



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 7, 2020 – 02:59 am BST

PDB ID : 6HU1  
Title : Crystal structure of Schistosoma mansoni HDAC8 complexed with a benzohydroxamate inhibitor 10  
Authors : Shaik, T.B.; Marek, M.; Romier, C.  
Deposited on : 2018-10-05  
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

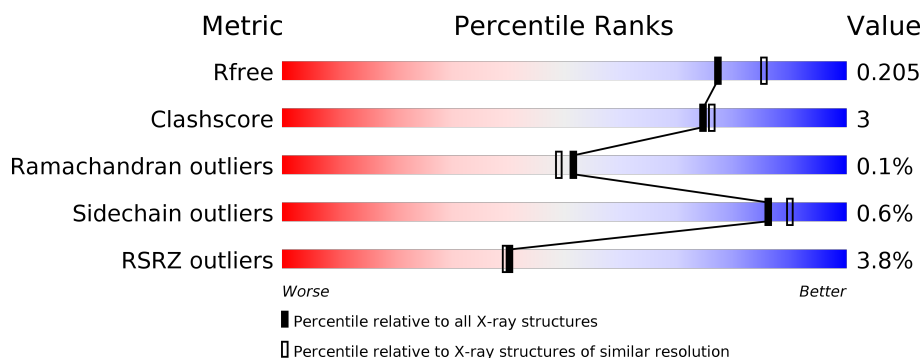
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	447	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>5%</div> <div>11%</div> </div> </div>
1	B	447	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>9%</div> </div> </div>
1	C	447	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>5%</div> <div>8%</div> </div> </div>
1	D	447	<div> <div>4%</div> <div> <div></div> <div>84%</div> <div>5%</div> <div>11%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	DMF	C	505	-	-	X	-
5	DMF	C	508	-	-	X	-
5	DMF	D	507	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 14026 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Histone deacetylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	399	Total	C	N	O	S	0	1	0
			3196	2059	534	588	15			
1	B	409	Total	C	N	O	S	0	1	0
			3268	2107	547	599	15			
1	C	410	Total	C	N	O	S	0	2	0
			3290	2122	550	602	16			
1	D	397	Total	C	N	O	S	0	4	0
			3198	2059	535	588	16			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	HIS	-	expression tag	UNP A5H660
A	441	GLY	-	expression tag	UNP A5H660
A	442	SER	-	expression tag	UNP A5H660
A	443	LEU	-	expression tag	UNP A5H660
A	444	VAL	-	expression tag	UNP A5H660
A	445	PRO	-	expression tag	UNP A5H660
A	446	ARG	-	expression tag	UNP A5H660
B	0	HIS	-	expression tag	UNP A5H660
B	441	GLY	-	expression tag	UNP A5H660
B	442	SER	-	expression tag	UNP A5H660
B	443	LEU	-	expression tag	UNP A5H660
B	444	VAL	-	expression tag	UNP A5H660
B	445	PRO	-	expression tag	UNP A5H660
B	446	ARG	-	expression tag	UNP A5H660
C	0	HIS	-	expression tag	UNP A5H660
C	441	GLY	-	expression tag	UNP A5H660
C	442	SER	-	expression tag	UNP A5H660
C	443	LEU	-	expression tag	UNP A5H660
C	444	VAL	-	expression tag	UNP A5H660
C	445	PRO	-	expression tag	UNP A5H660
C	446	ARG	-	expression tag	UNP A5H660

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	0	HIS	-	expression tag	UNP A5H660
D	441	GLY	-	expression tag	UNP A5H660
D	442	SER	-	expression tag	UNP A5H660
D	443	LEU	-	expression tag	UNP A5H660
D	444	VAL	-	expression tag	UNP A5H660
D	445	PRO	-	expression tag	UNP A5H660
D	446	ARG	-	expression tag	UNP A5H660

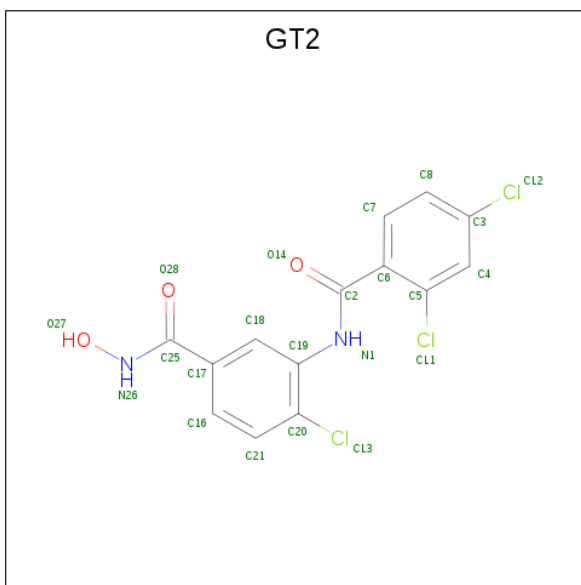
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Zn 1 1	0	0
2	A	1	Total Zn 1 1	0	0
2	D	1	Total Zn 1 1	0	0
2	C	1	Total Zn 1 1	0	0

- Molecule 3 is POTASSIUM ION (three-letter code: K) (formula: K).

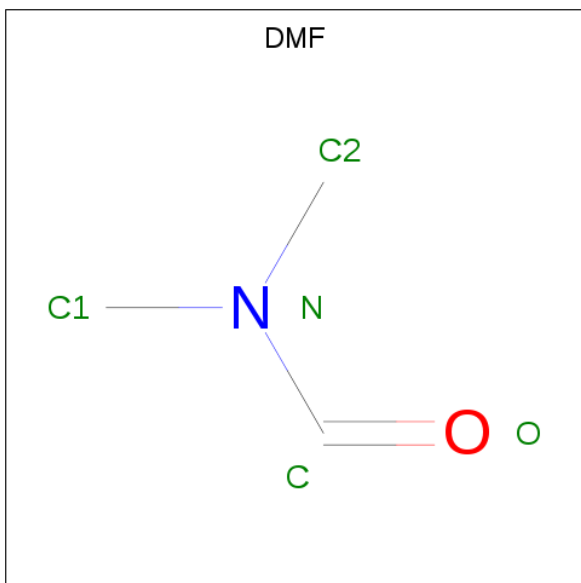
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total K 2 2	0	0
3	A	2	Total K 2 2	0	0
3	D	2	Total K 2 2	0	0
3	C	2	Total K 2 2	0	0

- Molecule 4 is 4-chloranyl-3-[(2,4-dichlorophenyl)carbonylamino]- {N}-oxidanyl-benzamide (three-letter code: GT2) (formula: C<sub>14</sub>H<sub>9</sub>Cl<sub>3</sub>N<sub>2</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 22	C 14	Cl 3	N 2	O 3	0	0
4	B	1	Total 22	C 14	Cl 3	N 2	O 3	0	0
4	C	1	Total 22	C 14	Cl 3	N 2	O 3	0	0
4	D	1	Total 22	C 14	Cl 3	N 2	O 3	0	0

- Molecule 5 is DIMETHYLFORMAMIDE (three-letter code: DMF) (formula:  $\text{C}_3\text{H}_7\text{NO}$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	A	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	B	1	Total	C	N	O	0	0
			5	3	1	1		
5	C	1	Total	C	N	O	0	0
			5	3	1	1		
5	C	1	Total	C	N	O	0	0
			5	3	1	1		
5	C	1	Total	C	N	O	0	0
			5	3	1	1		
5	C	1	Total	C	N	O	0	0
			5	3	1	1		
5	C	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	D	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		
5	D	1	Total	C	N	O	0	0
			5	3	1	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		
6	B	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		

Continued on next page...



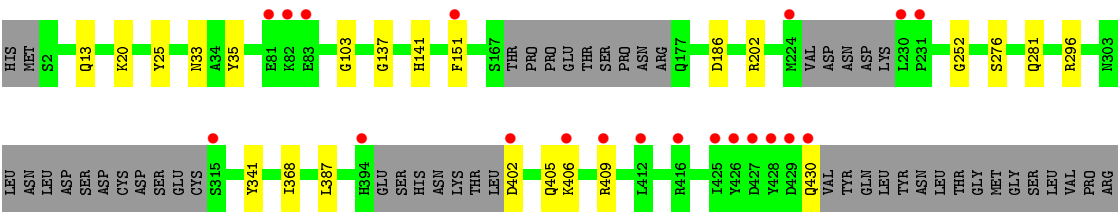
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			6	3	3		
6	D	1	Total	C	O	0	0
			6	3	3		
6	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	195	Total	O	0	0
			195	195		
7	B	211	Total	O	0	0
			211	211		
7	C	187	Total	O	0	0
			187	187		
7	D	194	Total	O	0	0
			194	194		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.22Å 71.19Å 99.03Å 78.27° 75.65° 85.73°	Depositor
Resolution (Å)	47.08 – 2.00 48.41 – 2.00	Depositor EDS
% Data completeness (in resolution range)	95.7 (47.08-2.00) 95.7 (48.41-2.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.44 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.160 , 0.205 0.161 , 0.205	Depositor DCC
$R_{free}$ test set	6007 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.1	Xtriage
Anisotropy	0.209	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 59.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.088 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14026	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.34% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, GT2, K, DMF

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.52	0/3289	0.59	1/4473 (0.0%)
1	B	0.56	0/3362	0.60	0/4573
1	C	0.52	0/3385	0.60	0/4603
1	D	0.52	0/3291	0.57	0/4475
All	All	0.53	0/13327	0.59	1/18124 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	191	ASP	CB-CG-OD1	5.05	122.85	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3196	0	3084	21	0
1	B	3268	0	3164	17	0
1	C	3290	0	3185	22	0
1	D	3198	0	3084	19	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	22	0	0	1	0
4	B	22	0	0	2	0
4	C	22	0	0	2	0
4	D	22	0	0	4	0
5	A	35	0	49	4	0
5	B	40	0	56	3	0
5	C	30	0	42	10	0
5	D	40	0	56	7	0
6	A	6	0	8	0	0
6	B	6	0	8	0	0
6	C	18	0	24	4	0
6	D	12	0	16	0	0
7	A	195	0	0	2	0
7	B	211	0	0	1	0
7	C	187	0	0	1	0
7	D	194	0	0	3	0
All	All	14026	0	12776	82	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 3.

All (82) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:25:TYR:HB2	5:D:507:DMF:H13	1.64	0.80
1:A:413:GLU:HG3	1:A:416:ARG:HH12	1.45	0.79
1:C:12:ARG:HH21	5:C:508:DMF:H23	1.51	0.75
1:A:93:ASP:O	5:A:511:DMF:H23	1.86	0.74
1:D:202:ARG:NH1	7:D:601:HOH:O	2.15	0.69
1:D:35:TYR:CE1	1:D:368:ILE:HG23	2.28	0.68
1:B:145:ARG:HE	5:B:510:DMF:H21	1.57	0.68
1:A:12:ARG:HH11	5:A:507:DMF:H23	1.60	0.66
1:D:341:TYR:OH	4:D:504:GT2:O28	2.11	0.64
1:C:2:SER:N	1:C:131:GLU:OE2	2.31	0.64
1:B:35:TYR:CE1	1:B:368:ILE:HG23	2.33	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:2:SER:N	1:B:131:GLU:OE2	2.33	0.61
1:D:20:LYS:HE2	1:D:151[B]:PHE:CE2	2.37	0.59
1:C:413:GLU:HG3	1:C:416:ARG:NH2	2.19	0.58
1:D:33:ASN:HB2	5:D:506:DMF:H13	1.85	0.57
1:C:12:ARG:HE	5:C:508:DMF:H12	1.70	0.57
1:A:94:SER:HA	5:A:511:DMF:H12	1.87	0.57
1:D:103:GLY:H	5:D:511:DMF:C	2.17	0.57
1:B:200:SER:HB3	7:B:751:HOH:O	2.07	0.55
1:A:375:HIS:O	5:A:509:DMF:HC	2.06	0.55
1:C:413:GLU:HG3	1:C:416:ARG:HH22	1.70	0.54
1:C:20:LYS:NZ	4:C:504:GT2:CL1	2.77	0.53
1:D:20:LYS:HE3	1:D:341:TYR:HE1	1.73	0.53
1:A:239:ARG:NH1	5:C:505:DMF:H11	2.23	0.53
1:D:430:GLN:OE1	1:D:430:GLN:N	2.34	0.53
1:A:341:TYR:OH	4:A:504:GT2:O28	2.17	0.53
1:B:341:TYR:OH	4:B:504:GT2:O28	2.24	0.53
5:C:508:DMF:O	7:C:601:HOH:O	2.19	0.53
1:D:402:ASP:N	1:D:405:GLN:OE1	2.43	0.52
1:A:394:HIS:CG	1:A:395:GLU:H	2.28	0.52
1:C:234:LEU:HB2	6:C:511:GOL:H12	1.92	0.51
1:C:187:LEU:HD21	1:C:294:ILE:HD12	1.92	0.51
1:D:20:LYS:HE3	1:D:341:TYR:CE1	2.46	0.51
1:A:296:ARG:HD3	1:A:391:TYR:OH	2.11	0.50
1:D:20:LYS:NZ	4:D:504:GT2:CL1	2.77	0.50
1:B:46:PRO:HB2	5:B:506:DMF:HC	1.93	0.50
1:A:386:GLU:OE2	7:A:601:HOH:O	2.20	0.50
1:C:179:ARG:NH2	1:C:424:LEU:HD21	2.28	0.49
1:B:187:LEU:HD21	1:B:294:ILE:HD12	1.94	0.49
1:A:423:LYS:NZ	1:C:386:GLU:OE1	2.43	0.49
1:A:35:TYR:CE1	1:A:368:ILE:HG23	2.47	0.49
1:C:341:TYR:OH	4:C:504:GT2:O28	2.20	0.48
1:B:413:GLU:OE2	1:B:416:ARG:NH2	2.46	0.48
1:C:35:TYR:CD1	1:C:368:ILE:HG23	2.48	0.48
1:A:35:TYR:CZ	1:A:373:PRO:HD3	2.49	0.48
1:C:390:ASP:OD1	5:C:505:DMF:H12	2.14	0.47
1:A:239:ARG:HH12	5:C:505:DMF:H11	1.78	0.47
1:B:73:LYS:HA	1:B:76:MET:HE2	1.96	0.47
1:C:10:GLN:HB2	6:C:512:GOL:H12	1.96	0.47
1:B:368:ILE:HG21	1:B:387:LEU:CD2	2.45	0.47
1:A:141:HIS:H	1:A:141:HIS:CD2	2.33	0.46
1:A:413:GLU:HG3	1:A:416:ARG:NH1	2.23	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:13:GLN:OE1	5:D:510:DMF:H21	2.16	0.46
1:A:367:THR:HG23	7:A:673:HOH:O	2.15	0.45
1:C:141:HIS:H	1:C:141:HIS:CD2	2.33	0.45
1:B:222:TRP:HB3	1:B:231:PRO:HB2	1.98	0.45
1:D:186:ASP:HB2	1:D:281:GLN:OE1	2.17	0.45
1:A:111:SER:OG	1:A:153:TYR:HB2	2.17	0.44
1:D:252:GLY:O	1:D:296:ARG:HD2	2.17	0.44
1:B:35:TYR:CD1	1:B:368:ILE:HG23	2.52	0.44
1:C:24:ARG:HH21	6:C:513:GOL:H31	1.82	0.44
1:C:20:LYS:HE2	1:C:151[B]:PHE:CE2	2.54	0.43
1:B:7:TYR:CD2	5:B:507:DMF:H13	2.54	0.43
1:B:368:ILE:HG21	1:B:387:LEU:HD22	2.01	0.43
1:C:24:ARG:HE	6:C:513:GOL:C3	2.32	0.43
1:A:35:TYR:CD1	1:A:368:ILE:HG23	2.54	0.42
5:D:507:DMF:H22	7:D:679:HOH:O	2.19	0.42
1:C:12:ARG:NH2	5:C:508:DMF:H23	2.28	0.42
5:D:507:DMF:H12	7:D:684:HOH:O	2.18	0.42
1:B:253:ILE:HG22	1:B:295:PHE:CD1	2.55	0.42
4:D:504:GT2:CL1	4:D:504:GT2:O14	2.74	0.41
1:A:239:ARG:NH2	5:C:505:DMF:H11	2.35	0.41
1:D:137:GLY:HA2	5:D:507:DMF:H11	2.02	0.41
1:C:35:TYR:CE1	1:C:368:ILE:HG23	2.56	0.41
1:D:368:ILE:HG21	1:D:387:LEU:HD22	2.01	0.41
1:D:151[B]:PHE:HE2	4:D:504:GT2:C19	2.34	0.41
1:B:341:TYR:CD1	4:B:504:GT2:CL1	3.11	0.41
1:C:384:ASP:CB	5:C:505:DMF:H21	2.50	0.41
1:A:239:ARG:CZ	5:C:505:DMF:H11	2.52	0.40
1:B:139:GLY:HA2	1:B:157:ILE:HD11	2.03	0.40
1:C:289:THR:HA	1:C:293:ARG:NH1	2.36	0.40
1:D:406:LYS:HD3	1:D:409:ARG:HH21	1.87	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.



The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	390/447 (87%)	385 (99%)	5 (1%)	0	100	100
1	B	398/447 (89%)	390 (98%)	8 (2%)	0	100	100
1	C	402/447 (90%)	394 (98%)	7 (2%)	1 (0%)	47	44
1	D	391/447 (88%)	386 (99%)	5 (1%)	0	100	100
All	All	1581/1788 (88%)	1555 (98%)	25 (2%)	1 (0%)	51	49

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	99	TYR

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	347/392 (88%)	345 (99%)	2 (1%)	86	90
1	B	355/392 (91%)	353 (99%)	2 (1%)	86	90
1	C	357/392 (91%)	355 (99%)	2 (1%)	86	90
1	D	348/392 (89%)	346 (99%)	2 (1%)	86	90
All	All	1407/1568 (90%)	1399 (99%)	8 (1%)	86	90

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	141	HIS
1	A	177	GLN
1	B	141	HIS
1	B	230	LEU
1	C	80	GLU
1	C	141	HIS
1	D	141	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	276	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 52 ligands modelled in this entry, 12 are monoatomic - leaving 40 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	DMF	C	505	-	4,4,4	0.27	0	4,4,4	0.16	0
5	DMF	D	507	-	4,4,4	0.23	0	4,4,4	0.61	0
4	GT2	D	504	2	23,23,23	0.28	0	32,32,32	0.38	0
5	DMF	A	509	-	4,4,4	0.41	0	4,4,4	0.53	0
5	DMF	B	507	-	4,4,4	0.38	0	4,4,4	0.50	0
5	DMF	D	506	-	4,4,4	0.38	0	4,4,4	0.29	0
5	DMF	D	508	-	4,4,4	0.36	0	4,4,4	0.51	0
5	DMF	D	512	-	4,4,4	0.33	0	4,4,4	0.54	0
5	DMF	D	509	-	4,4,4	0.35	0	4,4,4	0.41	0
5	DMF	B	508	-	4,4,4	0.37	0	4,4,4	0.51	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	DMF	A	505	-	4,4,4	0.41	0	4,4,4	0.33	0
6	GOL	C	512	-	5,5,5	0.31	0	5,5,5	0.49	0
5	DMF	A	511	-	4,4,4	0.31	0	4,4,4	0.69	0
5	DMF	B	506	-	4,4,4	0.28	0	4,4,4	0.37	0
5	DMF	B	505	-	4,4,4	0.37	0	4,4,4	0.31	0
5	DMF	C	510	-	4,4,4	0.39	0	4,4,4	0.64	0
5	DMF	A	507	-	4,4,4	0.33	0	4,4,4	0.37	0
5	DMF	D	510	-	4,4,4	0.31	0	4,4,4	0.21	0
6	GOL	C	513	-	5,5,5	0.51	0	5,5,5	0.73	0
5	DMF	C	509	-	4,4,4	0.38	0	4,4,4	0.88	0
6	GOL	D	514	-	5,5,5	0.39	0	5,5,5	0.47	0
6	GOL	B	513	-	5,5,5	0.38	0	5,5,5	0.36	0
6	GOL	C	511	-	5,5,5	0.47	0	5,5,5	0.65	0
4	GT2	C	504	2	23,23,23	0.27	0	32,32,32	0.33	0
5	DMF	B	510	-	4,4,4	0.32	0	4,4,4	0.05	0
5	DMF	D	511	-	4,4,4	0.30	0	4,4,4	0.33	0
5	DMF	D	505	-	4,4,4	0.41	0	4,4,4	0.64	0
5	DMF	A	508	-	4,4,4	0.38	0	4,4,4	0.65	0
5	DMF	C	508	-	4,4,4	0.34	0	4,4,4	0.49	0
5	DMF	A	506	-	4,4,4	0.35	0	4,4,4	0.41	0
5	DMF	B	511	-	4,4,4	0.33	0	4,4,4	0.69	0
6	GOL	A	512	-	5,5,5	0.36	0	5,5,5	0.17	0
6	GOL	D	513	-	5,5,5	0.48	0	5,5,5	0.41	0
5	DMF	B	509	-	4,4,4	0.33	0	4,4,4	0.76	0
4	GT2	B	504	2	23,23,23	0.26	0	32,32,32	0.28	0
5	DMF	B	512	-	4,4,4	0.42	0	4,4,4	0.38	0
5	DMF	C	507	-	4,4,4	0.37	0	4,4,4	0.35	0
4	GT2	A	504	2	23,23,23	0.25	0	32,32,32	0.32	0
5	DMF	A	510	-	4,4,4	0.35	0	4,4,4	0.41	0
5	DMF	C	506	-	4,4,4	0.45	0	4,4,4	0.19	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DMF	C	505	-	-	2/2/2/2	-
5	DMF	D	507	-	-	2/2/2/2	-
4	GT2	D	504	2	-	2/14/14/14	0/2/2/2
5	DMF	A	509	-	-	2/2/2/2	-
5	DMF	B	507	-	-	0/2/2/2	-

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	DMF	D	506	-	-	2/2/2/2	-
5	DMF	D	508	-	-	2/2/2/2	-
5	DMF	D	512	-	-	2/2/2/2	-
5	DMF	D	509	-	-	2/2/2/2	-
5	DMF	B	508	-	-	0/2/2/2	-
5	DMF	A	505	-	-	0/2/2/2	-
6	GOL	C	512	-	-	2/4/4/4	-
5	DMF	A	511	-	-	0/2/2/2	-
5	DMF	B	506	-	-	2/2/2/2	-
5	DMF	B	505	-	-	2/2/2/2	-
5	DMF	C	510	-	-	0/2/2/2	-
5	DMF	A	507	-	-	2/2/2/2	-
5	DMF	D	510	-	-	2/2/2/2	-
6	GOL	C	513	-	-	4/4/4/4	-
5	DMF	C	509	-	-	2/2/2/2	-
6	GOL	D	514	-	-	2/4/4/4	-
6	GOL	B	513	-	-	0/4/4/4	-
6	GOL	C	511	-	-	4/4/4/4	-
4	GT2	C	504	2	-	4/14/14/14	0/2/2/2
5	DMF	B	510	-	-	2/2/2/2	-
5	DMF	D	511	-	-	2/2/2/2	-
5	DMF	D	505	-	-	0/2/2/2	-
5	DMF	A	508	-	-	2/2/2/2	-
5	DMF	C	508	-	-	0/2/2/2	-
5	DMF	A	506	-	-	0/2/2/2	-
5	DMF	B	511	-	-	2/2/2/2	-
6	GOL	A	512	-	-	0/4/4/4	-
6	GOL	D	513	-	-	2/4/4/4	-
5	DMF	B	509	-	-	0/2/2/2	-
4	GT2	B	504	2	-	2/14/14/14	0/2/2/2
5	DMF	B	512	-	-	0/2/2/2	-
5	DMF	C	507	-	-	0/2/2/2	-
4	GT2	A	504	2	-	4/14/14/14	0/2/2/2
5	DMF	A	510	-	-	0/2/2/2	-
5	DMF	C	506	-	-	0/2/2/2	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (58) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	C	513	GOL	C1-C2-C3-O3
4	D	504	GT2	N1-C2-C6-C5
4	D	504	GT2	O14-C2-C6-C5
6	C	511	GOL	O1-C1-C2-C3
6	C	511	GOL	C1-C2-C3-O3
4	C	504	GT2	C20-C19-N1-C2
5	D	506	DMF	O-C-N-C2
5	B	506	DMF	O-C-N-C1
5	D	506	DMF	O-C-N-C1
5	D	510	DMF	O-C-N-C1
5	B	511	DMF	O-C-N-C2
5	D	510	DMF	O-C-N-C2
5	C	505	DMF	O-C-N-C2
5	A	507	DMF	O-C-N-C1
5	B	506	DMF	O-C-N-C2
5	B	511	DMF	O-C-N-C1
5	C	505	DMF	O-C-N-C1
5	A	507	DMF	O-C-N-C2
6	D	514	GOL	O1-C1-C2-C3
6	C	512	GOL	C1-C2-C3-O3
5	D	509	DMF	O-C-N-C2
6	C	513	GOL	O2-C2-C3-O3
6	D	514	GOL	O1-C1-C2-O2
6	C	512	GOL	O2-C2-C3-O3
6	C	511	GOL	O1-C1-C2-O2
6	C	511	GOL	O2-C2-C3-O3
5	D	512	DMF	O-C-N-C1
5	D	508	DMF	O-C-N-C2
5	D	509	DMF	O-C-N-C1
4	B	504	GT2	C20-C19-N1-C2
5	A	508	DMF	O-C-N-C2
4	A	504	GT2	O14-C2-C6-C5
6	C	513	GOL	O1-C1-C2-O2
4	C	504	GT2	C18-C19-N1-C2
5	D	512	DMF	O-C-N-C2
5	D	508	DMF	O-C-N-C1
4	B	504	GT2	C18-C19-N1-C2
5	A	508	DMF	O-C-N-C1
5	A	509	DMF	O-C-N-C1
6	C	513	GOL	O1-C1-C2-C3

*Continued on next page...*

*Continued from previous page...*

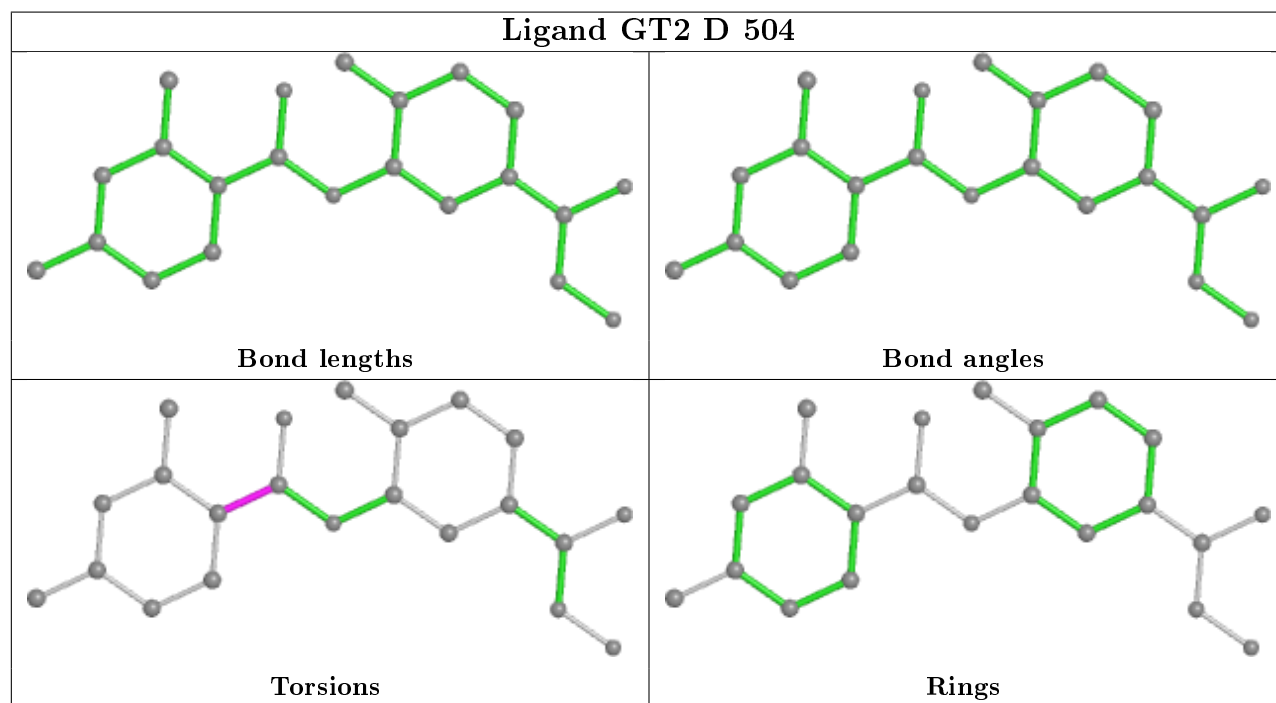
Mol	Chain	Res	Type	Atoms
5	C	509	DMF	O-C-N-C1
5	B	510	DMF	O-C-N-C1
5	D	511	DMF	O-C-N-C1
4	A	504	GT2	N1-C2-C6-C5
6	D	513	GOL	O2-C2-C3-O3
5	A	509	DMF	O-C-N-C2
4	A	504	GT2	C18-C19-N1-C2
4	A	504	GT2	C20-C19-N1-C2
5	B	510	DMF	O-C-N-C2
5	D	511	DMF	O-C-N-C2
5	C	509	DMF	O-C-N-C2
5	B	505	DMF	O-C-N-C2
4	C	504	GT2	N1-C2-C6-C5
6	D	513	GOL	C1-C2-C3-O3
5	B	505	DMF	O-C-N-C1
4	C	504	GT2	O14-C2-C6-C5
5	D	507	DMF	O-C-N-C1
5	D	507	DMF	O-C-N-C2

There are no ring outliers.

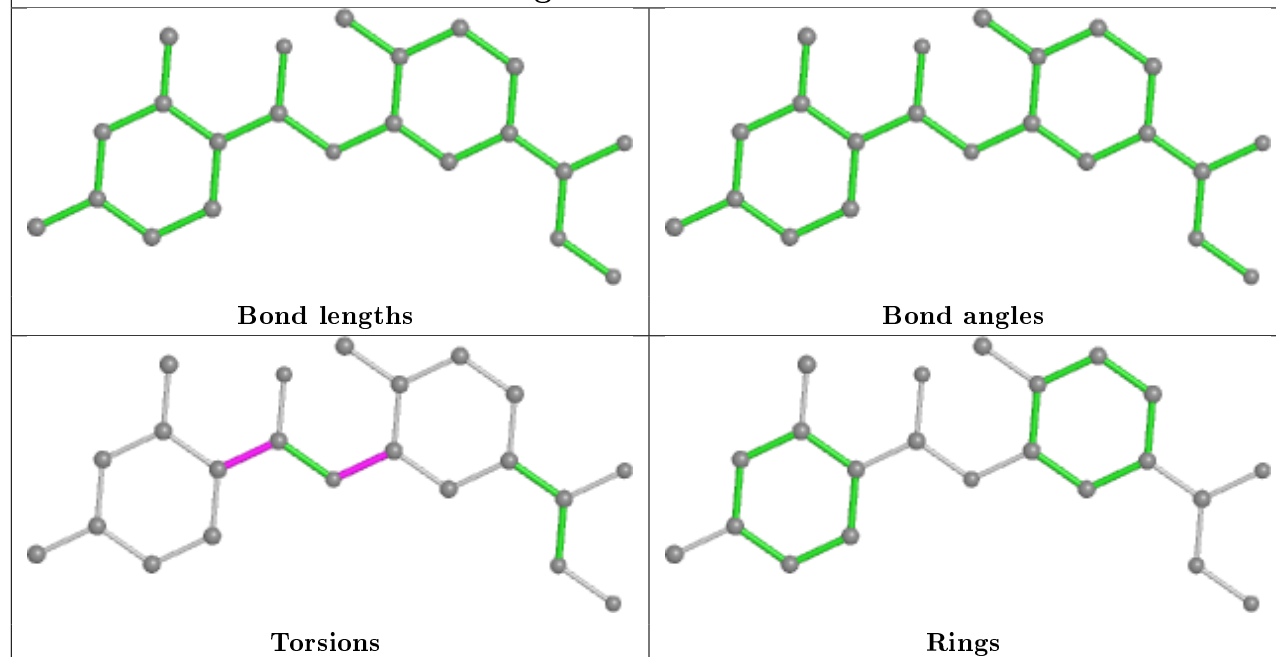
19 monomers are involved in 37 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	505	DMF	6	0
5	D	507	DMF	4	0
4	D	504	GT2	4	0
5	A	509	DMF	1	0
5	B	507	DMF	1	0
5	D	506	DMF	1	0
6	C	512	GOL	1	0
5	A	511	DMF	2	0
5	B	506	DMF	1	0
5	A	507	DMF	1	0
5	D	510	DMF	1	0
6	C	513	GOL	2	0
6	C	511	GOL	1	0
4	C	504	GT2	2	0
5	B	510	DMF	1	0
5	D	511	DMF	1	0
5	C	508	DMF	4	0
4	B	504	GT2	2	0
4	A	504	GT2	1	0

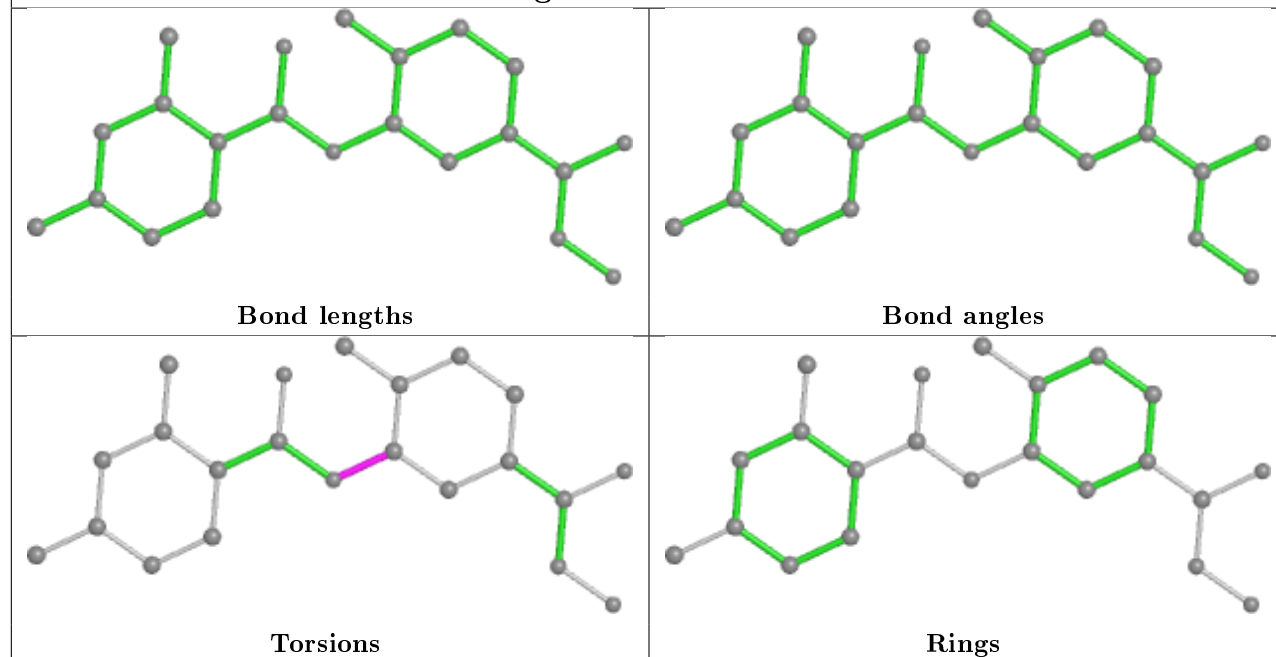
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



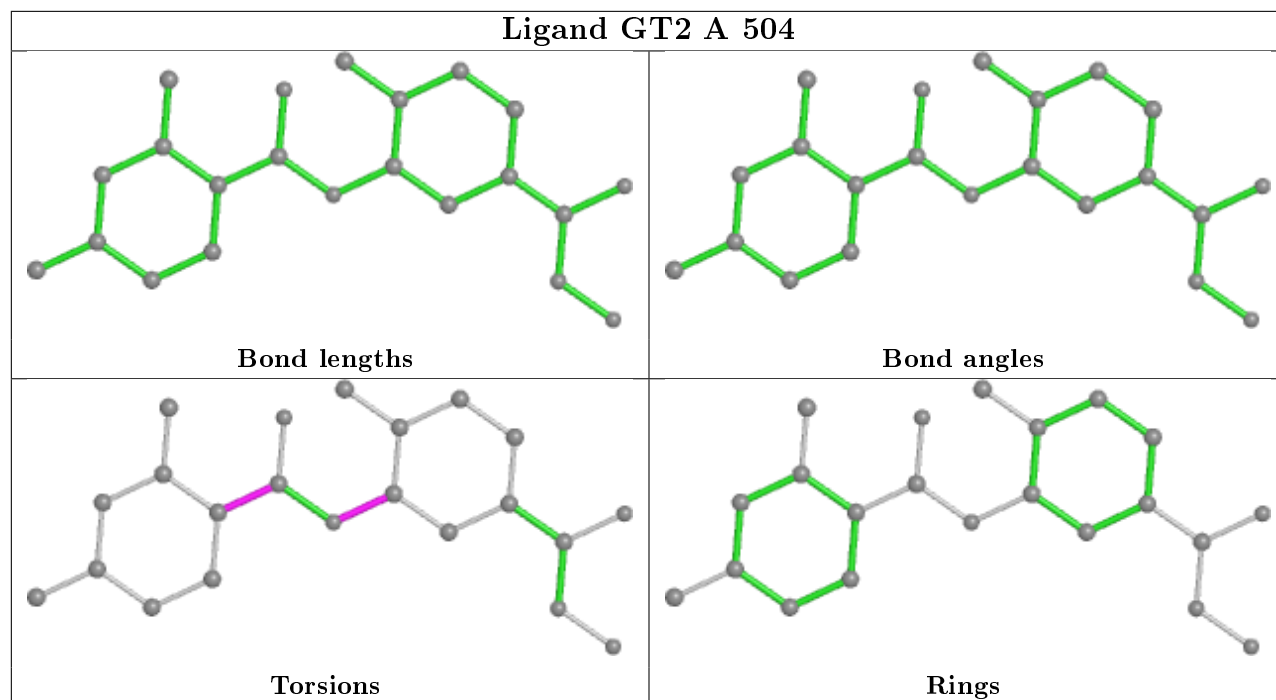
## Ligand GT2 C 504



## Ligand GT2 B 504







## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å <sup>2</sup> )	Q<0.9
1	A	399/447 (89%)	-0.06	18 (4%)	33 32	16, 25, 58, 115	0
1	B	409/447 (91%)	-0.09	10 (2%)	59 57	15, 23, 47, 79	0
1	C	410/447 (91%)	-0.06	13 (3%)	47 46	15, 24, 50, 122	0
1	D	397/447 (88%)	-0.06	20 (5%)	28 28	15, 24, 56, 118	0
All	All	1615/1788 (90%)	-0.07	61 (3%)	40 39	15, 24, 52, 122	0

All (61) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	432	TYR	12.5
1	C	81	GLU	7.9
1	D	428	TYR	7.5
1	D	426	TYR	7.4
1	A	425	ILE	7.3
1	A	426	TYR	7.1
1	A	428	TYR	7.1
1	D	427	ASP	6.9
1	D	425	ILE	6.6
1	A	433	GLN	5.6
1	C	82	LYS	5.3
1	A	81	GLU	5.2
1	A	429	ASP	5.0
1	D	81	GLU	5.0
1	C	425	ILE	4.7
1	A	427	ASP	4.2
1	D	82	LYS	3.9
1	B	425	ILE	3.9
1	B	168	THR	3.8
1	B	230	LEU	3.8
1	D	230	LEU	3.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	168	THR	3.5
1	B	446	ARG	3.5
1	D	394	HIS	3.4
1	A	409	ARG	3.3
1	A	230	LEU	3.2
1	C	83	GLU	3.2
1	D	402	ASP	3.1
1	C	426	TYR	3.1
1	C	401	LEU	3.1
1	D	430	GLN	3.0
1	B	99	TYR	3.0
1	D	409	ARG	2.9
1	A	431	VAL	2.9
1	A	404	ILE	2.9
1	A	82	LYS	2.9
1	D	416	ARG	2.8
1	A	394	HIS	2.7
1	A	395	GLU	2.6
1	A	231	PRO	2.6
1	D	406	LYS	2.6
1	C	446	ARG	2.5
1	B	177	GLN	2.5
1	D	429	ASP	2.5
1	D	224	MET	2.5
1	B	84	LEU	2.5
1	A	412	LEU	2.4
1	C	177	GLN	2.4
1	C	445	PRO	2.3
1	B	231	PRO	2.2
1	C	444	VAL	2.2
1	D	151[A]	PHE	2.2
1	D	231	PRO	2.2
1	D	83	GLU	2.2
1	D	315	SER	2.1
1	C	84	LEU	2.1
1	D	412	LEU	2.1
1	B	426	TYR	2.1
1	A	430	GLN	2.1
1	B	445	PRO	2.1
1	C	128	ARG	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	GOL	C	513	6/6	0.74	0.23	29,42,49,53	0
4	GT2	A	504	22/22	0.79	0.23	32,63,88,97	0
4	GT2	C	504	22/22	0.80	0.29	28,67,83,90	0
5	DMF	C	506	5/5	0.82	0.15	23,38,41,45	0
5	DMF	D	505	5/5	0.83	0.19	46,47,52,53	0
4	GT2	B	504	22/22	0.83	0.30	36,63,97,100	0
4	GT2	D	504	22/22	0.84	0.18	26,46,71,82	22
5	DMF	D	511	5/5	0.84	0.30	42,50,67,70	0
5	DMF	A	507	5/5	0.85	0.12	50,55,59,61	0
5	DMF	B	506	5/5	0.87	0.19	30,41,46,50	0
5	DMF	A	506	5/5	0.88	0.23	53,61,65,69	0
5	DMF	B	511	5/5	0.88	0.20	48,54,55,57	0
5	DMF	A	505	5/5	0.88	0.15	32,33,39,40	0
5	DMF	D	509	5/5	0.88	0.15	58,59,60,60	0
6	GOL	D	514	6/6	0.88	0.16	37,39,41,51	0
5	DMF	A	509	5/5	0.89	0.12	36,39,42,43	0
5	DMF	D	506	5/5	0.89	0.13	35,40,43,45	0
6	GOL	C	512	6/6	0.89	0.16	37,39,42,43	0
5	DMF	D	507	5/5	0.90	0.23	40,48,55,58	0
5	DMF	B	510	5/5	0.90	0.15	29,39,42,46	0
5	DMF	D	510	5/5	0.90	0.11	52,56,60,61	0
5	DMF	A	510	5/5	0.91	0.14	43,44,49,54	0
5	DMF	C	505	5/5	0.91	0.18	43,49,61,71	0
5	DMF	B	505	5/5	0.92	0.10	29,32,38,41	0
5	DMF	C	508	5/5	0.93	0.11	55,55,56,57	0
5	DMF	B	508	5/5	0.94	0.14	34,35,39,39	0
6	GOL	A	512	6/6	0.94	0.09	25,32,33,34	0

*Continued on next page...*

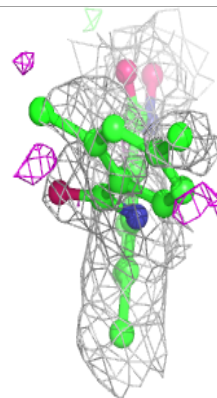
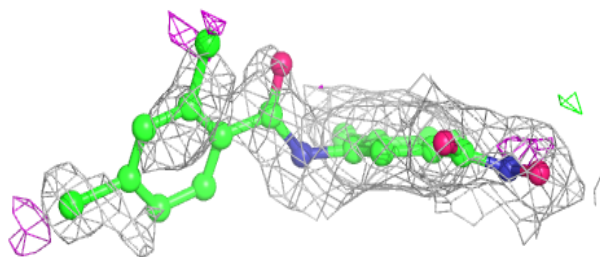
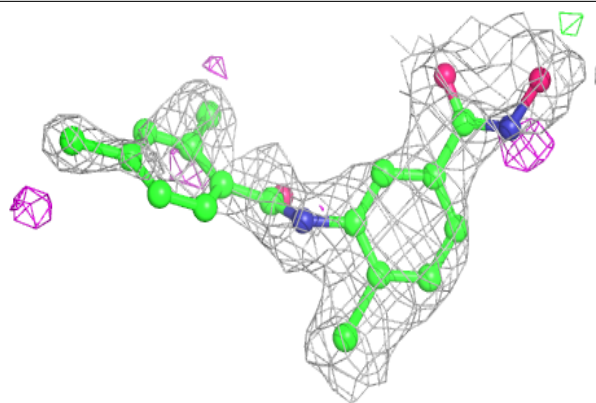
*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	GOL	D	513	6/6	0.94	0.11	23,26,28,36	0
5	DMF	B	509	5/5	0.94	0.11	29,32,41,45	0
5	DMF	C	509	5/5	0.94	0.11	34,34,37,42	0
5	DMF	B	512	5/5	0.94	0.11	31,35,44,48	0
5	DMF	D	512	5/5	0.94	0.13	44,44,46,47	0
5	DMF	B	507	5/5	0.94	0.13	45,45,46,55	0
5	DMF	C	510	5/5	0.95	0.12	27,30,32,36	0
5	DMF	A	511	5/5	0.95	0.10	27,39,48,57	0
5	DMF	C	507	5/5	0.95	0.09	30,34,35,38	0
5	DMF	A	508	5/5	0.96	0.16	33,36,42,42	0
5	DMF	D	508	5/5	0.96	0.13	28,40,42,44	0
6	GOL	C	511	6/6	0.97	0.09	22,31,37,44	0
3	K	D	502	1/1	0.98	0.12	26,26,26,26	0
6	GOL	B	513	6/6	0.98	0.07	22,26,26,28	0
2	ZN	D	501	1/1	0.99	0.04	38,38,38,38	0
3	K	C	503	1/1	0.99	0.09	24,24,24,24	0
2	ZN	C	501	1/1	0.99	0.04	39,39,39,39	0
2	ZN	B	501	1/1	0.99	0.05	38,38,38,38	0
2	ZN	A	501	1/1	1.00	0.04	38,38,38,38	0
3	K	C	502	1/1	1.00	0.11	16,16,16,16	0
3	K	B	502	1/1	1.00	0.12	22,22,22,22	0
3	K	D	503	1/1	1.00	0.10	17,17,17,17	0
3	K	A	502	1/1	1.00	0.11	19,19,19,19	0
3	K	A	503	1/1	1.00	0.09	22,22,22,22	0
3	K	B	503	1/1	1.00	0.12	18,18,18,18	0

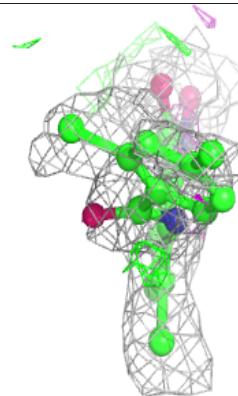
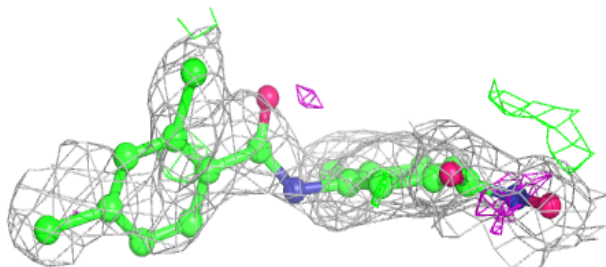
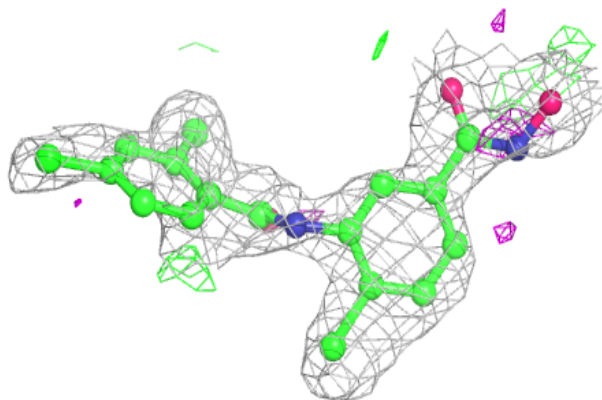
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around GT2 A 504:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

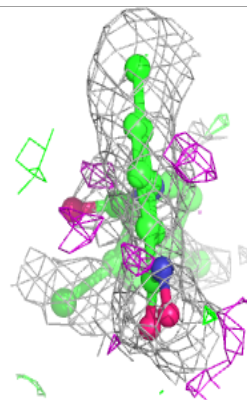
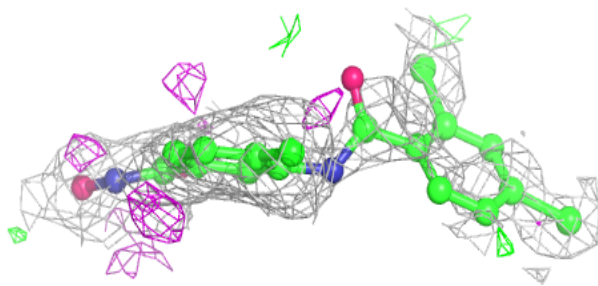
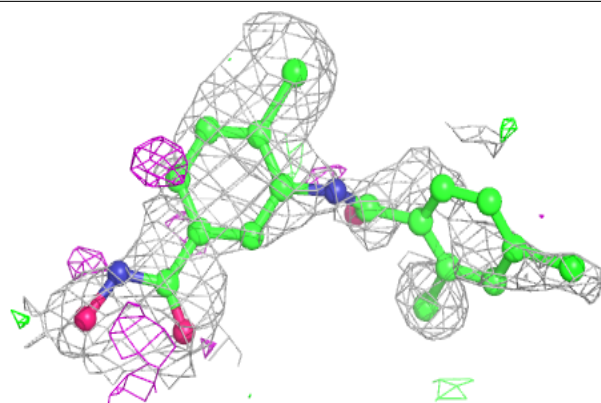
**Electron density around GT2 C 504:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

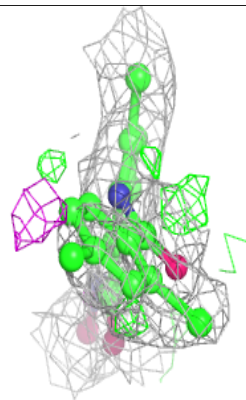
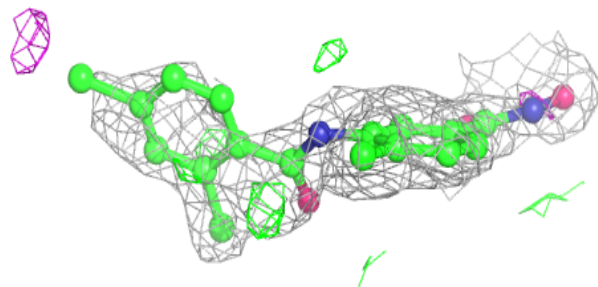
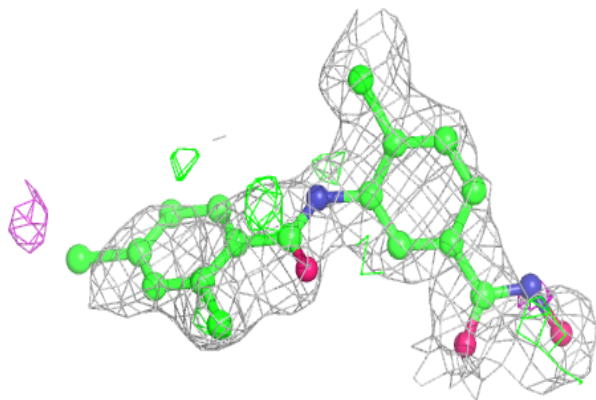


**Electron density around GT2 B 504:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around GT2 D 504:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers

There are no such residues in this entry.