: @neysanev