



# Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 09:25 am BST

PDB ID : 6DUH  
Title : Crystal structure of HIV-1 reverse transcriptase Y181I mutant in complex with non-nucleoside inhibitor 25a  
Authors : Yang, Y.; Nguyen, L.A.; Smithline, Z.B.; Steitz, T.A.  
Deposited on : 2018-06-20  
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

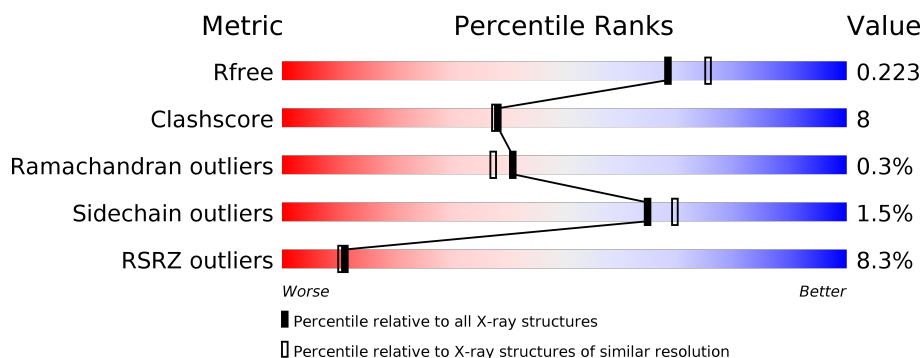
# 1 Overall quality at a glance

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	557	<div> <div>7%</div> <div>83%</div> <div>16%</div> </div>
2	B	428	<div> <div>9%</div> <div>82%</div> <div>13%</div> <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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	EDO	A	609	-	-	-	X
6	EDO	A	616	-	-	-	X
6	EDO	B	514	-	-	X	X
6	EDO	B	516	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 8723 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 p66 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	556	Total	C	N	O	S	0	0	0
			4512	2920	750	834	8			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	initiating methionine	UNP P03366
A	0	VAL	-	expression tag	UNP P03366
A	172	ALA	LYS	engineered mutation	UNP P03366
A	173	ALA	LYS	engineered mutation	UNP P03366
A	181	ILE	TYR	engineered mutation	UNP P03366
A	280	SER	CYS	engineered mutation	UNP P03366

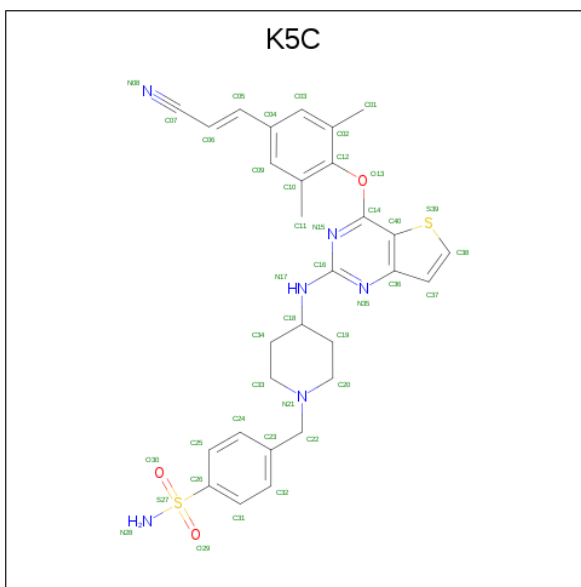
- Molecule 2 is a protein called p51 RT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	412	Total	C	N	O	S	0	0	0
			3412	2224	564	617	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 3 is 4-({4-[(4-{4-[(E)-2-cyanoethenyl]-2,6-dimethylphenoxy}thieno[3,2-d]pyrimidin-2-yl)amino]piperidin-1-yl)methyl)benzene-1-sulfonamide (three-letter code: K5C) (formula: C<sub>29</sub>H<sub>30</sub>N<sub>6</sub>O<sub>3</sub>S<sub>2</sub>) (labeled as "Ligand of Interest" by author).

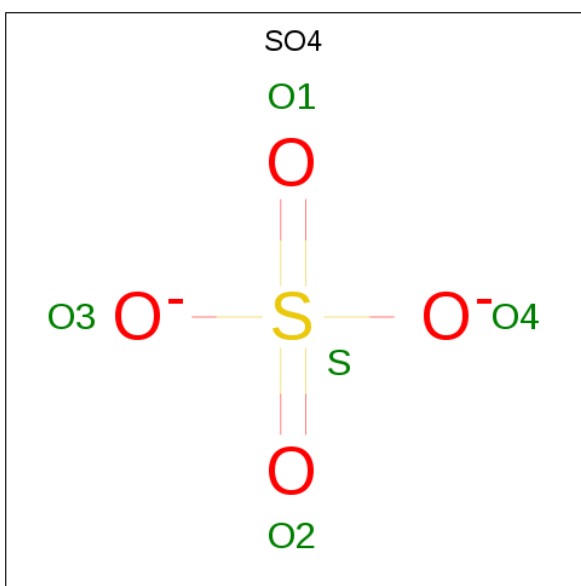


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			40	29	6	3	2		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

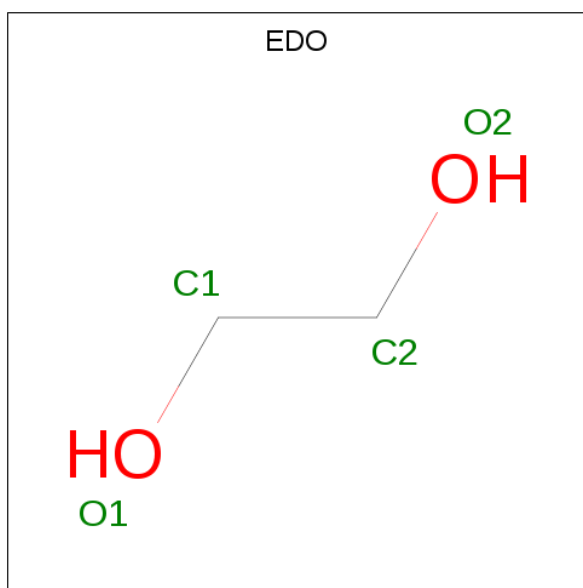
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mg	0	0
			1	1		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	A	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0
5	B	1	Total O S 5 4 1	0	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0
6	A	1	Total C O 4 2 2	0	0

*Continued on next page...*

*Continued from previous page...*

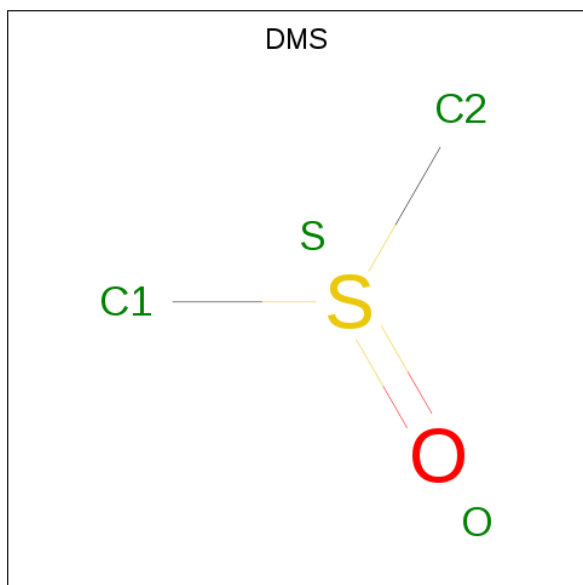
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	A	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0
6	B	1	Total 4	C 2	O 2	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		
6	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C<sub>2</sub>H<sub>6</sub>OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	B	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 8 is water.

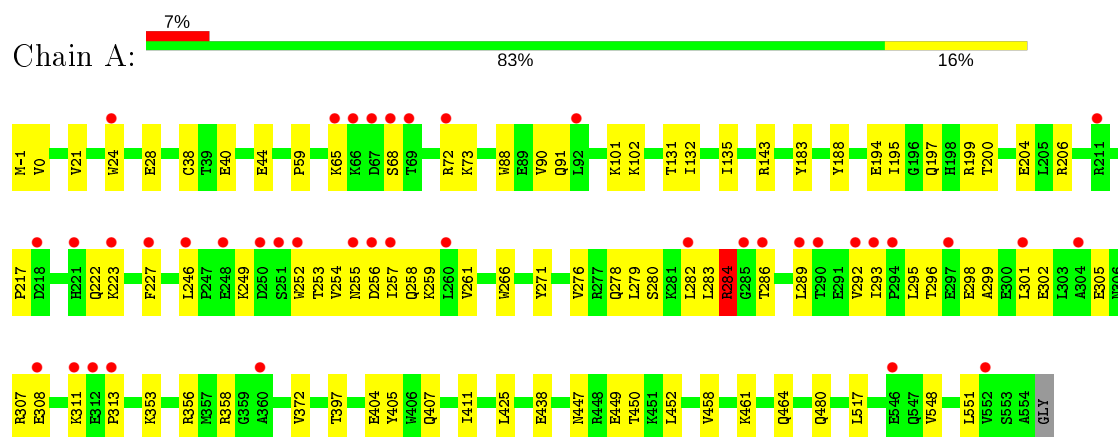
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	367	Total	O	0	0
			367	367		
8	B	235	Total	O	0	0
			235	235		



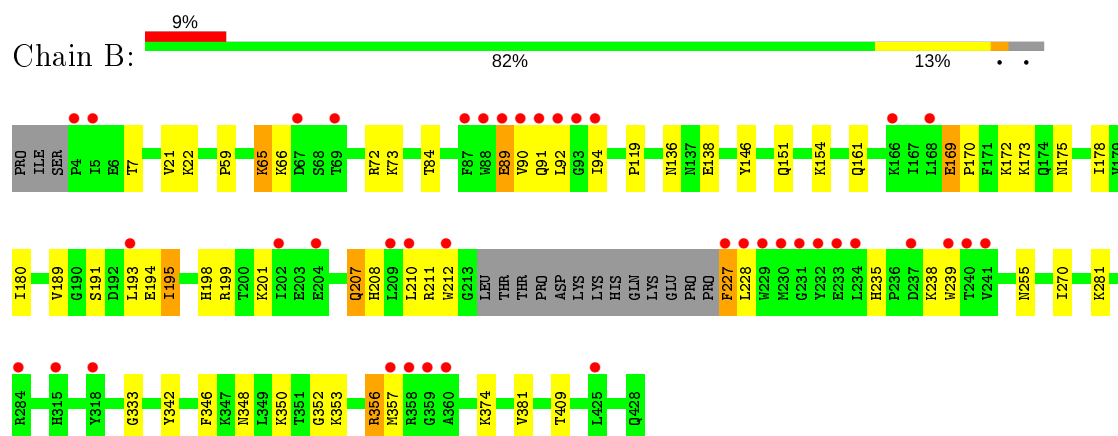
### 3 Residue-property plots [i](#)

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: p66 RT



#### • Molecule 2: p51 RT



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	162.90Å 73.23Å 109.38Å 90.00° 100.30° 90.00°	Depositor
Resolution (Å)	107.62 – 2.00 107.62 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.0 (107.62-2.00) 99.1 (107.62-2.00)	Depositor EDS
$R_{merge}$	0.03	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.99 (at 2.00Å)	Xtriage
Refinement program	PHENIX (1.13_2998)	Depositor
R, $R_{free}$	0.192 , 0.223 0.192 , 0.223	Depositor DCC
$R_{free}$ test set	4233 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.4	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 60.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8723	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.21% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, DMS, K5C, EDO, SO4

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	0.28	0/4629	0.48	0/6292
2	B	0.31	0/3511	0.50	0/4768
All	All	0.29	0/8140	0.49	0/11060

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

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	4512	0	4574	68	0
2	B	3412	0	3440	55	0
3	A	40	0	0	1	0
4	A	1	0	0	0	0
5	A	25	0	0	0	0
5	B	15	0	0	0	0
6	A	56	0	84	5	0
6	B	56	0	84	14	0
7	B	4	0	6	0	0
8	A	367	0	0	3	0
8	B	235	0	0	4	0
All	All	8723	0	8188	124	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 8.

All (124) 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:A:101:LYS:HE2	8:A:865:HOH:O	1.43	1.18
1:A:255:ASN:CG	1:A:289:LEU:HD23	1.75	1.05
1:A:254:VAL:HG12	1:A:293:ILE:CD1	1.94	0.97
2:B:191:SER:HG	2:B:198:HIS:HD1	1.17	0.91
1:A:255:ASN:CB	1:A:289:LEU:HD23	2.01	0.91
2:B:94:ILE:HD11	2:B:161:GLN:HE22	1.44	0.82
1:A:254:VAL:HG22	1:A:289:LEU:HA	1.62	0.81
2:B:333:GLY:HA2	6:B:518:EDO:H11	1.64	0.79
1:A:254:VAL:HG12	1:A:293:ILE:HD11	1.64	0.79
1:A:255:ASN:HB2	1:A:289:LEU:HD23	1.63	0.78
1:A:195:ILE:O	1:A:199:ARG:HG3	1.87	0.74
1:A:255:ASN:CG	1:A:289:LEU:CD2	2.54	0.73
1:A:548:VAL:HA	1:A:551:LEU:HD12	1.73	0.70
1:A:450:THR:HG23	1:A:452:LEU:H	1.58	0.68
1:A:254:VAL:HG12	1:A:293:ILE:HD12	1.74	0.68
2:B:65:LYS:NZ	8:B:601:HOH:O	2.28	0.66
1:A:254:VAL:CG2	1:A:289:LEU:HA	2.27	0.64
1:A:194:GLU:OE2	1:A:197:GLN:N	2.26	0.64
1:A:246:LEU:HD12	1:A:307:ARG:HG2	1.80	0.64
1:A:194:GLU:OE2	1:A:197:GLN:HG3	1.99	0.63
2:B:255:ASN:HB3	6:B:516:EDO:H21	1.81	0.62
2:B:346:PHE:H	6:B:514:EDO:H12	1.64	0.62
2:B:73:LYS:NZ	2:B:146:TYR:OH	2.33	0.62
1:A:183:TYR:HB3	1:A:188:TYR:HE2	1.64	0.62
2:B:89:GLU:HG2	2:B:90:VAL:H	1.64	0.62
2:B:175:ASN:HB3	2:B:178:ILE:HG12	1.83	0.60
1:A:296:THR:HG23	1:A:299:ALA:H	1.66	0.59
1:A:372:VAL:HG11	1:A:411:ILE:HG23	1.86	0.58
2:B:180:ILE:HG23	2:B:189:VAL:HG22	1.85	0.57
1:A:227:PHE:HB3	3:A:601:K5C:N08	2.20	0.55
1:A:254:VAL:CG1	1:A:293:ILE:HD11	2.35	0.55
2:B:352:GLY:HA2	6:B:509:EDO:H22	1.88	0.54
1:A:480:GLN:HG2	1:A:517:LEU:HD21	1.89	0.54
2:B:7:THR:HG22	2:B:119:PRO:HG2	1.88	0.54
2:B:207:GLN:OE1	2:B:207:GLN:HA	2.08	0.53
2:B:239:TRP:HB3	2:B:350:LYS:HE2	1.89	0.53
1:A:252:TRP:CD1	1:A:295:LEU:HD11	2.44	0.52

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:356:ARG:HH11	2:B:356:ARG:HG2	1.74	0.52
2:B:169:GLU:O	2:B:173:LYS:HG2	2.09	0.52
1:A:284:ARG:NE	1:A:284:ARG:O	2.43	0.51
1:A:353:LYS:HG2	6:A:612:EDO:H12	1.92	0.51
2:B:255:ASN:ND2	8:B:617:HOH:O	2.42	0.51
2:B:374:LYS:HG3	6:B:510:EDO:H21	1.94	0.50
2:B:208:HIS:O	2:B:212:TRP:HD1	1.95	0.50
2:B:238:LYS:HB2	2:B:239:TRP:HD1	1.77	0.50
1:A:283:LEU:O	1:A:286:THR:HG22	2.12	0.49
2:B:21:VAL:HG13	6:B:517:EDO:H22	1.94	0.49
1:A:131:THR:HG23	1:A:143:ARG:HG3	1.95	0.49
2:B:210:LEU:HA	2:B:227:PHE:HE1	1.77	0.49
1:A:284:ARG:C	1:A:284:ARG:HE	2.16	0.49
2:B:374:LYS:CG	6:B:510:EDO:H21	2.43	0.49
1:A:266:TRP:CD1	6:A:619:EDO:H11	2.48	0.49
2:B:208:HIS:O	2:B:212:TRP:CD1	2.66	0.49
1:A:358:ARG:NH2	8:A:726:HOH:O	2.45	0.49
2:B:84:THR:HB	2:B:154:LYS:HE2	1.95	0.48
2:B:169:GLU:HA	2:B:172:LYS:HG3	1.95	0.48
1:A:258:GLN:O	1:A:261:VAL:HG12	2.14	0.48
1:A:405:TYR:CE2	1:A:407:GLN:HB2	2.49	0.48
2:B:195:ILE:HD11	2:B:199:ARG:NH2	2.28	0.48
1:A:253:THR:HG22	1:A:292:VAL:HG22	1.94	0.48
1:A:254:VAL:O	8:A:701:HOH:O	2.20	0.48
1:A:90:VAL:N	1:A:91:GLN:OE1	2.46	0.48
2:B:136:ASN:HB3	2:B:138:GLU:HG3	1.96	0.48
6:B:514:EDO:H21	8:B:720:HOH:O	2.13	0.48
1:A:284:ARG:NE	1:A:284:ARG:C	2.67	0.47
1:A:73:LYS:HE3	1:A:73:LYS:HB3	1.66	0.47
1:A:257:ILE:HD13	1:A:282:LEU:HD23	1.96	0.47
2:B:22:LYS:HB3	6:B:517:EDO:H11	1.97	0.47
1:A:458:VAL:HG23	1:A:464:GLN:HG2	1.96	0.47
2:B:193:LEU:HD12	2:B:194:GLU:O	2.15	0.46
2:B:91:GLN:HG3	2:B:92:LEU:N	2.31	0.46
1:A:296:THR:OG1	1:A:298:GLU:OE2	2.30	0.46
2:B:193:LEU:HD11	2:B:198:HIS:N	2.31	0.46
2:B:353:LYS:O	6:B:509:EDO:H11	2.16	0.46
1:A:301:LEU:O	1:A:305:GLU:HG3	2.16	0.45
1:A:91:GLN:N	1:A:91:GLN:OE1	2.49	0.45
2:B:356:ARG:HD3	2:B:356:ARG:O	2.15	0.45
2:B:348:ASN:OD1	6:B:514:EDO:O2	2.23	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:72:ARG:NH1	2:B:151:GLN:OE1	2.50	0.45
1:A:88:TRP:CH2	6:A:615:EDO:H21	2.52	0.45
2:B:356:ARG:HG2	2:B:356:ARG:NH1	2.32	0.45
1:A:24:TRP:CE3	1:A:24:TRP:HA	2.52	0.44
2:B:170:PRO:HG2	2:B:212:TRP:CZ2	2.53	0.44
2:B:89:GLU:HG2	2:B:90:VAL:N	2.30	0.44
2:B:270:ILE:HG12	2:B:346:PHE:HB3	1.98	0.44
2:B:89:GLU:OE1	2:B:89:GLU:N	2.51	0.44
1:A:276:VAL:O	1:A:280:SER:OG	2.26	0.44
2:B:239:TRP:HZ3	2:B:381:VAL:HB	1.82	0.44
1:A:249:LYS:HB2	1:A:252:TRP:CE2	2.53	0.43
1:A:447:ASN:OD1	1:A:449:GLU:HG2	2.18	0.43
2:B:169:GLU:OE2	2:B:169:GLU:C	2.56	0.43
2:B:21:VAL:HB	2:B:59:PRO:HD3	1.98	0.43
1:A:258:GLN:HA	1:A:261:VAL:HG12	2.01	0.43
1:A:255:ASN:ND2	1:A:289:LEU:HD23	2.31	0.43
2:B:72:ARG:NH2	8:B:619:HOH:O	2.46	0.43
1:A:200:THR:O	1:A:204:GLU:HG3	2.17	0.43
1:A:279:LEU:HG	1:A:302:GLU:OE1	2.18	0.43
1:A:206:ARG:HH21	1:A:217:PRO:HA	1.84	0.43
1:A:28:GLU:HG3	1:A:135:ILE:HD12	2.01	0.43
1:A:438:GLU:HG3	1:A:461:LYS:HG3	2.01	0.43
1:A:38:CYS:SG	1:A:132:ILE:HD11	2.59	0.42
1:A:72:ARG:HD3	1:A:72:ARG:HA	1.88	0.42
2:B:346:PHE:H	6:B:514:EDO:C1	2.30	0.42
1:A:65:LYS:HB2	1:A:68:SER:HB2	2.00	0.42
1:A:40:GLU:O	1:A:44:GLU:HG3	2.20	0.42
2:B:201:LYS:HD3	2:B:201:LYS:HA	1.83	0.42
2:B:227:PHE:HB2	2:B:228:LEU:H	1.44	0.42
2:B:235:HIS:H	6:B:512:EDO:H22	1.83	0.42
1:A:404:GLU:O	6:A:616:EDO:H22	2.19	0.42
1:A:296:THR:HG23	1:A:299:ALA:N	2.32	0.41
1:A:271:TYR:OH	1:A:313:PRO:HA	2.20	0.41
2:B:357:MET:SD	2:B:374:LYS:NZ	2.91	0.41
2:B:65:LYS:HG3	2:B:66:LYS:H	1.85	0.41
1:A:353:LYS:H	6:A:612:EDO:H12	1.85	0.41
2:B:342:TYR:CE1	6:B:514:EDO:H11	2.55	0.41
1:A:308:GLU:O	1:A:311:LYS:HG2	2.21	0.41
1:A:-1:MET:HE2	1:A:0:VAL:O	2.21	0.41
2:B:170:PRO:HG2	2:B:212:TRP:CH2	2.56	0.40
2:B:211:ARG:CZ	2:B:211:ARG:HB2	2.51	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:253:THR:N	1:A:256:ASP:OD2	2.34	0.40
1:A:397:THR:HG23	1:A:425:LEU:HG	2.03	0.40
1:A:255:ASN:ND2	1:A:289:LEU:CD2	2.84	0.40
1:A:21:VAL:HB	1:A:59:PRO:HD3	2.04	0.40
2:B:72:ARG:NH2	2:B:409:THR:HG22	2.37	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	554/557 (100%)	538 (97%)	15 (3%)	1 (0%)	47	44
2	B	408/428 (95%)	394 (97%)	12 (3%)	2 (0%)	29	23
All	All	962/985 (98%)	932 (97%)	27 (3%)	3 (0%)	41	37

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	89	GLU
1	A	284	ARG
2	B	195	ILE

### 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	495/495 (100%)	488 (99%)	7 (1%)	67	72
2	B	374/390 (96%)	368 (98%)	6 (2%)	62	67
All	All	869/885 (98%)	856 (98%)	13 (2%)	65	69

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	LYS
1	A	222	GLN
1	A	223	LYS
1	A	259	LYS
1	A	278	GLN
1	A	284	ARG
1	A	356	ARG
2	B	65	LYS
2	B	169	GLU
2	B	207	GLN
2	B	227	PHE
2	B	281	LYS
2	B	356	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	182	GLN
1	A	197	GLN
1	A	258	GLN
1	A	278	GLN
1	A	306	ASN
2	B	161	GLN
2	B	394	GLN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 39 ligands modelled in this entry, 1 is monoatomic - leaving 38 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
6	EDO	B	510	-	3,3,3	0.45	0	2,2,2	0.32	0
6	EDO	A	613	-	3,3,3	0.44	0	2,2,2	0.41	0
6	EDO	A	620	-	3,3,3	0.46	0	2,2,2	0.27	0
6	EDO	B	508	-	3,3,3	0.43	0	2,2,2	0.37	0
5	SO4	A	606	-	4,4,4	0.13	0	6,6,6	0.08	0
6	EDO	B	516	-	3,3,3	0.44	0	2,2,2	0.33	0
6	EDO	B	509	-	3,3,3	0.47	0	2,2,2	0.25	0
5	SO4	A	603	-	4,4,4	0.15	0	6,6,6	0.08	0
6	EDO	A	621	-	3,3,3	0.47	0	2,2,2	0.27	0
6	EDO	B	506	-	3,3,3	0.43	0	2,2,2	0.33	0
6	EDO	A	617	-	3,3,3	0.47	0	2,2,2	0.22	0
6	EDO	B	505	-	3,3,3	0.50	0	2,2,2	0.15	0
6	EDO	A	612	-	3,3,3	0.48	0	2,2,2	0.21	0
7	DMS	B	501	-	3,3,3	0.59	0	3,3,3	0.56	0
6	EDO	A	616	-	3,3,3	0.46	0	2,2,2	0.23	0
6	EDO	B	507	-	3,3,3	0.47	0	2,2,2	0.30	0
6	EDO	A	619	-	3,3,3	0.46	0	2,2,2	0.22	0
6	EDO	B	514	-	3,3,3	0.42	0	2,2,2	0.33	0
5	SO4	A	605	-	4,4,4	0.14	0	6,6,6	0.06	0
6	EDO	B	515	-	3,3,3	0.45	0	2,2,2	0.32	0
6	EDO	A	610	-	3,3,3	0.46	0	2,2,2	0.33	0
6	EDO	A	611	-	3,3,3	0.43	0	2,2,2	0.38	0
6	EDO	A	614	-	3,3,3	0.46	0	2,2,2	0.33	0
6	EDO	A	615	-	3,3,3	0.46	0	2,2,2	0.31	0
5	SO4	B	504	-	4,4,4	0.15	0	6,6,6	0.04	0
3	K5C	A	601	-	42,44,44	1.62	3 (7%)	51,63,63	2.28	20 (39%)
5	SO4	A	607	-	4,4,4	0.15	0	6,6,6	0.04	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	EDO	B	513	-	3,3,3	0.46	0	2,2,2	0.34	0
6	EDO	B	517	-	3,3,3	0.45	0	2,2,2	0.35	0
5	SO4	B	502	-	4,4,4	0.14	0	6,6,6	0.06	0
6	EDO	A	618	-	3,3,3	0.46	0	2,2,2	0.32	0
6	EDO	B	511	-	3,3,3	0.47	0	2,2,2	0.30	0
6	EDO	A	608	-	3,3,3	0.46	0	2,2,2	0.35	0
6	EDO	B	518	-	3,3,3	0.44	0	2,2,2	0.22	0
6	EDO	B	512	-	3,3,3	0.50	0	2,2,2	0.27	0
5	SO4	A	604	-	4,4,4	0.14	0	6,6,6	0.08	0
6	EDO	A	609	-	3,3,3	0.46	0	2,2,2	0.25	0
5	SO4	B	503	-	4,4,4	0.17	0	6,6,6	0.09	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
6	EDO	B	510	-	-	0/1/1/1	-
6	EDO	A	613	-	-	0/1/1/1	-
6	EDO	A	620	-	-	0/1/1/1	-
6	EDO	B	508	-	-	0/1/1/1	-
6	EDO	B	516	-	-	0/1/1/1	-
6	EDO	A	619	-	-	0/1/1/1	-
6	EDO	B	509	-	-	1/1/1/1	-
6	EDO	A	621	-	-	1/1/1/1	-
6	EDO	B	506	-	-	0/1/1/1	-
6	EDO	A	617	-	-	0/1/1/1	-
6	EDO	B	505	-	-	0/1/1/1	-
6	EDO	A	612	-	-	0/1/1/1	-
6	EDO	A	616	-	-	0/1/1/1	-
6	EDO	B	507	-	-	0/1/1/1	-
3	K5C	A	601	-	-	3/21/32/32	0/5/5/5
6	EDO	B	514	-	-	0/1/1/1	-
6	EDO	B	515	-	-	0/1/1/1	-
6	EDO	A	610	-	-	0/1/1/1	-
6	EDO	A	611	-	-	0/1/1/1	-
6	EDO	A	614	-	-	0/1/1/1	-
6	EDO	A	615	-	-	0/1/1/1	-
6	EDO	B	513	-	-	0/1/1/1	-
6	EDO	B	517	-	-	0/1/1/1	-
6	EDO	A	618	-	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	B	511	-	-	0/1/1/1	-
6	EDO	A	608	-	-	0/1/1/1	-
6	EDO	B	518	-	-	0/1/1/1	-
6	EDO	B	512	-	-	0/1/1/1	-
6	EDO	A	609	-	-	0/1/1/1	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	601	K5C	C16-N17	7.64	1.44	1.34
3	A	601	K5C	S27-N28	2.74	1.65	1.60
3	A	601	K5C	C14-N15	2.08	1.35	1.31

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	601	K5C	C31-C26-C25	-6.02	112.04	120.44
3	A	601	K5C	C32-C31-C26	5.75	125.40	119.45
3	A	601	K5C	C22-N21-C20	-5.45	98.93	111.06
3	A	601	K5C	C24-C25-C26	4.75	124.36	119.45
3	A	601	K5C	C34-C18-N17	-3.76	104.74	110.60
3	A	601	K5C	C31-C26-S27	3.60	124.95	119.73
3	A	601	K5C	O29-S27-C26	3.55	111.31	107.35
3	A	601	K5C	O30-S27-O29	-3.21	113.48	118.76
3	A	601	K5C	O13-C14-C40	3.03	120.62	115.69
3	A	601	K5C	C16-N17-C18	-2.68	119.78	124.31
3	A	601	K5C	N17-C16-N35	-2.66	113.19	117.18
3	A	601	K5C	C01-C02-C03	2.60	124.30	119.49
3	A	601	K5C	C09-C04-C03	-2.44	115.81	119.12
3	A	601	K5C	C22-C23-C32	2.35	125.19	120.77
3	A	601	K5C	C37-C36-C40	2.31	112.31	110.62
3	A	601	K5C	C25-C26-S27	2.26	123.01	119.73
3	A	601	K5C	N17-C16-N15	2.13	120.38	117.18
3	A	601	K5C	C34-C18-C19	2.10	114.46	110.82
3	A	601	K5C	C19-C18-N17	2.06	113.81	110.60
3	A	601	K5C	O13-C12-C02	2.06	122.84	118.35

There are no chirality outliers.

All (5) torsion outliers are listed below:

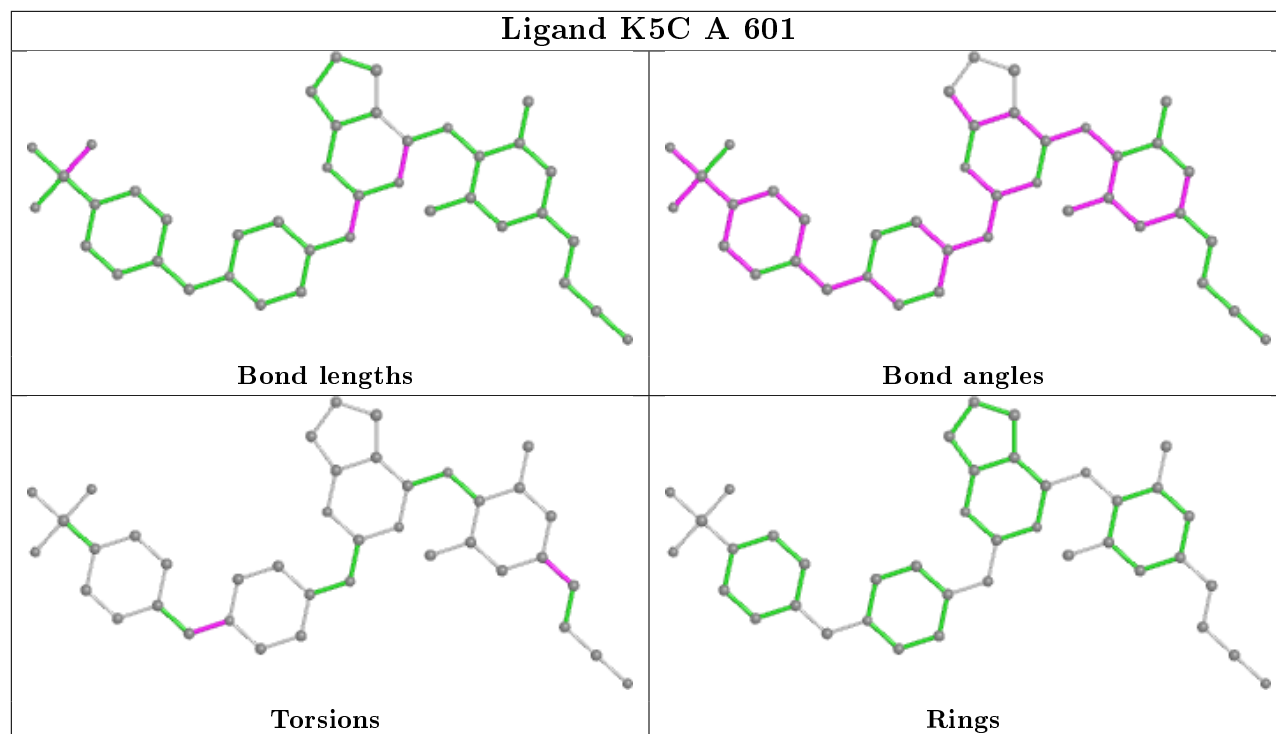
Mol	Chain	Res	Type	Atoms
3	A	601	K5C	C03-C04-C05-C06
3	A	601	K5C	C09-C04-C05-C06
6	B	509	EDO	O1-C1-C2-O2
6	A	621	EDO	O1-C1-C2-O2
3	A	601	K5C	C23-C22-N21-C20

There are no ring outliers.

12 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	510	EDO	2	0
6	B	516	EDO	1	0
6	B	509	EDO	2	0
6	A	612	EDO	2	0
6	A	616	EDO	1	0
6	A	619	EDO	1	0
6	B	514	EDO	5	0
6	A	615	EDO	1	0
3	A	601	K5C	1	0
6	B	517	EDO	2	0
6	B	518	EDO	1	0
6	B	512	EDO	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.



## 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	556/557 (99%)	0.50	40 (7%) 15 14	38, 67, 151, 218	0
2	B	412/428 (96%)	0.63	40 (9%) 7 7	38, 64, 141, 218	0
All	All	968/985 (98%)	0.56	80 (8%) 11 10	38, 66, 143, 218	0

All (80) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	90	VAL	12.0
1	A	67	ASP	10.4
1	A	290	THR	9.9
2	B	91	GLN	9.1
2	B	4	PRO	8.8
1	A	286	THR	7.6
2	B	358	ARG	7.4
2	B	92	LEU	7.1
2	B	88	TRP	6.9
2	B	239	TRP	6.9
2	B	89	GLU	6.1
1	A	257	ILE	6.1
2	B	359	GLY	5.9
2	B	232	TYR	5.8
1	A	66	LYS	5.7
2	B	315	HIS	5.3
2	B	228	LEU	5.3
1	A	223	LYS	5.0
1	A	285	GLY	4.9
1	A	68	SER	4.8
1	A	294	PRO	4.6
1	A	221	HIS	4.5
2	B	230	MET	4.2
1	A	69	THR	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	425	LEU	4.1
2	B	229	TRP	4.1
1	A	311	LYS	4.1
1	A	92	LEU	3.9
2	B	231	GLY	3.6
1	A	24	TRP	3.5
2	B	357	MET	3.5
1	A	292	VAL	3.4
1	A	256	ASP	3.4
2	B	237	ASP	3.3
1	A	293	ILE	3.3
1	A	252	TRP	3.2
1	A	297	GLU	3.2
1	A	301	LEU	3.2
1	A	250	ASP	3.2
1	A	289	LEU	3.2
2	B	87	PHE	3.0
2	B	318	TYR	3.0
1	A	260	LEU	2.9
2	B	168	LEU	2.9
2	B	5	ILE	2.8
2	B	166	LYS	2.8
1	A	218	ASP	2.8
2	B	360	ALA	2.7
2	B	69	THR	2.7
2	B	94	ILE	2.7
1	A	308	GLU	2.7
1	A	304	ALA	2.7
2	B	209	LEU	2.6
1	A	546	GLU	2.5
1	A	227	PHE	2.5
2	B	204	GLU	2.4
2	B	202	ILE	2.4
2	B	234	LEU	2.4
1	A	312	GLU	2.4
1	A	251	SER	2.4
2	B	210	LEU	2.3
2	B	212	TRP	2.3
2	B	93	GLY	2.3
2	B	241	VAL	2.3
1	A	72	ARG	2.3
1	A	552	VAL	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	282	LEU	2.2
2	B	227	PHE	2.2
1	A	211	ARG	2.2
2	B	233	GLU	2.2
2	B	67	ASP	2.2
1	A	255	ASN	2.2
1	A	246	LEU	2.1
2	B	193	LEU	2.1
1	A	313	PRO	2.1
1	A	248	GLU	2.1
2	B	240	THR	2.1
2	B	284	ARG	2.1
1	A	65	LYS	2.1
1	A	360	ALA	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(Å <sup>2</sup> )	Q<0.9
6	EDO	A	612	4/4	0.29	0.20	105,105,107,107	0
6	EDO	B	511	4/4	0.29	0.31	105,107,109,110	0
6	EDO	A	621	4/4	0.33	0.35	104,107,110,112	0
6	EDO	A	609	4/4	0.56	0.48	110,112,112,112	0
6	EDO	A	616	4/4	0.57	0.63	111,117,122,124	0
6	EDO	B	510	4/4	0.58	0.22	92,92,93,93	0
6	EDO	B	509	4/4	0.61	0.27	112,114,114,116	0
4	MG	A	602	1/1	0.61	0.21	109,109,109,109	0
6	EDO	A	610	4/4	0.67	0.34	84,86,89,91	0

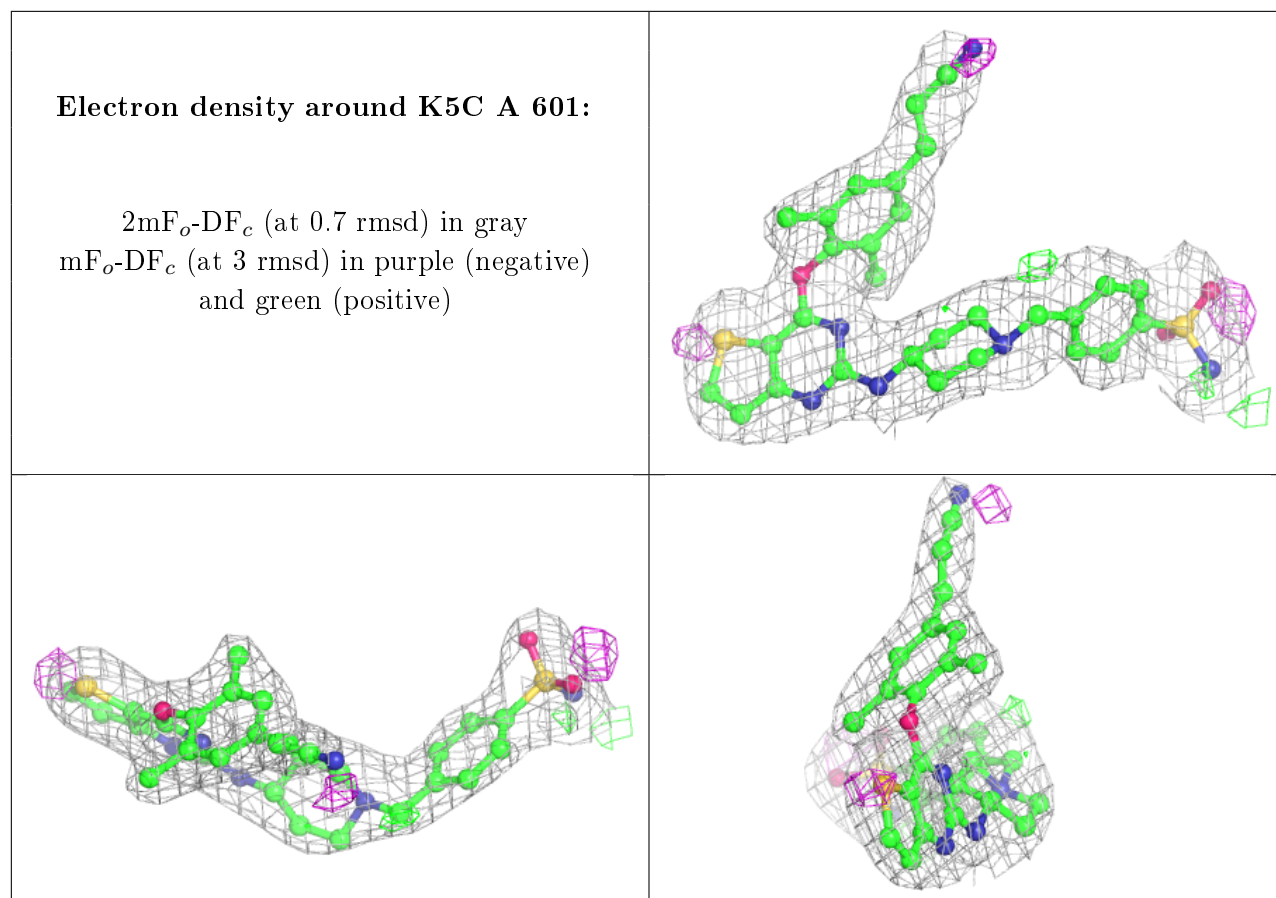
*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SO4	A	605	5/5	0.68	0.24	168,169,170,171	0
6	EDO	A	614	4/4	0.68	0.29	81,82,85,89	0
6	EDO	A	619	4/4	0.70	0.26	91,92,93,93	0
6	EDO	B	514	4/4	0.75	0.45	84,87,94,98	0
6	EDO	A	618	4/4	0.75	0.19	92,93,95,96	0
5	SO4	B	504	5/5	0.76	0.38	162,163,163,164	0
6	EDO	B	507	4/4	0.77	0.26	92,95,96,96	0
6	EDO	B	512	4/4	0.77	0.25	86,86,88,89	0
6	EDO	B	515	4/4	0.77	0.20	97,98,98,99	0
6	EDO	B	505	4/4	0.78	0.22	67,74,75,78	0
6	EDO	B	516	4/4	0.79	0.42	94,94,95,97	0
6	EDO	B	506	4/4	0.84	0.19	66,66,70,74	0
6	EDO	B	518	4/4	0.84	0.29	96,96,96,100	0
5	SO4	A	604	5/5	0.85	0.16	120,122,125,126	0
6	EDO	A	611	4/4	0.86	0.14	75,77,83,90	0
6	EDO	A	608	4/4	0.86	0.15	93,93,94,95	0
5	SO4	B	503	5/5	0.86	0.35	157,157,158,159	0
6	EDO	B	513	4/4	0.87	0.24	90,91,92,94	0
6	EDO	B	508	4/4	0.87	0.25	65,68,69,70	0
6	EDO	A	617	4/4	0.87	0.19	85,87,88,88	0
5	SO4	A	607	5/5	0.88	0.27	163,164,164,165	0
6	EDO	A	615	4/4	0.89	0.14	87,87,88,90	0
6	EDO	B	517	4/4	0.89	0.23	88,88,92,96	0
5	SO4	A	606	5/5	0.90	0.25	95,95,99,102	0
6	EDO	A	620	4/4	0.92	0.20	86,88,90,90	0
3	K5C	A	601	40/40	0.92	0.14	50,71,90,92	0
7	DMS	B	501	4/4	0.93	0.21	60,71,71,77	0
5	SO4	B	502	5/5	0.94	0.25	122,122,124,126	0
5	SO4	A	603	5/5	0.96	0.15	119,120,120,121	0
6	EDO	A	613	4/4	0.97	0.20	58,62,67,74	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.



## 6.5 Other polymers ⓘ

There are no such residues in this entry.