Full wwPDB X-ray Structure Validation Report ⓘ

Sep 25, 2023 – 02:45 AM EDT

PDB ID	:	5WFC
Title	:	Humanized mutant of the Chaetomium thermophilum Polycomb Repressive Complex 2 bound to the inhibitor GSK343
Authors	:	Bratkowski, M.A.; Liu, X.
Deposited on	:	2017-07-11
Resolution	:	2.28 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

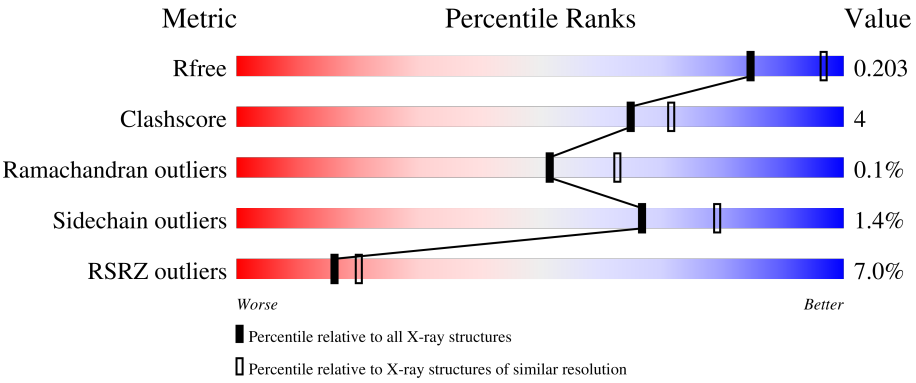
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.28 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)
RSRZ outliers	127900	6849 (2.30-2.26)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	605	5% 69% 7% 24%
2	B	936	5% 72% 10% 18%
3	D	11	36% 55% 9% 36%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 10563 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 Polycomb Protein EED.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	460	3643	2328	628	668	19	0	2	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-39	MET	-	initiating methionine	UNP G0S8H7
A	-38	ALA	-	expression tag	UNP G0S8H7
A	-37	SER	-	expression tag	UNP G0S8H7
A	-36	ALA	-	expression tag	UNP G0S8H7
A	-35	TRP	-	expression tag	UNP G0S8H7
A	-34	SER	-	expression tag	UNP G0S8H7
A	-33	HIS	-	expression tag	UNP G0S8H7
A	-32	PRO	-	expression tag	UNP G0S8H7
A	-31	GLN	-	expression tag	UNP G0S8H7
A	-30	PHE	-	expression tag	UNP G0S8H7
A	-29	GLU	-	expression tag	UNP G0S8H7
A	-28	LYS	-	expression tag	UNP G0S8H7
A	-27	GLY	-	expression tag	UNP G0S8H7
A	-26	GLY	-	expression tag	UNP G0S8H7
A	-25	GLY	-	expression tag	UNP G0S8H7
A	-24	SER	-	expression tag	UNP G0S8H7
A	-23	GLY	-	expression tag	UNP G0S8H7
A	-22	GLY	-	expression tag	UNP G0S8H7
A	-21	GLY	-	expression tag	UNP G0S8H7
A	-20	SER	-	expression tag	UNP G0S8H7
A	-19	GLY	-	expression tag	UNP G0S8H7
A	-18	GLY	-	expression tag	UNP G0S8H7
A	-17	SER	-	expression tag	UNP G0S8H7
A	-16	ALA	-	expression tag	UNP G0S8H7
A	-15	TRP	-	expression tag	UNP G0S8H7
A	-14	SER	-	expression tag	UNP G0S8H7
A	-13	HIS	-	expression tag	UNP G0S8H7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	-12	PRO	-	expression tag	UNP G0S8H7
A	-11	GLN	-	expression tag	UNP G0S8H7
A	-10	PHE	-	expression tag	UNP G0S8H7
A	-9	GLU	-	expression tag	UNP G0S8H7
A	-8	LYS	-	expression tag	UNP G0S8H7
A	-7	LEU	-	expression tag	UNP G0S8H7
A	-6	GLU	-	expression tag	UNP G0S8H7
A	-5	VAL	-	expression tag	UNP G0S8H7
A	-4	LEU	-	expression tag	UNP G0S8H7
A	-3	PHE	-	expression tag	UNP G0S8H7
A	-2	GLN	-	expression tag	UNP G0S8H7
A	-1	GLY	-	expression tag	UNP G0S8H7
A	0	PRO	-	expression tag	UNP G0S8H7

- Molecule 2 is a protein called Histone-lysine-N-methyltransferase EZH2, Polycomb protein SUZ12 chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	769	Total	C	N	O	S	0	0	0
			6185	3898	1116	1130	41			

There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	182	SER	-	expression tag	UNP G0SDW4
B	183	ASN	-	expression tag	UNP G0SDW4
B	184	HIS	-	expression tag	UNP G0SDW4
B	185	HIS	-	expression tag	UNP G0SDW4
B	186	HIS	-	expression tag	UNP G0SDW4
B	187	HIS	-	expression tag	UNP G0SDW4
B	188	HIS	-	expression tag	UNP G0SDW4
B	189	HIS	-	expression tag	UNP G0SDW4
B	190	ALA	-	expression tag	UNP G0SDW4
B	302	ILE	PRO	engineered mutation	UNP G0SDW4
B	304	TYR	ARG	engineered mutation	UNP G0SDW4
B	?	-	GLU	deletion	UNP G0SDW4
B	850	LYS	GLU	engineered mutation	UNP G0SDW4
B	851	TYR	ASN	engineered mutation	UNP G0SDW4
B	852	MET	LYS	engineered mutation	UNP G0SDW4
B	853	CYS	VAL	engineered mutation	UNP G0SDW4
B	855	PHE	TYR	engineered mutation	UNP G0SDW4
B	2524	LEU	-	linker	UNP G0SDW4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	2525	VAL	-	linker	UNP G0SDW4
B	2526	PRO	-	linker	UNP G0SDW4
B	2527	ARG	-	linker	UNP G0SDW4
B	2528	GLY	-	linker	UNP G0SDW4
B	2529	SER	-	linker	UNP G0SDW4

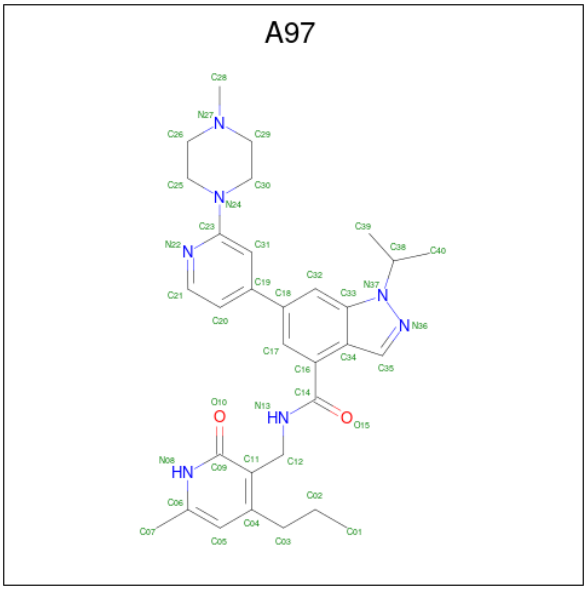
- Molecule 3 is a protein called Histone H3.1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	D	7	Total	C	N	O	0	0	0
			51	32	11	8			

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	8	Total	Zn	0	0
			8	8		

- Molecule 5 is N-[(6-methyl-2-oxo-4-propyl-1,2-dihydropyridin-3-yl)methyl]-6-[2-(4-methylpiperazin-1-yl)pyridin-4-yl]-1-(propan-2-yl)-1H-indazole-4-carboxamide (three-letter code: A97) (formula: $C_{31}H_{39}N_7O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			40	31	7	2		

- Molecule 6 is water.

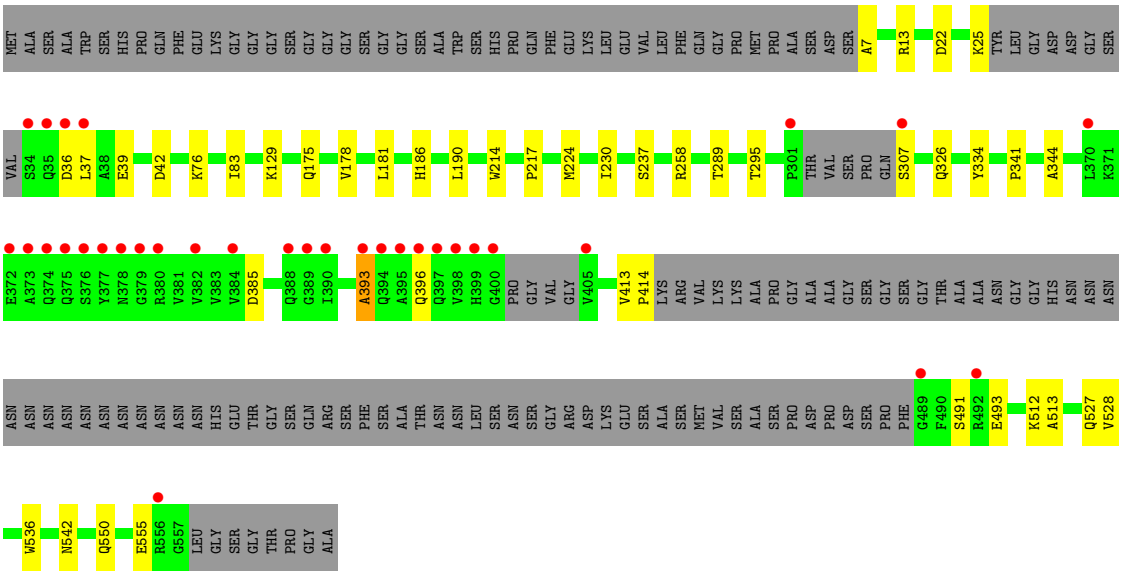
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	258	Total O 258 258	0	0
6	B	377	Total O 377 377	0	0
6	D	1	Total O 1 1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

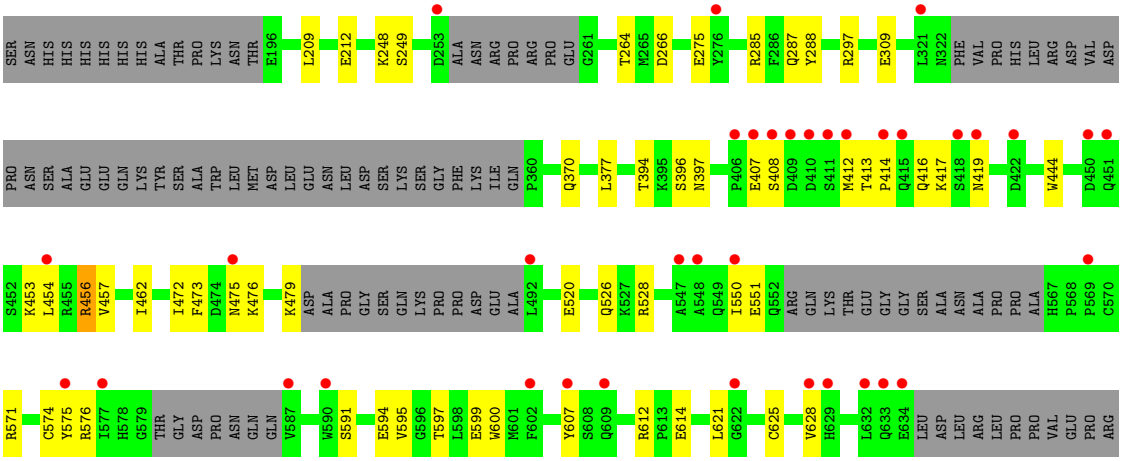
• Molecule 1: Polycomb Protein EED

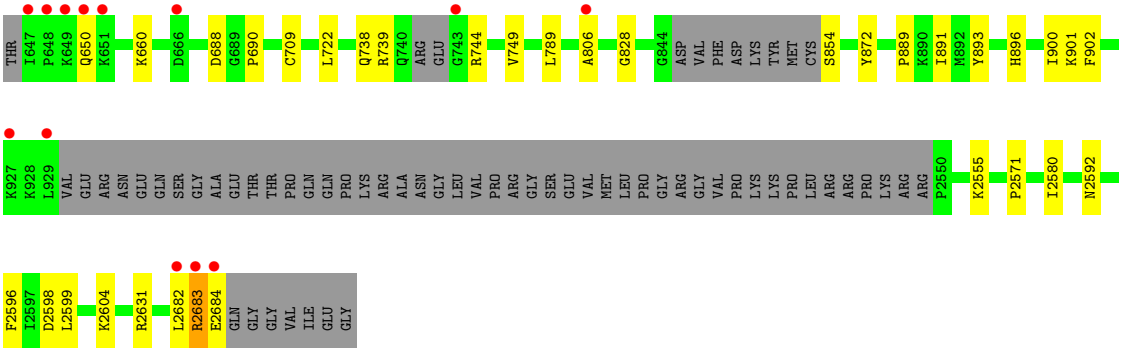
Chain A:



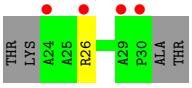
• Molecule 2: Histone-lysine-N-methyltransferase EZH2, Polycomb protein SUZ12 chimera

Chain B:





• Molecule 3: Histone H3.1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	117.59Å 137.79Å 222.33Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.72 – 2.28 44.72 – 2.28	Depositor EDS
% Data completeness (in resolution range)	96.5 (44.72-2.28) 96.5 (44.72-2.28)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.67 (at 2.29Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, R_{free}	0.167 , 0.203 0.167 , 0.203	Depositor DCC
R_{free} test set	4072 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	29.3	Xtriage
Anisotropy	0.023	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 49.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	10563	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.71% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: A97, ZN, M3L

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	0.42	0/3753	0.59	0/5106
2	B	0.40	0/6325	0.54	0/8541
3	D	0.65	0/39	0.43	0/52
All	All	0.41	0/10117	0.56	0/13699

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	3643	0	3510	26	0
2	B	6185	0	6076	57	0
3	D	51	0	58	1	0
4	B	8	0	0	0	0
5	B	40	0	0	0	0
6	A	258	0	0	3	0
6	B	377	0	0	4	0
6	D	1	0	0	0	0
All	All	10563	0	9644	81	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 4.

All (81) 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)
2:B:248:LYS:HD2	2:B:249:SER:H	1.30	0.91
2:B:2683:ARG:O	2:B:2684:GLU:C	2.21	0.79
2:B:248:LYS:HD2	2:B:249:SER:N	2.07	0.68
1:A:22:ASP:OD2	1:A:36:ASP:HA	1.94	0.67
2:B:412:MET:HE2	2:B:417:LYS:HG2	1.76	0.66
2:B:891:ILE:HD13	2:B:900:ILE:HG12	1.76	0.66
2:B:476:LYS:HE2	2:B:2596:PHE:CE1	2.32	0.65
2:B:248:LYS:CD	2:B:249:SER:H	2.09	0.64
1:A:326:GLN:OE1	3:D:26:ARG:NH1	2.30	0.62
2:B:309:GLU:HG2	2:B:2580:ILE:HG12	1.83	0.60
1:A:186:HIS:CG	1:A:190:LEU:HD21	2.39	0.58
2:B:248:LYS:HE2	2:B:249:SER:HB3	1.86	0.56
2:B:419:ASN:HB3	2:B:472:ILE:HG12	1.86	0.56
2:B:407:GLU:OE2	2:B:407:GLU:N	2.35	0.56
1:A:289:THR:HG22	6:A:717:HOH:O	2.05	0.55
2:B:413:THR:HG23	2:B:414:PRO:O	2.08	0.54
1:A:237:SER:HB2	1:A:295:THR:HG23	1.89	0.54
2:B:475:ASN:O	2:B:526:GLN:NE2	2.41	0.52
2:B:612:ARG:HG3	2:B:612:ARG:HH11	1.75	0.52
2:B:595:VAL:O	2:B:599:GLU:HG3	2.10	0.52
1:A:536:TRP:CZ3	1:A:550:GLN:HB2	2.45	0.51
1:A:491:SER:HB2	1:A:493:GLU:OE2	2.11	0.50
2:B:614:GLU:HB3	2:B:625:CYS:SG	2.52	0.49
2:B:453:LYS:HD3	2:B:454:LEU:N	2.28	0.49
2:B:520:GLU:HB3	2:B:528:ARG:HG3	1.95	0.49
1:A:83:ILE:HB	2:B:288:TYR:CE2	2.48	0.48
2:B:574:CYS:HA	2:B:625:CYS:HB3	1.95	0.48
1:A:39:GLU:HB3	1:A:542:ASN:HA	1.94	0.48
1:A:493:GLU:CD	1:A:493:GLU:H	2.17	0.48
2:B:370:GLN:HG2	2:B:456:ARG:O	2.14	0.48
1:A:13:ARG:HD2	2:B:275:GLU:HA	1.95	0.48
2:B:889:PRO:HB3	2:B:902:PHE:CE1	2.49	0.47
2:B:412:MET:HG2	2:B:473:PHE:CE2	2.49	0.47
1:A:181:LEU:HD13	1:A:214:TRP:CG	2.50	0.47
1:A:217:PRO:HD3	1:A:230:ILE:HD11	1.96	0.47
2:B:575:TYR:CE1	2:B:576:ARG:HG3	2.50	0.47
2:B:2555:LYS:HG3	2:B:2571:PRO:HB3	1.97	0.47
2:B:413:THR:O	2:B:416:GLN:HB2	2.16	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:412:MET:HG2	2:B:473:PHE:HE2	1.80	0.46
2:B:444:TRP:CZ2	2:B:462:ILE:HD12	2.50	0.46
2:B:2682:LEU:O	2:B:2683:ARG:C	2.53	0.46
2:B:266:ASP:OD1	2:B:287:GLN:HG2	2.16	0.46
2:B:453:LYS:HD3	2:B:454:LEU:H	1.79	0.46
2:B:738:GLN:HA	2:B:749:VAL:HG11	1.98	0.45
2:B:551:GLU:HG2	2:B:551:GLU:O	2.16	0.45
2:B:394:THR:HG22	2:B:396:SER:N	2.32	0.45
1:A:42:ASP:OD1	1:A:527:GLN:HG2	2.17	0.45
2:B:650:GLN:OE1	2:B:690:PRO:HB3	2.17	0.45
2:B:2592:ASN:HB3	6:B:8237:HOH:O	2.17	0.45
2:B:2599:LEU:O	2:B:2604:LYS:HE3	2.17	0.44
2:B:550:ILE:HD11	2:B:607:TYR:CD2	2.53	0.44
1:A:129:LYS:HB3	1:A:129:LYS:HE2	1.66	0.44
1:A:307:SER:N	6:A:612:HOH:O	2.51	0.44
1:A:512:LYS:NZ	1:A:513:ALA:O	2.39	0.43
1:A:76:LYS:HG2	1:A:555:GLU:HB3	2.00	0.43
2:B:893:TYR:CE2	2:B:896:HIS:HA	2.54	0.43
1:A:7:ALA:N	6:A:610:HOH:O	2.50	0.43
1:A:178:VAL:HB	1:A:224:MET:HE1	2.01	0.43
2:B:209:LEU:O	2:B:212:GLU:HG2	2.19	0.43
2:B:709:CYS:N	2:B:722:LEU:O	2.49	0.42
2:B:660:LYS:O	2:B:744:ARG:NH1	2.43	0.42
1:A:334:TYR:O	1:A:341:PRO:HA	2.19	0.42
2:B:597:THR:HA	2:B:600:TRP:CZ2	2.55	0.42
1:A:413:VAL:HA	1:A:414:PRO:HD3	1.93	0.42
2:B:2682:LEU:O	2:B:2684:GLU:N	2.53	0.42
2:B:377:LEU:HD12	2:B:457:VAL:HG11	2.02	0.41
2:B:571:ARG:HD3	6:B:8351:HOH:O	2.20	0.41
2:B:789:LEU:HD11	2:B:901:LYS:HB3	2.02	0.41
2:B:806:ALA:N	6:B:8125:HOH:O	2.51	0.41
2:B:408:SER:O	2:B:412:MET:HB2	2.21	0.41
2:B:476:LYS:HA	2:B:476:LYS:HD2	1.81	0.41
2:B:479:LYS:NZ	6:B:8105:HOH:O	2.34	0.41
2:B:394:THR:HG22	2:B:396:SER:H	1.86	0.41
2:B:2598:ASP:OD1	2:B:2598:ASP:N	2.53	0.41
2:B:394:THR:HB	2:B:397:ASN:HB2	2.02	0.41
1:A:178:VAL:HG12	1:A:224:MET:HE3	2.02	0.41
1:A:344:ALA:HB3	1:A:528:VAL:HG11	2.03	0.41
1:A:393:ALA:N	1:A:396:GLN:HG3	2.36	0.41
2:B:591:SER:OG	2:B:594:GLU:HG3	2.20	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:36:ASP:O	1:A:37:LEU:HD13	2.21	0.40
2:B:828:GLY:HA2	2:B:872:TYR:O	2.22	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	452/605 (75%)	442 (98%)	9 (2%)	1 (0%)	47 57
2	B	749/936 (80%)	726 (97%)	23 (3%)	0	100 100
3	D	4/11 (36%)	4 (100%)	0	0	100 100
All	All	1205/1552 (78%)	1172 (97%)	32 (3%)	1 (0%)	51 63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	393	ALA

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	389/495 (79%)	385 (99%)	4 (1%)	76 86
2	B	672/815 (82%)	661 (98%)	11 (2%)	62 76

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	D	3/6 (50%)	3 (100%)	0	100 100
All	All	1064/1316 (81%)	1049 (99%)	15 (1%)	67 79

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LYS
1	A	175	GLN
1	A	258	ARG
1	A	385	ASP
2	B	264	THR
2	B	285	ARG
2	B	297	ARG
2	B	456	ARG
2	B	621	LEU
2	B	628	VAL
2	B	688	ASP
2	B	739	ARG
2	B	854	SER
2	B	2631	ARG
2	B	2683	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	394	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	$\#\|Z\| > 2$
3	M3L	D	27	3	10,11,12	0.56	0	9,14,16	0.50	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
3	M3L	D	27	3	-	0/9/10/12	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 8 are monoatomic - leaving 1 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	A97	B	8009	-	39,44,44	2.69	14 (35%)	45,63,63	4.55	17 (37%)

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	A97	B	8009	-	-	3/22/34/34	0/5/5/5

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	8009	A97	C23-N24	7.83	1.53	1.37
5	B	8009	A97	O10-C09	7.03	1.37	1.23
5	B	8009	A97	C30-N24	6.15	1.56	1.46
5	B	8009	A97	C25-N24	4.87	1.54	1.46
5	B	8009	A97	C05-C06	3.47	1.44	1.36
5	B	8009	A97	C35-C34	3.09	1.47	1.40
5	B	8009	A97	C14-N13	2.87	1.40	1.33
5	B	8009	A97	C31-C23	2.85	1.43	1.39
5	B	8009	A97	O15-C14	2.79	1.29	1.23
5	B	8009	A97	C03-C04	2.64	1.58	1.51
5	B	8009	A97	C20-C21	2.61	1.43	1.38
5	B	8009	A97	C07-C06	2.43	1.54	1.50
5	B	8009	A97	C32-C33	2.17	1.45	1.40
5	B	8009	A97	C17-C18	2.16	1.43	1.39

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	8009	A97	C19-C31-C23	18.68	124.83	118.25
5	B	8009	A97	C11-C09-N08	15.36	127.17	115.69
5	B	8009	A97	C09-N08-C06	-7.33	118.09	125.20
5	B	8009	A97	O10-C09-C11	-7.09	115.88	124.54
5	B	8009	A97	C31-C23-N24	-6.32	114.78	122.29
5	B	8009	A97	N22-C23-N24	6.25	126.15	116.79
5	B	8009	A97	C16-C14-N13	4.73	122.70	116.30
5	B	8009	A97	C12-N13-C14	4.65	132.38	121.33
5	B	8009	A97	C29-C30-N24	4.54	119.52	110.70
5	B	8009	A97	C30-C29-N27	4.25	115.61	110.80
5	B	8009	A97	C34-C16-C14	3.53	125.80	120.68
5	B	8009	A97	O15-C14-C16	-3.32	117.13	121.72
5	B	8009	A97	C20-C19-C31	-2.96	113.99	118.16
5	B	8009	A97	C28-N27-C29	-2.77	106.51	110.66
5	B	8009	A97	C02-C03-C04	2.23	122.53	114.28

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	8009	A97	C31-C23-N22	-2.11	119.07	122.73
5	B	8009	A97	C30-N24-C23	-2.05	115.62	120.39

There are no chirality outliers.

All (3) torsion outliers are listed below:

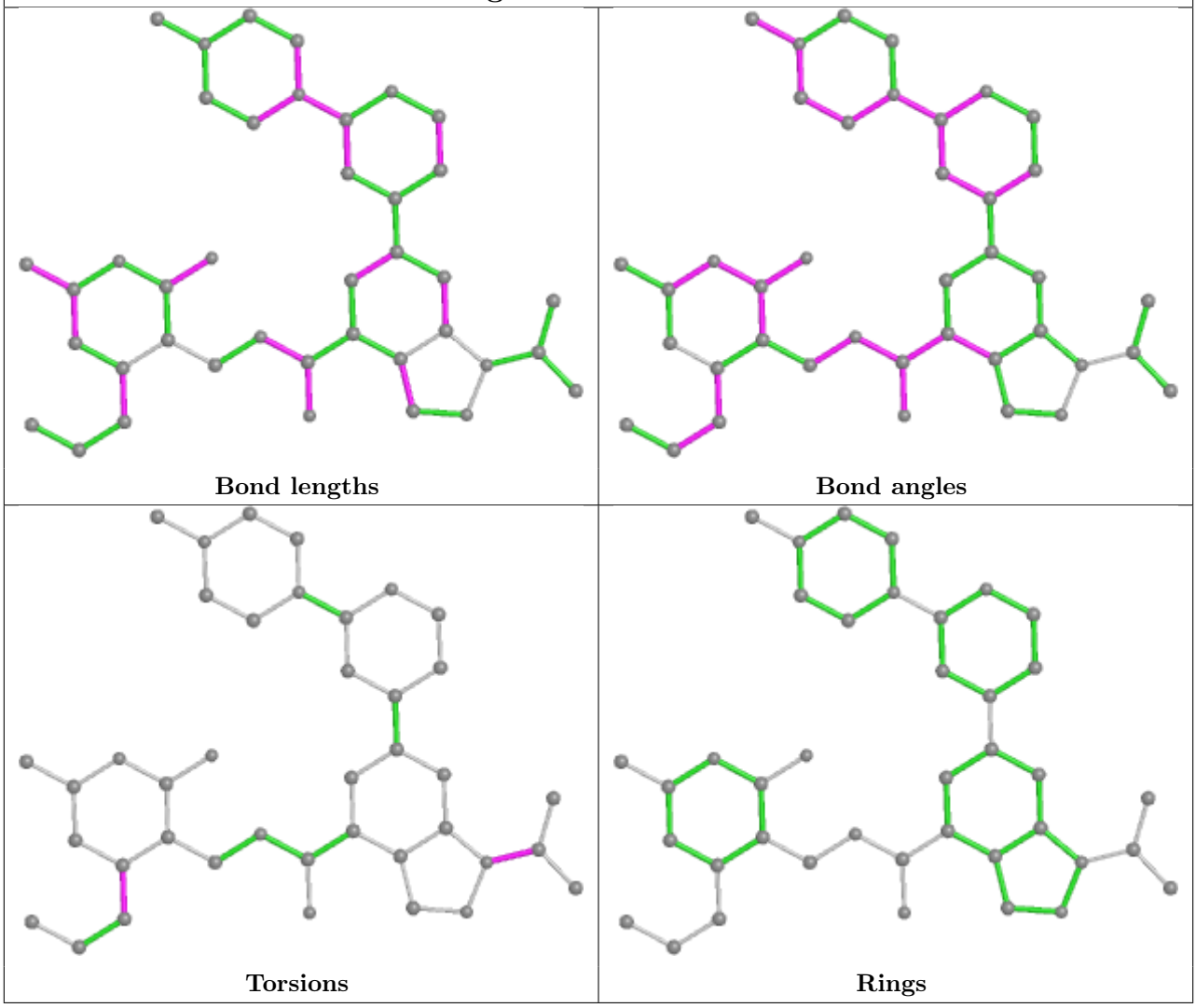
Mol	Chain	Res	Type	Atoms
5	B	8009	A97	C02-C03-C04-C11
5	B	8009	A97	C02-C03-C04-C05
5	B	8009	A97	C39-C38-N37-C33

There are no ring outliers.

No monomer is involved in short contacts.

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 then 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 A97 B 8009



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^{th} 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^2)	Q<0.9
1	A	460/605 (76%)	0.14	33 (7%)	15	19	15, 29, 82, 109	0
2	B	769/936 (82%)	0.02	50 (6%)	18	23	16, 40, 77, 109	0
3	D	6/11 (54%)	2.32	4 (66%)	0	0	46, 54, 70, 80	0
All	All	1235/1552 (79%)	0.08	87 (7%)	16	20	15, 36, 80, 109	0

All (87) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	379	GLY	6.2
2	B	410	ASP	4.7
2	B	550	ILE	4.5
1	A	393	ALA	4.4
1	A	389	GLY	4.3
1	A	375	GLN	4.3
1	A	397	GLN	4.2
1	A	373	ALA	3.9
1	A	378	ASN	3.9
2	B	929	LEU	3.9
2	B	628	VAL	3.8
1	A	556	ARG	3.8
1	A	390	ILE	3.7
1	A	36	ASP	3.7
1	A	492	ARG	3.7
3	D	29	ALA	3.7
2	B	632	LEU	3.6
2	B	927	LYS	3.5
2	B	609	GLN	3.5
2	B	634	GLU	3.4
3	D	30	PRO	3.3
1	A	399	HIS	3.3
2	B	633	GLN	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	400	GLY	3.1
2	B	743	GLY	3.1
2	B	607	TYR	3.1
2	B	454	LEU	3.0
2	B	651	LYS	3.0
2	B	587	VAL	3.0
1	A	394	GLN	2.9
1	A	395	ALA	2.9
2	B	590	TRP	2.8
2	B	415	GLN	2.8
1	A	380	ARG	2.8
2	B	547	ALA	2.8
2	B	409	ASP	2.8
3	D	24	ALA	2.8
2	B	451	GLN	2.8
1	A	34	SER	2.7
2	B	629	HIS	2.7
1	A	377	TYR	2.7
2	B	622	GLY	2.7
2	B	406	PRO	2.7
2	B	414	PRO	2.7
2	B	649	LYS	2.7
2	B	276	TYR	2.7
2	B	411	SER	2.7
1	A	307	SER	2.7
2	B	666	ASP	2.6
2	B	412	MET	2.6
2	B	806	ALA	2.6
1	A	37	LEU	2.5
1	A	405	VAL	2.5
2	B	408	SER	2.5
2	B	575	TYR	2.5
1	A	374	GLN	2.4
1	A	489	GLY	2.4
2	B	647	ILE	2.4
2	B	419	ASN	2.4
2	B	407	GLU	2.4
2	B	602	PHE	2.4
1	A	35	GLN	2.3
2	B	475	ASN	2.3
1	A	388	GLN	2.3
2	B	321	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	548	ALA	2.3
3	D	26	ARG	2.3
1	A	376	SER	2.3
2	B	418	SER	2.3
1	A	370	LEU	2.3
2	B	2683	ARG	2.3
2	B	422	ASP	2.3
1	A	372	GLU	2.3
2	B	569	PRO	2.2
2	B	648	PRO	2.2
2	B	577	ILE	2.2
2	B	650	GLN	2.2
1	A	396	GLN	2.2
2	B	450	ASP	2.1
2	B	492	LEU	2.1
2	B	2682	LEU	2.1
1	A	382	VAL	2.1
1	A	398	VAL	2.1
2	B	2684	GLU	2.1
1	A	301	PRO	2.0
2	B	253	ASP	2.0
1	A	384	VAL	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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^2)	Q<0.9
3	M3L	D	27	12/13	0.84	0.17	35,43,48,51	0

6.3 Carbohydrates (i)

There are no monosaccharides 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,

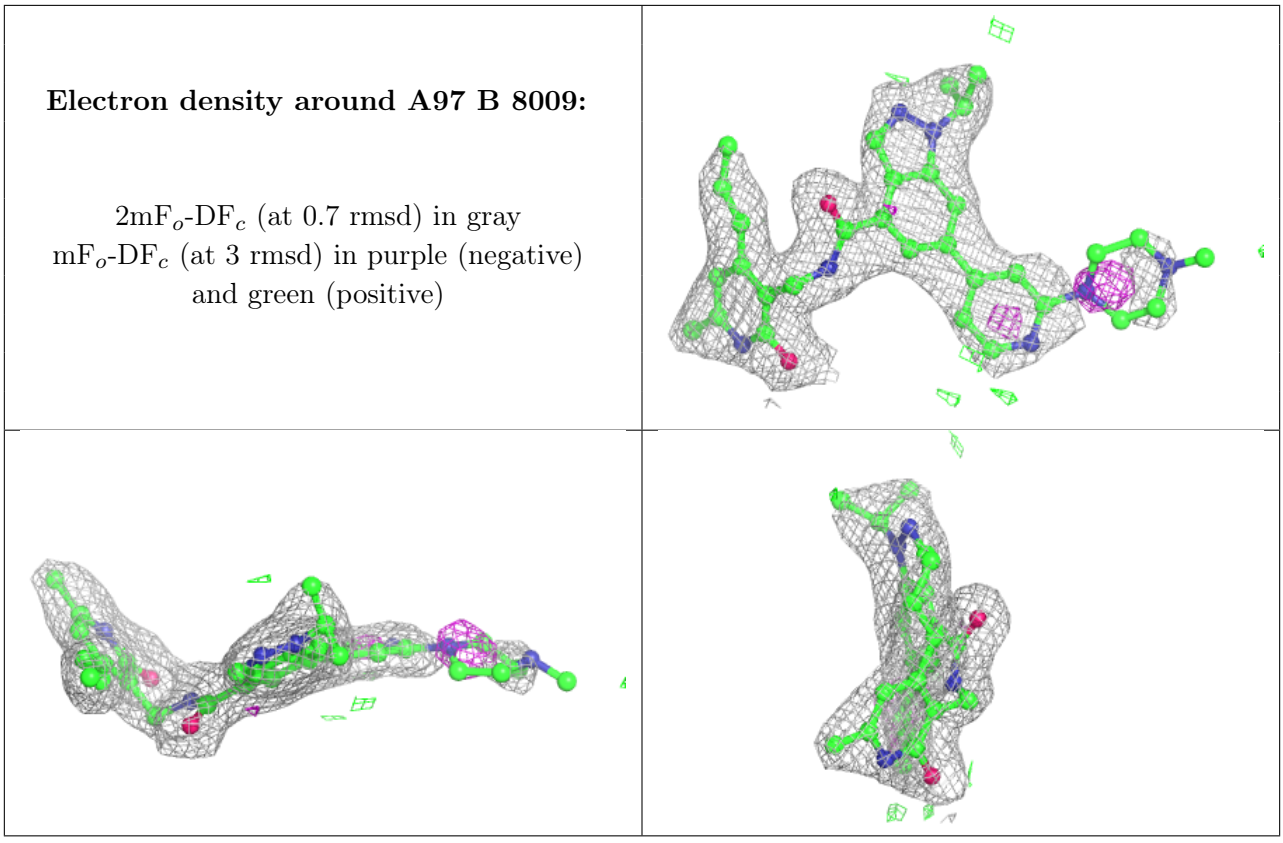
median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	A97	B	8009	40/40	0.89	0.17	24,41,74,75	0
4	ZN	B	8008	1/1	0.98	0.04	55,55,55,55	0
4	ZN	B	8002	1/1	0.99	0.16	28,28,28,28	0
4	ZN	B	8004	1/1	0.99	0.06	37,37,37,37	0
4	ZN	B	8005	1/1	1.00	0.05	36,36,36,36	0
4	ZN	B	8006	1/1	1.00	0.08	33,33,33,33	0
4	ZN	B	8007	1/1	1.00	0.15	23,23,23,23	0
4	ZN	B	8003	1/1	1.00	0.12	30,30,30,30	0
4	ZN	B	8001	1/1	1.00	0.14	28,28,28,28	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 A97 B 8009:

$2mF_o$-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.