



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 08:24 am BST

PDB ID : 6HQY  
Title : Crystal structure of Schistosoma mansoni HDAC8 complexed with PCI-34051  
Authors : Marek, M.; Shaik, T.B.; Romier, C.  
Deposited on : 2018-09-25  
Resolution : 2.50 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

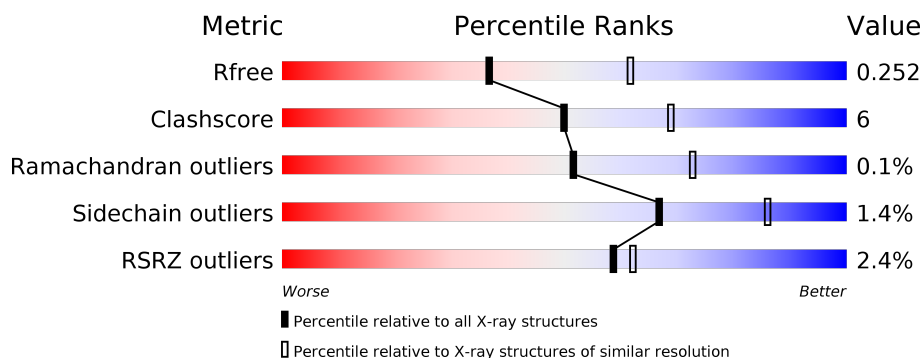
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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>76%</div> <div>12%</div> <div>11%</div> </div> </div>
1	B	447	<div> <div>%</div> <div> <div></div> <div>78%</div> <div>14%</div> <div>9%</div> </div> </div>
1	C	447	<div> <div>%</div> <div> <div></div> <div>80%</div> <div>12%</div> <div>8%</div> </div> </div>
1	D	447	<div> <div>2%</div> <div> <div></div> <div>76%</div> <div>12%</div> <div>12%</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	GOL	B	505	-	-	X	-
5	GOL	C	505	-	-	X	-
5	GOL	D	505	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 13230 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	397	Total	C	N	O	S	0	1	0
			3188	2057	532	584	15			
1	B	409	Total	C	N	O	S	0	0	0
			3265	2105	547	598	15			
1	C	411	Total	C	N	O	S	0	0	0
			3280	2114	549	601	16			
1	D	395	Total	C	N	O	S	0	0	0
			3164	2042	527	581	14			

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

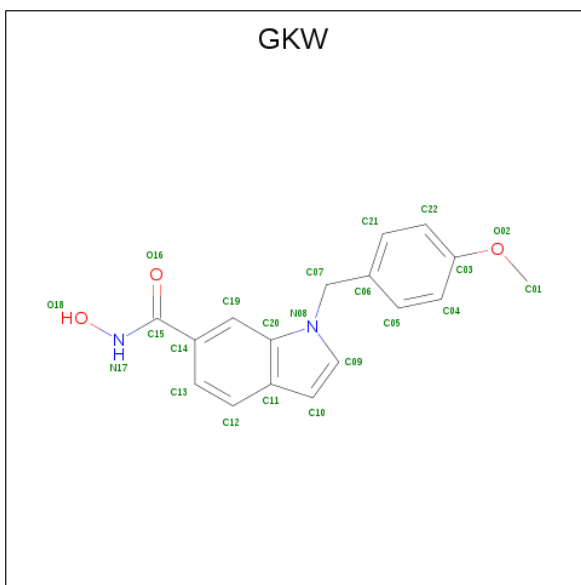
- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

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).

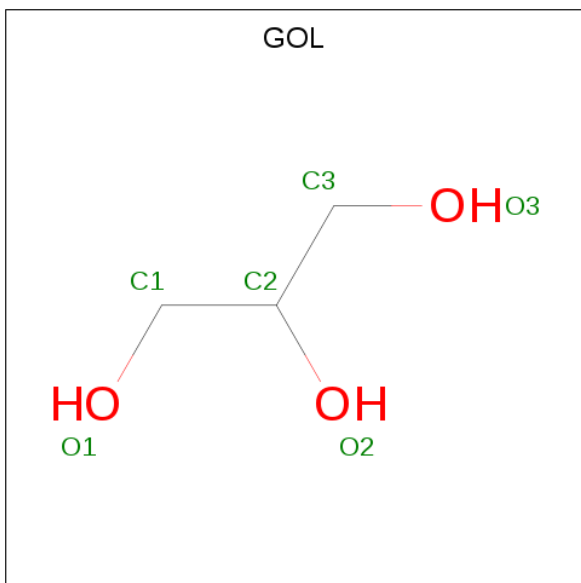
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total K 2 2	0	0
3	A	2	Total K 2 2	0	0
3	D	2	Total K 2 2	0	0
3	C	2	Total K 2 2	0	0

- Molecule 4 is 1-[(4-methoxyphenyl)methyl]-{N}-oxidanyl-indole-6-carboxamide (three-letter code: GKW) (formula: C<sub>17</sub>H<sub>16</sub>N<sub>2</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 22	C 17	N 2	O 3	0	0
4	B	1	Total 22	C 17	N 2	O 3	0	0
4	C	1	Total 22	C 17	N 2	O 3	0	0
4	D	1	Total 22	C 17	N 2	O 3	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $\text{C}_3\text{H}_8\text{O}_3$ ).



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

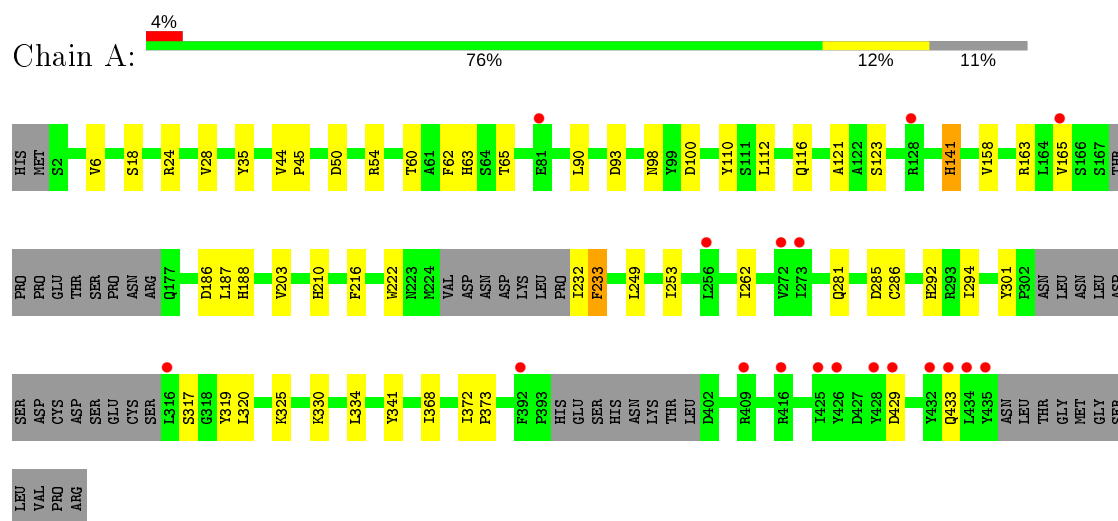
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	55	Total O 55 55	0	0
6	B	53	Total O 53 53	0	0
6	C	53	Total O 53 53	0	0
6	D	36	Total O 36 36	0	0

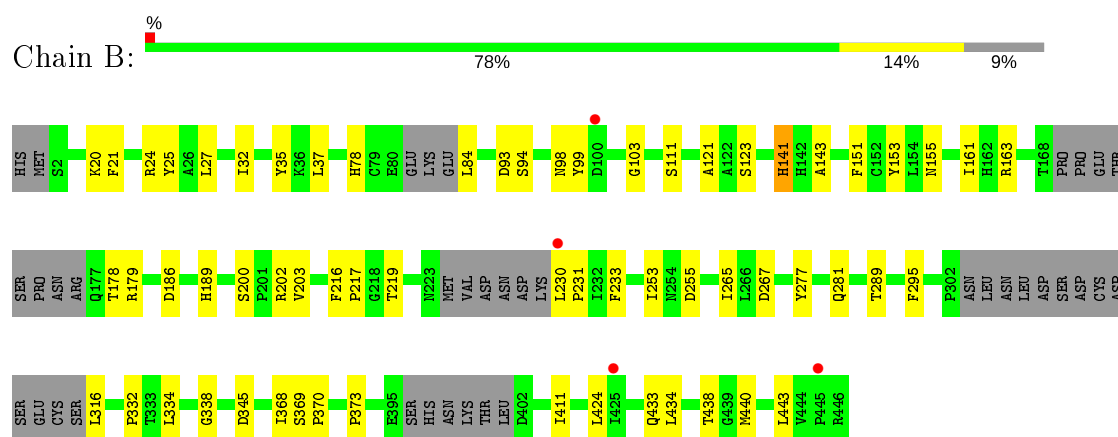
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

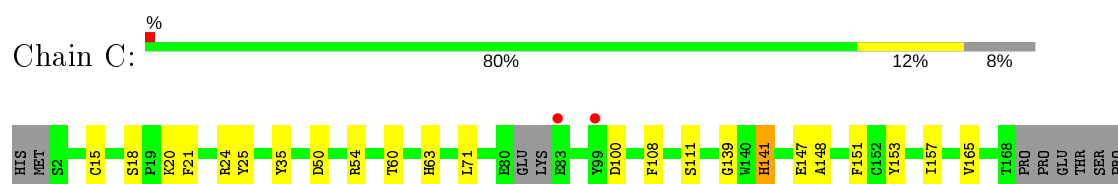
#### • Molecule 1: Histone deacetylase



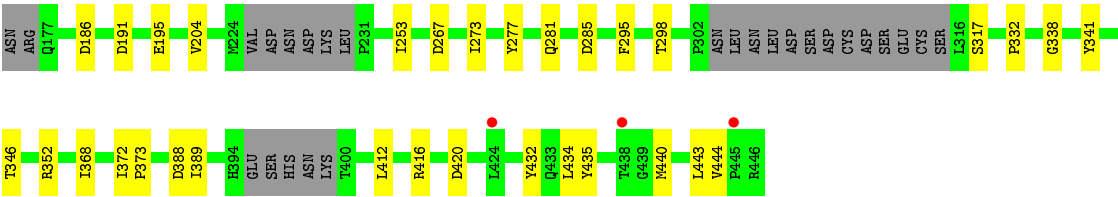
#### • Molecule 1: Histone deacetylase



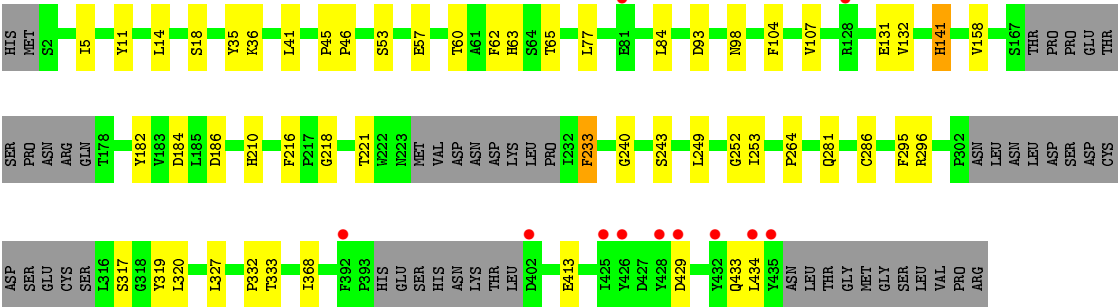
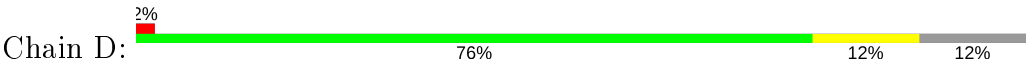
#### • Molecule 1: Histone deacetylase







● Molecule 1: Histone deacetylase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.88Å 71.05Å 98.20Å 78.06° 75.61° 85.45°	Depositor
Resolution (Å)	42.76 – 2.50 49.90 – 2.50	Depositor EDS
% Data completeness (in resolution range)	97.1 (42.76-2.50) 97.1 (49.90-2.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.01 (at 2.51Å)	Xtriage
Refinement program	PHENIX dev_1796	Depositor
R, $R_{free}$	0.180 , 0.254 0.186 , 0.252	Depositor DCC
$R_{free}$ test set	3089 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.2	Xtriage
Anisotropy	0.227	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 60.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.119 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13230	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.70% 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, K, GKW

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	0.40	0/3281	0.52	0/4461
1	B	0.43	0/3356	0.54	0/4565
1	C	0.43	0/3371	0.54	0/4584
1	D	0.40	0/3253	0.53	0/4424
All	All	0.42	0/13261	0.53	0/18034

There are no bond length outliers.

There are no bond angle outliers.

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	3188	0	3079	35	0
1	B	3265	0	3160	38	0
1	C	3280	0	3177	39	0
1	D	3164	0	3056	35	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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	0	0
4	C	22	0	0	1	0
4	D	22	0	0	0	0
5	A	18	0	24	3	0
5	B	6	0	8	5	0
5	C	6	0	8	7	0
5	D	6	0	8	4	0
6	A	55	0	0	1	0
6	B	53	0	0	1	0
6	C	53	0	0	2	0
6	D	36	0	0	0	0
All	All	13230	0	12520	147	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 6.

All (147) 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:B:24:ARG:HE	5:B:505:GOL:H31	1.30	0.96
1:C:24:ARG:HE	5:C:505:GOL:H12	1.27	0.94
1:A:249:LEU:HD13	1:A:253:ILE:HD13	1.63	0.81
1:B:267:ASP:HB3	1:B:434:LEU:HD21	1.67	0.75
1:A:216:PHE:O	5:A:506:GOL:H32	1.90	0.72
1:A:100:ASP:HB2	5:A:505:GOL:H32	1.73	0.71
1:A:317:SER:HB3	1:A:320:LEU:HG	1.73	0.68
1:B:93:ASP:OD1	1:B:98:ASN:ND2	2.26	0.67
1:C:21:PHE:HB3	5:C:505:GOL:H31	1.76	0.66
1:B:35:TYR:CE1	1:B:368:ILE:HG23	2.34	0.62
1:C:35:TYR:CE1	1:C:368:ILE:HG23	2.34	0.62
1:D:317:SER:HB3	1:D:320:LEU:HG	1.81	0.61
1:B:178:THR:HB	1:B:202:ARG:HH21	1.65	0.60
1:B:24:ARG:NE	5:B:505:GOL:H31	2.10	0.59
1:C:186:ASP:HB2	1:C:281:GLN:OE1	2.03	0.58
1:B:141:HIS:CD2	1:B:141:HIS:H	2.21	0.58
1:D:249:LEU:HD13	1:D:253:ILE:HD13	1.86	0.57
1:C:253:ILE:HG22	1:C:295:PHE:CD1	2.40	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:123:SER:OG	1:B:163:ARG:NH1	2.38	0.56
1:A:121:ALA:HB1	1:A:334:LEU:HD11	1.87	0.56
1:D:216:PHE:O	5:D:505:GOL:H11	2.05	0.56
1:B:21:PHE:HB3	5:B:505:GOL:H12	1.88	0.55
1:B:265:ILE:HD13	1:B:411:ILE:HD13	1.88	0.55
1:B:440:MET:HB3	1:B:443:LEU:HD12	1.89	0.55
1:B:289:THR:OG1	1:B:345:ASP:OD2	2.18	0.54
1:B:338:GLY:N	6:B:604:HOH:O	2.37	0.54
1:D:218:GLY:HA2	5:D:505:GOL:H32	1.90	0.54
1:B:25:TYR:N	5:B:505:GOL:O1	2.27	0.54
1:D:210:HIS:O	5:D:505:GOL:H12	2.08	0.53
1:C:15:CYS:O	5:C:505:GOL:H32	2.08	0.53
1:B:189:HIS:ND1	1:B:219:THR:OG1	2.32	0.53
1:C:20:LYS:HE3	1:C:341:TYR:HE2	1.74	0.53
1:D:60:THR:HG21	1:D:65:THR:HG22	1.91	0.52
1:B:277:TYR:CD1	1:B:332:PRO:HB2	2.44	0.52
1:C:141:HIS:H	1:C:141:HIS:CD2	2.28	0.51
1:C:24:ARG:NH2	1:C:338:GLY:O	2.27	0.51
1:B:267:ASP:CB	1:B:434:LEU:HD21	2.38	0.50
1:A:429:ASP:O	1:A:433:GLN:HG2	2.11	0.50
1:C:412:LEU:HD11	1:C:435:TYR:CZ	2.47	0.50
1:A:123:SER:OG	1:A:163:ARG:NH1	2.45	0.50
1:A:24:ARG:NH2	6:A:609:HOH:O	2.44	0.50
1:C:341:TYR:OH	4:C:504:GKW:O16	2.25	0.49
1:A:186:ASP:HB2	1:A:281:GLN:OE1	2.11	0.49
1:D:62:PHE:CG	1:D:158:VAL:HG11	2.48	0.49
1:D:240:GLY:O	1:D:243:SER:OG	2.29	0.49
1:A:93:ASP:OD1	1:A:98:ASN:ND2	2.35	0.49
1:C:277:TYR:CD1	1:C:332:PRO:HB2	2.48	0.49
1:B:20:LYS:HE2	1:B:20:LYS:HB3	1.59	0.48
1:B:143:ALA:HB3	1:B:155:ASN:HB2	1.95	0.48
1:B:161:ILE:HG21	1:B:203:VAL:HG11	1.95	0.48
1:D:286:CYS:HB2	1:D:319:TYR:OH	2.14	0.48
1:C:420:ASP:OD2	6:C:601:HOH:O	2.20	0.48
1:A:60:THR:HG21	1:A:65:THR:HG22	1.96	0.47
1:A:112:LEU:O	1:A:116:GLN:HG3	2.14	0.47
1:B:32:ILE:HG23	1:B:37:LEU:HB2	1.97	0.47
1:D:5:ILE:HG12	1:D:41:LEU:HB3	1.95	0.47
1:C:204:VAL:HG21	1:C:273:ILE:HD13	1.97	0.46
1:C:71:LEU:HG	1:C:108:PHE:HD1	1.80	0.46
1:C:24:ARG:NE	5:C:505:GOL:H12	2.12	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:252:GLY:O	1:D:296:ARG:HD2	2.15	0.46
1:D:53:SER:O	1:D:57:GLU:HG3	2.17	0.45
1:A:334:LEU:HD12	1:A:334:LEU:HA	1.85	0.45
1:A:60:THR:HA	1:A:63:HIS:O	2.17	0.45
1:D:186:ASP:HB2	1:D:281:GLN:OE1	2.17	0.45
1:D:35:TYR:CE1	1:D:368:ILE:HG23	2.50	0.45
1:D:429:ASP:O	1:D:433:GLN:HG2	2.16	0.45
1:B:121:ALA:HB1	1:B:334:LEU:HD11	1.98	0.45
1:A:222:TRP:HE3	1:A:232:ILE:HA	1.82	0.45
1:D:60:THR:HA	1:D:63:HIS:O	2.16	0.45
1:C:25:TYR:HB2	5:C:505:GOL:H11	1.98	0.45
1:D:62:PHE:CD1	1:D:158:VAL:HG11	2.51	0.45
1:C:298:THR:HA	1:C:389:ILE:HD11	1.99	0.45
1:B:277:TYR:HD1	1:B:332:PRO:HB2	1.82	0.45
1:C:139:GLY:HA2	1:C:157:ILE:HD11	1.99	0.45
1:A:18:SER:HB3	1:A:110:TYR:CE1	2.52	0.44
1:C:18:SER:HB2	1:C:151:PHE:HZ	1.82	0.44
1:C:100:ASP:O	6:C:602:HOH:O	2.21	0.44
1:D:264:PRO:HB3	1:D:434:LEU:O	2.17	0.44
1:D:77:LEU:HD13	1:D:84:LEU:HG	1.99	0.44
1:D:93:ASP:OD1	1:D:98:ASN:ND2	2.40	0.44
1:C:440:MET:HB3	1:C:443:LEU:HD12	2.00	0.44
1:C:412:LEU:HD11	1:C:435:TYR:CE1	2.53	0.44
1:D:45:PRO:HA	1:D:46:PRO:HD3	1.90	0.44
1:A:325:LYS:HB2	1:A:325:LYS:HE2	1.74	0.43
1:B:334:LEU:HD12	1:B:334:LEU:HA	1.84	0.43
1:A:286:CYS:HB2	1:A:319:TYR:OH	2.18	0.43
1:C:50:ASP:HB2	1:C:54:ARG:HD2	2.00	0.43
1:B:111:SER:OG	1:B:153:TYR:HB2	2.18	0.43
1:B:179:ARG:NH2	1:B:424:LEU:HD21	2.33	0.43
1:B:186:ASP:HB2	1:B:281:GLN:OE1	2.19	0.43
1:A:262:ILE:H	1:A:262:ILE:HG13	1.70	0.43
1:A:187:LEU:HD23	1:A:188:HIS:CE1	2.53	0.43
1:B:253:ILE:HG22	1:B:295:PHE:CD1	2.54	0.43
1:C:267:ASP:HB3	1:C:434:LEU:HD11	2.01	0.43
1:C:432:TYR:HA	1:C:444:VAL:HG21	2.01	0.43
1:A:35:TYR:CE1	1:A:368:ILE:HG23	2.54	0.43
1:A:210:HIS:O	5:A:506:GOL:H31	2.19	0.43
1:D:104:PHE:CE2	1:D:107:VAL:HA	2.53	0.43
1:B:20:LYS:HD3	1:B:151:PHE:CD1	2.54	0.43
1:B:20:LYS:HB2	1:B:151:PHE:CZ	2.54	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:141:HIS:CD2	1:D:141:HIS:H	2.37	0.43
1:B:78:HIS:CE1	1:B:103:GLY:HA3	2.54	0.42
1:D:131:GLU:O	1:D:332:PRO:HA	2.19	0.42
1:A:141:HIS:H	1:A:141:HIS:CD2	2.36	0.42
1:B:35:TYR:CZ	1:B:373:PRO:HD3	2.53	0.42
1:C:111:SER:OG	1:C:153:TYR:HB2	2.18	0.42
1:C:191:ASP:O	1:C:195:GLU:HB2	2.19	0.42
1:A:330:LYS:HD3	1:A:330:LYS:HA	1.90	0.42
1:A:6:VAL:HA	1:A:44:VAL:O	2.19	0.42
1:B:189:HIS:CE1	1:B:219:THR:HG1	2.33	0.42
1:D:14:LEU:HA	1:D:14:LEU:HD23	1.87	0.42
1:C:352:ARG:NH1	1:C:388:ASP:OD1	2.53	0.42
1:C:25:TYR:N	5:C:505:GOL:H11	2.35	0.42
1:D:182:TYR:CE2	1:D:184:ASP:HB2	2.55	0.42
1:A:165:VAL:HG11	1:A:203:VAL:CG2	2.50	0.42
1:C:191:ASP:OD1	1:C:191:ASP:N	2.53	0.42
1:D:253:ILE:HG22	1:D:295:PHE:CD1	2.54	0.42
1:C:277:TYR:HD1	1:C:332:PRO:HB2	1.85	0.41
1:D:221:THR:HA	1:D:233:PHE:CZ	2.55	0.41
1:D:36:LYS:HA	1:D:36:LYS:HD3	1.80	0.41
1:A:44:VAL:HA	1:A:45:PRO:HD2	1.90	0.41
1:A:50:ASP:HB2	1:A:54:ARG:HG3	2.03	0.41
1:B:216:PHE:HA	1:B:217:PRO:HA	1.81	0.41
1:B:24:ARG:HE	5:B:505:GOL:C3	2.16	0.41
1:B:369:SER:HA	1:B:370:PRO:HD3	1.91	0.41
1:D:327:LEU:HA	1:D:327:LEU:HD23	1.90	0.41
1:C:18:SER:HB2	1:C:151:PHE:CZ	2.56	0.41
1:A:341:TYR:OH	4:A:504:GKW:O16	2.27	0.41
1:C:24:ARG:HD2	1:C:346:THR:HG21	2.03	0.41
1:C:21:PHE:CB	5:C:505:GOL:H31	2.49	0.41
1:D:11:TYR:HB2	1:D:46:PRO:HG2	2.02	0.41
1:A:62:PHE:CD1	1:A:158:VAL:HG11	2.56	0.41
1:A:372:ILE:HA	1:A:373:PRO:HD3	1.92	0.41
1:C:298:THR:O	1:C:317:SER:OG	2.27	0.41
1:D:218:GLY:CA	5:D:505:GOL:H32	2.50	0.41
1:D:93:ASP:HA	1:D:98:ASN:ND2	2.35	0.41
1:A:90:LEU:HA	1:A:90:LEU:HD23	1.76	0.41
1:A:232:ILE:HB	1:A:233:PHE:H	1.68	0.41
1:A:292:HIS:HB3	1:A:294:ILE:HG12	2.03	0.41
1:C:372:ILE:HA	1:C:373:PRO:HD3	1.92	0.41
1:C:147:GLU:HG2	1:C:148:ALA:O	2.22	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:60:THR:HA	1:C:63:HIS:O	2.22	0.40
1:B:255:ASP:HB3	1:B:316:LEU:O	2.22	0.40
1:A:24:ARG:O	1:A:28:VAL:HG23	2.22	0.40
1:D:132:VAL:HG22	1:D:333:THR:HB	2.03	0.40
1:B:230:LEU:HA	1:B:231:PRO:HD3	1.97	0.40
1:D:317:SER:CB	1:D:320:LEU:HG	2.50	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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	388/447 (87%)	377 (97%)	11 (3%)	0	100	100
1	B	397/447 (89%)	387 (98%)	9 (2%)	1 (0%)	41	61
1	C	399/447 (89%)	386 (97%)	13 (3%)	0	100	100
1	D	385/447 (86%)	372 (97%)	13 (3%)	0	100	100
All	All	1569/1788 (88%)	1522 (97%)	46 (3%)	1 (0%)	51	73

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	99	TYR

### 5.3.2 Protein sidechains [i](#)

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	345/392 (88%)	341 (99%)	4 (1%)	71	88
1	B	354/392 (90%)	346 (98%)	8 (2%)	50	76
1	C	356/392 (91%)	352 (99%)	4 (1%)	73	89
1	D	342/392 (87%)	338 (99%)	4 (1%)	71	88
All	All	1397/1568 (89%)	1377 (99%)	20 (1%)	67	86

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	141	HIS
1	A	233	PHE
1	A	285	ASP
1	A	301	TYR
1	B	27	LEU
1	B	84	LEU
1	B	94	SER
1	B	141	HIS
1	B	200	SER
1	B	233	PHE
1	B	433	GLN
1	B	438	THR
1	C	141	HIS
1	C	165	VAL
1	C	285	ASP
1	C	416	ARG
1	D	18	SER
1	D	141	HIS
1	D	233	PHE
1	D	413	GLU

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 [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 12 are monoatomic - leaving 10 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	GOL	A	505	-	5,5,5	0.68	0	5,5,5	1.62	2 (40%)
5	GOL	B	505	-	5,5,5	0.79	0	5,5,5	1.48	1 (20%)
4	GKW	D	504	2	24,24,24	2.56	4 (16%)	30,33,33	0.92	2 (6%)
5	GOL	D	505	-	5,5,5	0.76	0	5,5,5	2.04	3 (60%)
4	GKW	B	504	2	24,24,24	2.65	4 (16%)	30,33,33	1.07	2 (6%)
5	GOL	A	507	-	5,5,5	0.88	0	5,5,5	1.16	0
4	GKW	A	504	2	24,24,24	2.57	5 (20%)	30,33,33	1.05	3 (10%)
5	GOL	A	506	-	5,5,5	0.71	0	5,5,5	1.72	2 (40%)
4	GKW	C	504	2	24,24,24	2.54	3 (12%)	30,33,33	1.12	3 (10%)
5	GOL	C	505	-	5,5,5	0.79	0	5,5,5	1.65	1 (20%)

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	GOL	A	505	-	-	3/4/4/4	-
5	GOL	B	505	-	-	0/4/4/4	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GKW	D	504	2	-	4/12/12/12	0/3/3/3
5	GOL	D	505	-	-	2/4/4/4	-
4	GKW	B	504	2	-	4/12/12/12	0/3/3/3
5	GOL	A	507	-	-	2/4/4/4	-
4	GKW	A	504	2	-	2/12/12/12	0/3/3/3
5	GOL	A	506	-	-	2/4/4/4	-
4	GKW	C	504	2	-	4/12/12/12	0/3/3/3
5	GOL	C	505	-	-	4/4/4/4	-

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	504	GKW	C15-N17	9.77	1.45	1.32
4	B	504	GKW	C15-N17	9.67	1.45	1.32
4	A	504	GKW	C15-N17	9.63	1.45	1.32
4	C	504	GKW	C15-N17	9.30	1.44	1.32
4	B	504	GKW	C14-C15	4.25	1.59	1.50
4	B	504	GKW	O18-N17	-4.18	1.29	1.40
4	C	504	GKW	O18-N17	-4.10	1.29	1.40
4	A	504	GKW	O18-N17	-4.04	1.29	1.40
4	D	504	GKW	O18-N17	-3.98	1.30	1.40
4	C	504	GKW	C14-C15	3.69	1.57	1.50
4	A	504	GKW	C14-C15	3.49	1.57	1.50
4	D	504	GKW	C14-C15	3.41	1.57	1.50
4	B	504	GKW	C20-N08	2.51	1.42	1.39
4	A	504	GKW	C20-N08	2.14	1.41	1.39
4	A	504	GKW	C22-C21	2.06	1.42	1.38
4	D	504	GKW	C20-N08	2.05	1.41	1.39

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	504	GKW	C14-C15-N17	3.68	122.06	116.16
4	B	504	GKW	O16-C15-N17	-3.34	116.71	122.94
4	D	504	GKW	C14-C15-N17	3.28	121.43	116.16
5	D	505	GOL	O3-C3-C2	3.11	125.13	110.20
4	C	504	GKW	O16-C15-N17	-3.05	117.24	122.94
4	A	504	GKW	C14-C15-N17	2.97	120.92	116.16
5	A	506	GOL	O1-C1-C2	2.62	122.78	110.20
4	A	504	GKW	C06-C07-N08	2.59	116.94	112.71

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	504	GKW	C14-C15-N17	2.58	120.29	116.16
5	C	505	GOL	O3-C3-C2	2.51	122.24	110.20
5	A	506	GOL	O3-C3-C2	2.44	121.88	110.20
5	D	505	GOL	O1-C1-C2	2.25	120.98	110.20
5	A	505	GOL	O3-C3-C2	2.24	120.95	110.20
4	C	504	GKW	C06-C07-N08	2.23	116.35	112.71
4	A	504	GKW	O16-C15-N17	-2.19	118.86	122.94
5	D	505	GOL	C3-C2-C1	2.15	120.05	111.70
4	D	504	GKW	O16-C15-N17	-2.11	119.00	122.94
5	B	505	GOL	O1-C1-C2	2.04	119.96	110.20
5	A	505	GOL	O1-C1-C2	2.00	119.80	110.20

There are no chirality outliers.

All (27) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	505	GOL	O1-C1-C2-C3
5	C	505	GOL	C1-C2-C3-O3
5	A	505	GOL	O2-C2-C3-O3
5	D	505	GOL	C1-C2-C3-O3
5	A	506	GOL	O1-C1-C2-C3
4	D	504	GKW	C22-C03-O02-C01
4	D	504	GKW	C04-C03-O02-C01
4	C	504	GKW	C22-C03-O02-C01
4	C	504	GKW	C04-C03-O02-C01
5	A	507	GOL	C1-C2-C3-O3
5	A	505	GOL	C1-C2-C3-O3
5	A	507	GOL	O2-C2-C3-O3
5	C	505	GOL	O1-C1-C2-O2
5	C	505	GOL	O2-C2-C3-O3
5	A	506	GOL	O1-C1-C2-O2
4	A	504	GKW	C21-C06-C07-N08
4	A	504	GKW	C05-C06-C07-N08
4	B	504	GKW	C04-C03-O02-C01
4	B	504	GKW	C22-C03-O02-C01
4	B	504	GKW	C21-C06-C07-N08
4	B	504	GKW	C05-C06-C07-N08
5	D	505	GOL	O2-C2-C3-O3
5	A	505	GOL	O1-C1-C2-O2
4	D	504	GKW	C05-C06-C07-N08
4	D	504	GKW	C21-C06-C07-N08
4	C	504	GKW	C05-C06-C07-N08

*Continued on next page...*

*Continued from previous page...*

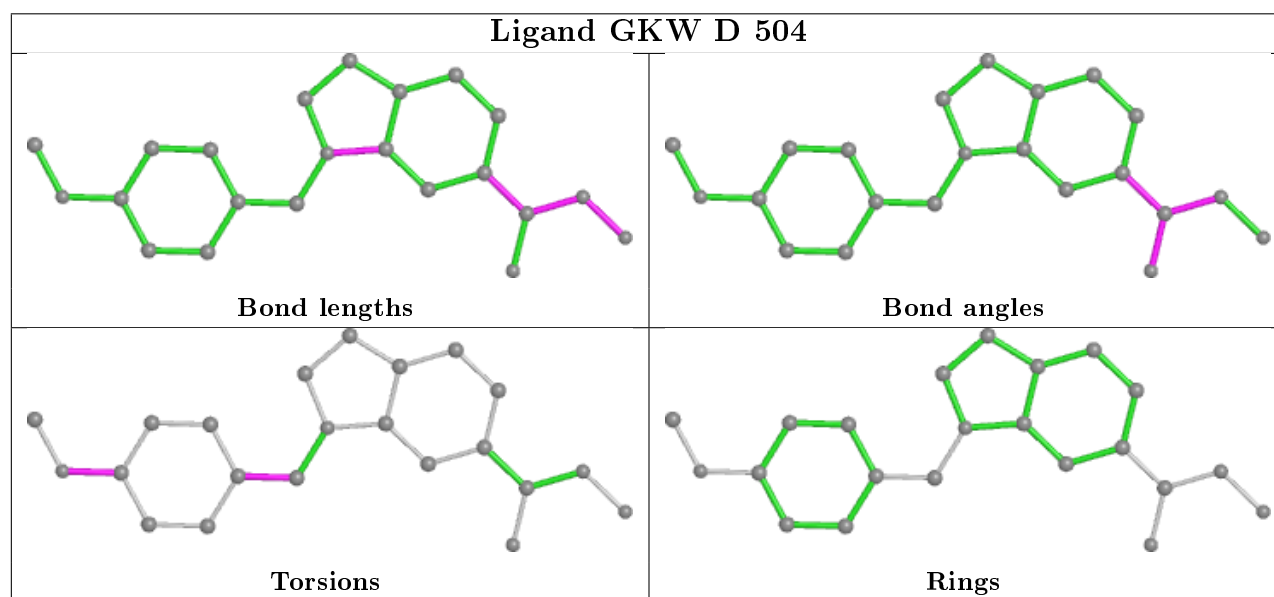
Mol	Chain	Res	Type	Atoms
4	C	504	GKW	C21-C06-C07-N08

There are no ring outliers.

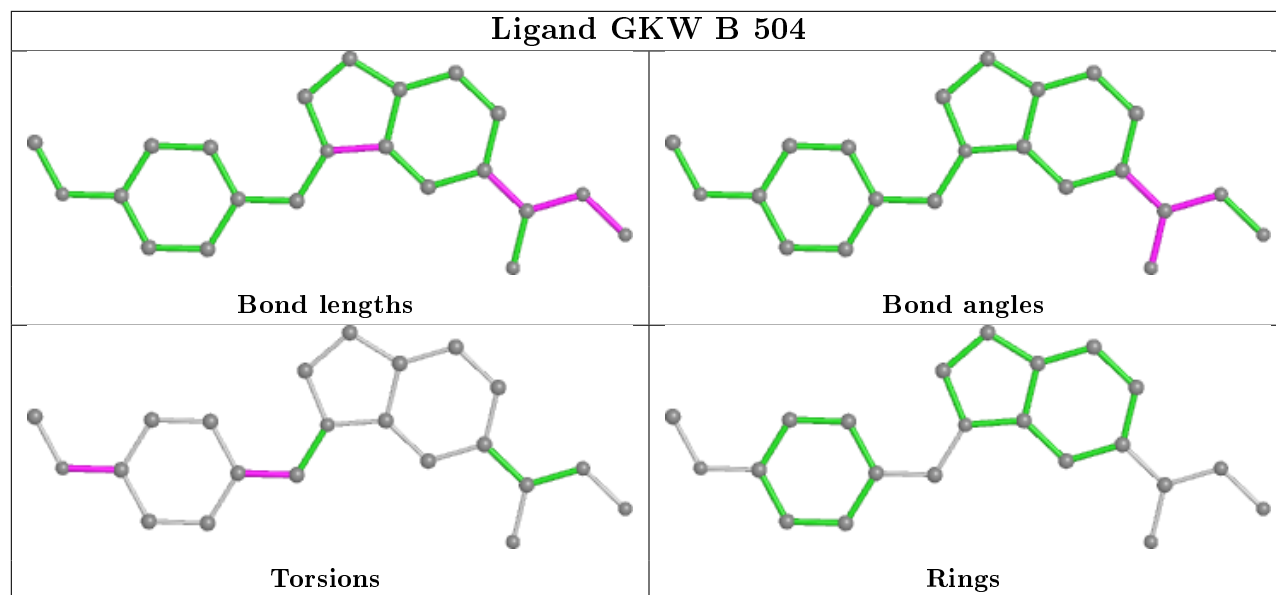
7 monomers are involved in 21 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	505	GOL	1	0
5	B	505	GOL	5	0
5	D	505	GOL	4	0
4	A	504	GKW	1	0
5	A	506	GOL	2	0
4	C	504	GKW	1	0
5	C	505	GOL	7	0

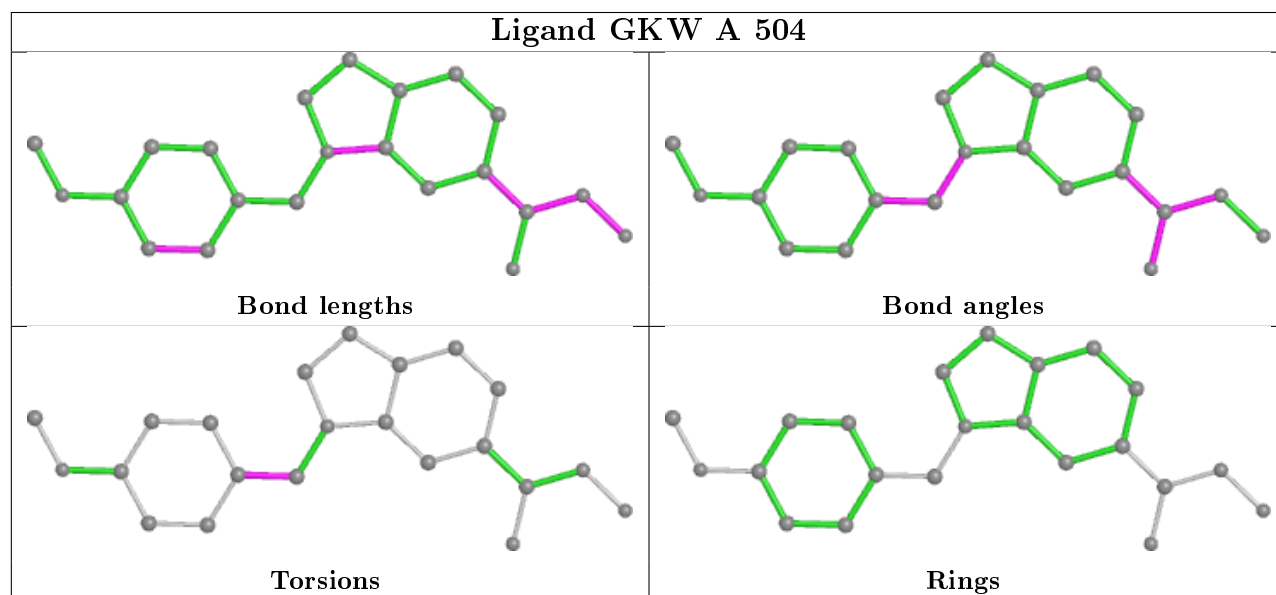
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.



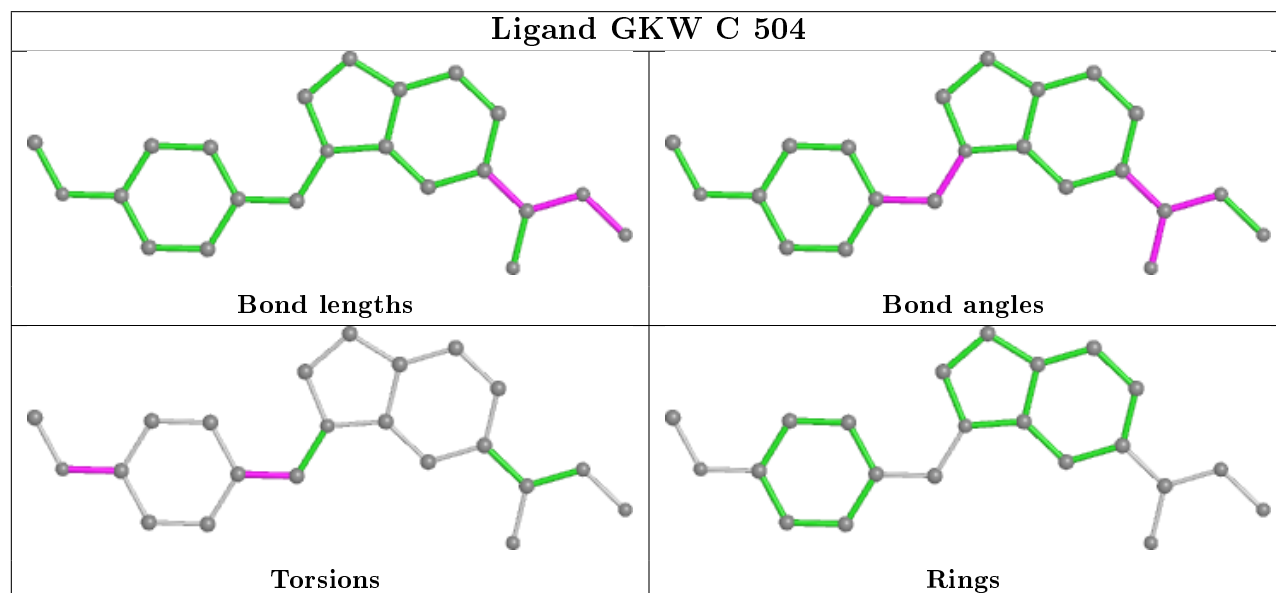
## Ligand GKW B 504



## Ligand GKW A 504



## Ligand GKW C 504



## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	397/447 (88%)	0.15	18 (4%) 33 36	14, 28, 58, 100	0
1	B	409/447 (91%)	-0.06	4 (0%) 82 84	9, 24, 49, 76	0
1	C	411/447 (91%)	-0.09	5 (1%) 79 80	10, 24, 50, 97	0
1	D	395/447 (88%)	0.12	11 (2%) 53 56	13, 27, 56, 100	0
All	All	1612/1788 (90%)	0.03	38 (2%) 59 62	9, 26, 54, 100	0

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	428	TYR	5.9
1	A	432	TYR	5.3
1	A	425	ILE	5.3
1	A	434	LEU	5.2
1	D	426	TYR	4.8
1	A	426	TYR	4.0
1	A	435	TYR	3.9
1	D	425	ILE	3.8
1	D	435	TYR	3.5
1	A	429	ASP	3.3
1	A	433	GLN	3.3
1	C	445	PRO	3.2
1	B	230	LEU	3.1
1	A	416	ARG	3.1
1	D	434	LEU	3.0
1	A	392	PHE	2.9
1	A	81	GLU	2.9
1	A	272	VAL	2.8
1	D	432	TYR	2.8
1	D	392	PHE	2.6
1	C	438	THR	2.6

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	165	VAL	2.5
1	C	99	TYR	2.5
1	D	81	GLU	2.5
1	B	425	ILE	2.5
1	D	428	TYR	2.3
1	A	256	LEU	2.3
1	A	273	ILE	2.3
1	D	402	ASP	2.2
1	C	424	LEU	2.2
1	D	429	ASP	2.1
1	D	128	ARG	2.1
1	A	128	ARG	2.1
1	B	445	PRO	2.1
1	A	316	LEU	2.1
1	A	409	ARG	2.0
1	C	83	GLU	2.0
1	B	100	ASP	2.0

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

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

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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.

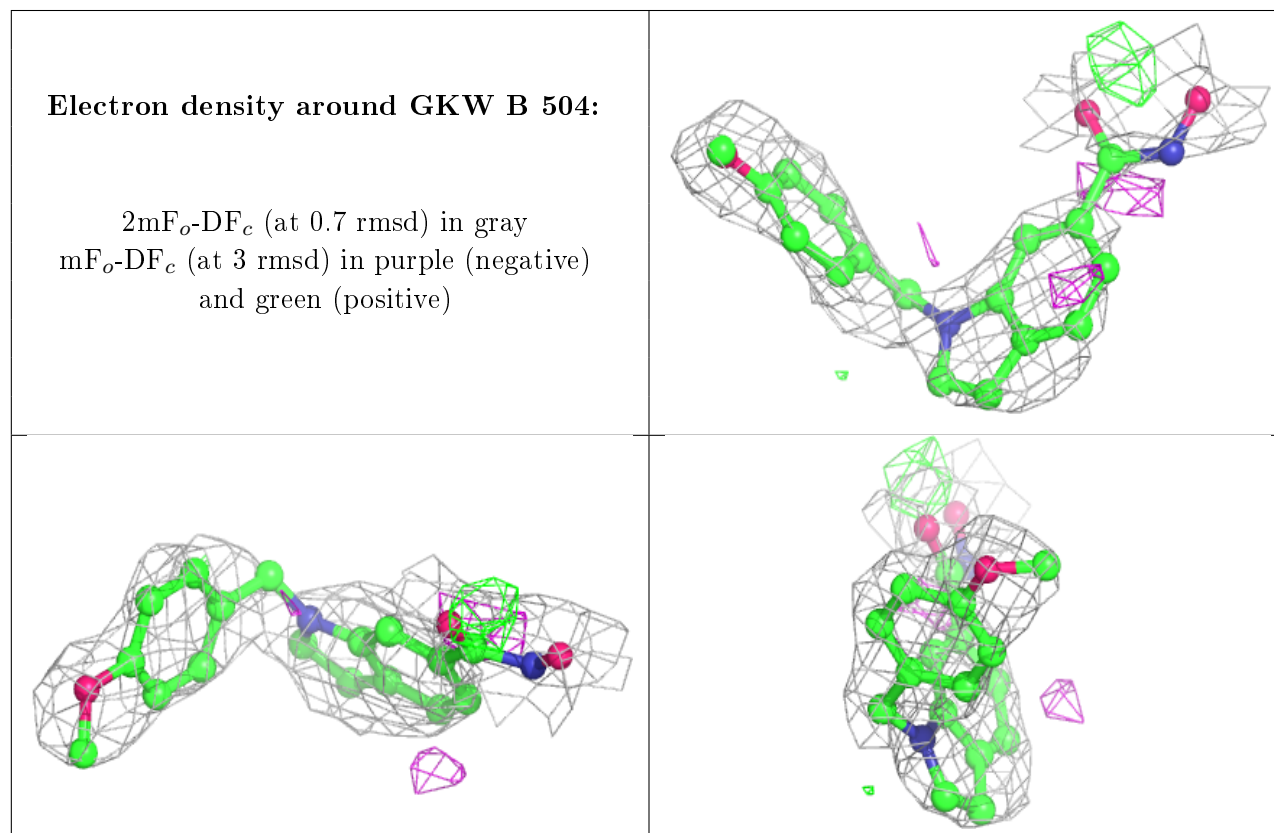
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	505	6/6	0.73	0.29	41,45,52,55	0
4	GKW	B	504	22/22	0.86	0.29	30,48,57,59	0
4	GKW	C	504	22/22	0.90	0.23	42,55,66,69	0
4	GKW	A	504	22/22	0.91	0.28	51,71,79,81	0
5	GOL	A	506	6/6	0.92	0.28	13,17,29,31	0
5	GOL	D	505	6/6	0.93	0.28	8,30,35,40	0

*Continued on next page...*

Continued from previous page...

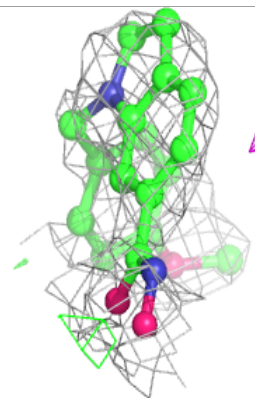
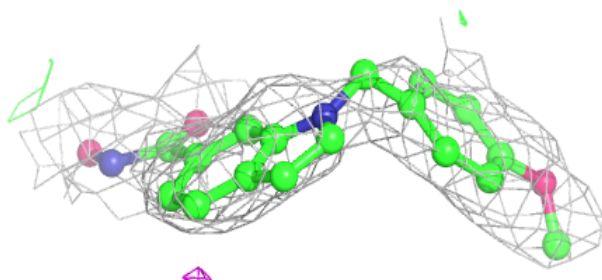
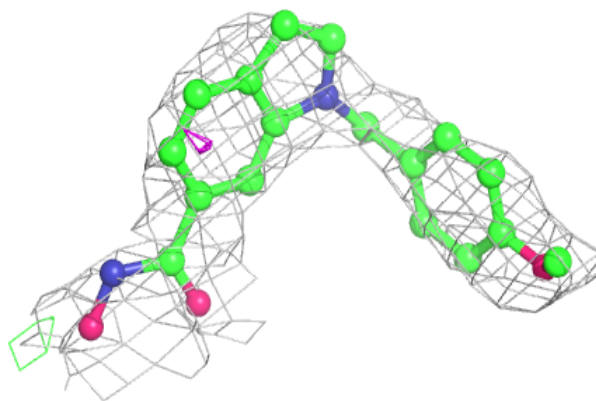
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	GOL	A	507	6/6	0.94	0.22	30,38,42,45	0
4	GKW	D	504	22/22	0.94	0.27	38,74,78,80	0
3	K	C	502	1/1	0.95	0.10	21,21,21,21	0
5	GOL	C	505	6/6	0.95	0.21	34,36,46,48	0
5	GOL	B	505	6/6	0.96	0.21	40,46,49,50	0
3	K	B	503	1/1	0.98	0.07	22,22,22,22	0
3	K	A	502	1/1	0.98	0.07	26,26,26,26	0
3	K	B	502	1/1	0.98	0.12	19,19,19,19	0
3	K	D	502	1/1	0.99	0.10	19,19,19,19	0
3	K	D	503	1/1	0.99	0.08	25,25,25,25	0
2	ZN	B	501	1/1	0.99	0.06	21,21,21,21	0
3	K	C	503	1/1	0.99	0.10	16,16,16,16	0
3	K	A	503	1/1	0.99	0.05	25,25,25,25	0
2	ZN	D	501	1/1	1.00	0.07	24,24,24,24	0
2	ZN	A	501	1/1	1.00	0.07	24,24,24,24	0
2	ZN	C	501	1/1	1.00	0.09	24,24,24,24	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.

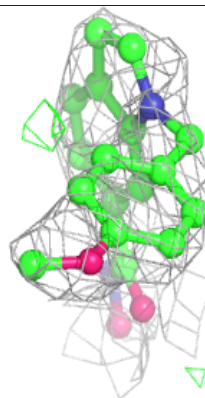
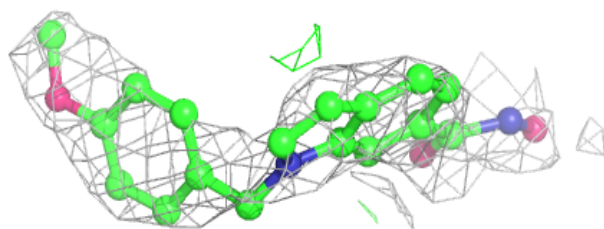
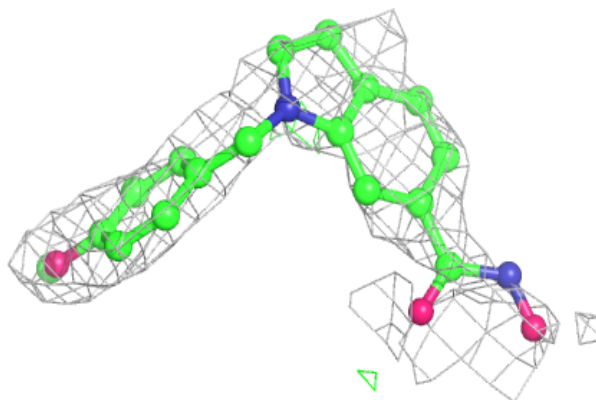


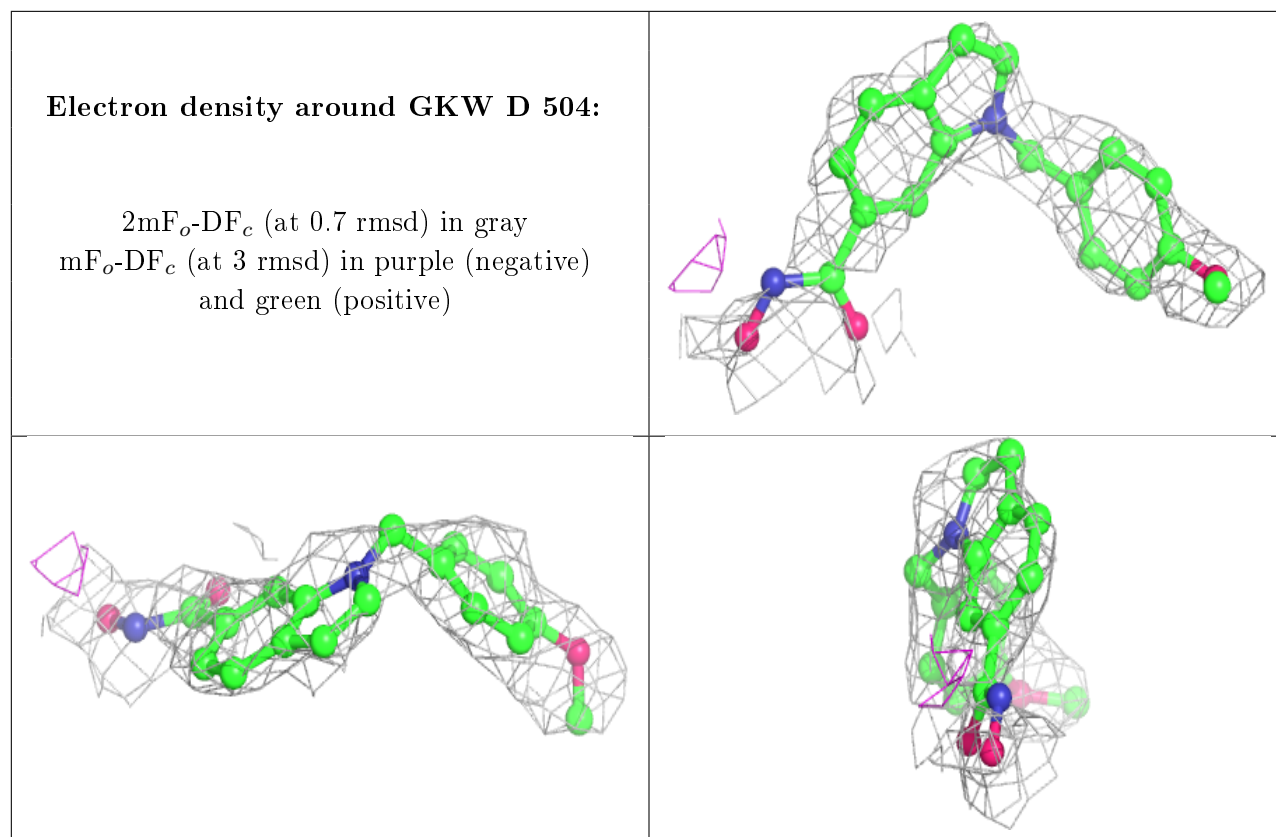
**Electron density around GKW C 504:**

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

**Electron density around GKW 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)





## 6.5 Other polymers [i](#)

There are no such residues in this entry.