Full wwPDB X-ray Structure Validation Report (i)

May 15, 2020 – 08:52 pm BST

PDB ID	:	5RAH
Title	:	PanDDA analysis group deposition – Crystal Structure of JMJD1B in complex with FM010032a
Authors	:	Snee, M.; Nowak, R.; Johansson, C.; Burgess-Brown, N.A.; Arrowsmith, C.H.; Bountra, C.; Edwards, A.M.; Oppermann, U.
Deposited on	:	2020-03-16
Resolution	:	1.66 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 (i) symbol.

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

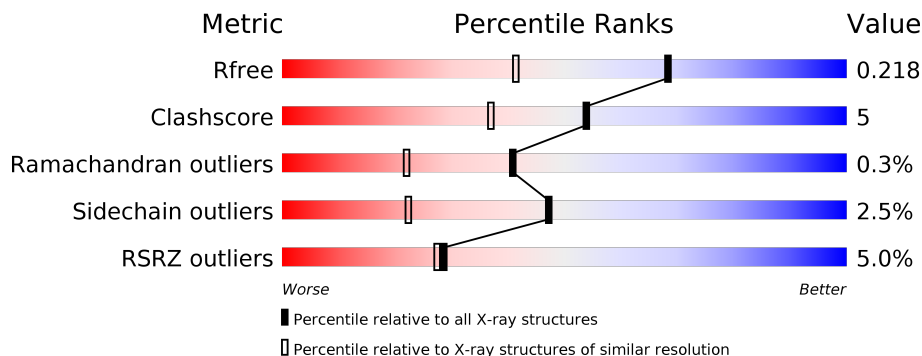|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & 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 1.66 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	Percentile Ranks	Value
Rfree		0.218
Clashscore		5
Ramachandran outliers		0.3%
Sidechain outliers		2.5%
RSRZ outliers		5.0%

Worse   *Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1827 (1.66-1.66)
Clashscore	141614	1931 (1.66-1.66)
Ramachandran outliers	138981	1891 (1.66-1.66)
Sidechain outliers	138945	1891 (1.66-1.66)
RSRZ outliers	127900	1791 (1.66-1.66)

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 >=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	372	5% 81% 9% • 9%
1	B	372	5% 84% 8% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6283 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 Lysine-specific demethylase 3B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	340	Total 2764	C 1749	N 482	O 517	S 16	0	4	0
1	B	344	Total 2798	C 1774	N 492	O 517	S 15	0	2	0

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1357	MET	-	initiating methionine	UNP Q7LBC6
A	1358	HIS	-	expression tag	UNP Q7LBC6
A	1359	HIS	-	expression tag	UNP Q7LBC6
A	1360	HIS	-	expression tag	UNP Q7LBC6
A	1361	HIS	-	expression tag	UNP Q7LBC6
A	1362	HIS	-	expression tag	UNP Q7LBC6
A	1363	HIS	-	expression tag	UNP Q7LBC6
A	1364	SER	-	expression tag	UNP Q7LBC6
A	1365	SER	-	expression tag	UNP Q7LBC6
A	1366	GLY	-	expression tag	UNP Q7LBC6
A	1367	VAL	-	expression tag	UNP Q7LBC6
A	1368	ASP	-	expression tag	UNP Q7LBC6
A	1369	LEU	-	expression tag	UNP Q7LBC6
A	1370	GLY	-	expression tag	UNP Q7LBC6
A	1371	THR	-	expression tag	UNP Q7LBC6
A	1372	GLU	-	expression tag	UNP Q7LBC6
A	1373	ASN	-	expression tag	UNP Q7LBC6
A	1374	LEU	-	expression tag	UNP Q7LBC6
A	1375	TYR	-	expression tag	UNP Q7LBC6
A	1376	PHE	-	expression tag	UNP Q7LBC6
A	1377	GLN	-	expression tag	UNP Q7LBC6
A	1378	SER	-	expression tag	UNP Q7LBC6
A	1379	MET	-	expression tag	UNP Q7LBC6
A	1601	GLU	GLN	conflict	UNP Q7LBC6
A	1606	HIS	GLY	conflict	UNP Q7LBC6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1357	MET	-	initiating methionine	UNP Q7LBC6
B	1358	HIS	-	expression tag	UNP Q7LBC6
B	1359	HIS	-	expression tag	UNP Q7LBC6
B	1360	HIS	-	expression tag	UNP Q7LBC6
B	1361	HIS	-	expression tag	UNP Q7LBC6
B	1362	HIS	-	expression tag	UNP Q7LBC6
B	1363	HIS	-	expression tag	UNP Q7LBC6
B	1364	SER	-	expression tag	UNP Q7LBC6
B	1365	SER	-	expression tag	UNP Q7LBC6
B	1366	GLY	-	expression tag	UNP Q7LBC6
B	1367	VAL	-	expression tag	UNP Q7LBC6
B	1368	ASP	-	expression tag	UNP Q7LBC6
B	1369	LEU	-	expression tag	UNP Q7LBC6
B	1370	GLY	-	expression tag	UNP Q7LBC6
B	1371	THR	-	expression tag	UNP Q7LBC6
B	1372	GLU	-	expression tag	UNP Q7LBC6
B	1373	ASN	-	expression tag	UNP Q7LBC6
B	1374	LEU	-	expression tag	UNP Q7LBC6
B	1375	TYR	-	expression tag	UNP Q7LBC6
B	1376	PHE	-	expression tag	UNP Q7LBC6
B	1377	GLN	-	expression tag	UNP Q7LBC6
B	1378	SER	-	expression tag	UNP Q7LBC6
B	1379	MET	-	expression tag	UNP Q7LBC6
B	1601	GLU	GLN	conflict	UNP Q7LBC6
B	1606	HIS	GLY	conflict	UNP Q7LBC6

- Molecule 2 is cyclopropyl-[4-(4-fluorophenyl)piperazin-1-yl]methanone (three-letter code: GX4) (formula: $C_{14}H_{17}FN_2O$) (labeled as "Ligand of Interest" by author).

GX4
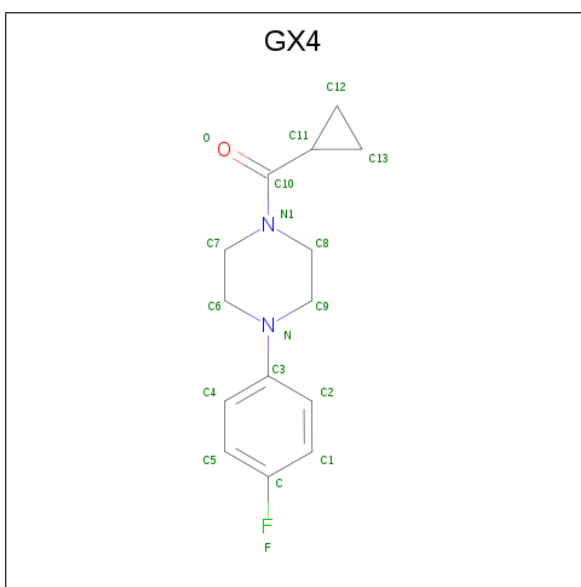
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	F	N	O		
2	A	1	18	14	1	2	1	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	A	4	Total	Cl	0	0
			4	4		

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Mn	0	0
			1	1		
4	A	1	Total	Mn	0	0
			1	1		

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	375	Total	O	0	0
			375	375		
5	B	321	Total	O	0	0
			321	321		

3 Residue-property plots ⓘ

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: Lysine-specific demethylase 3B
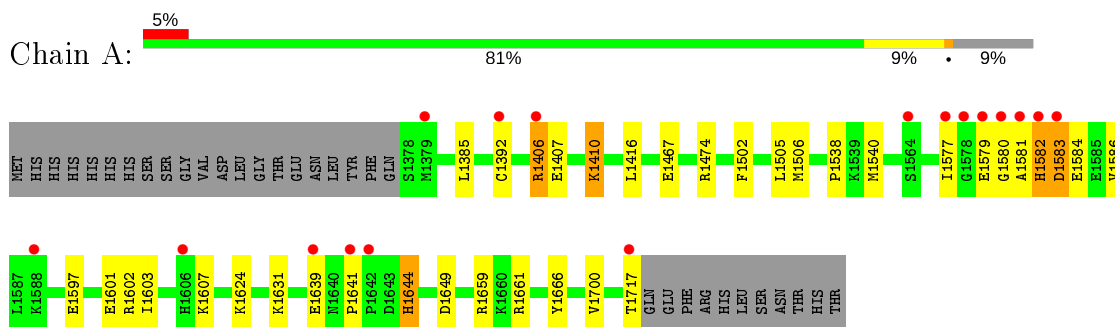
- Molecule 1: Lysine-specific demethylase 3B
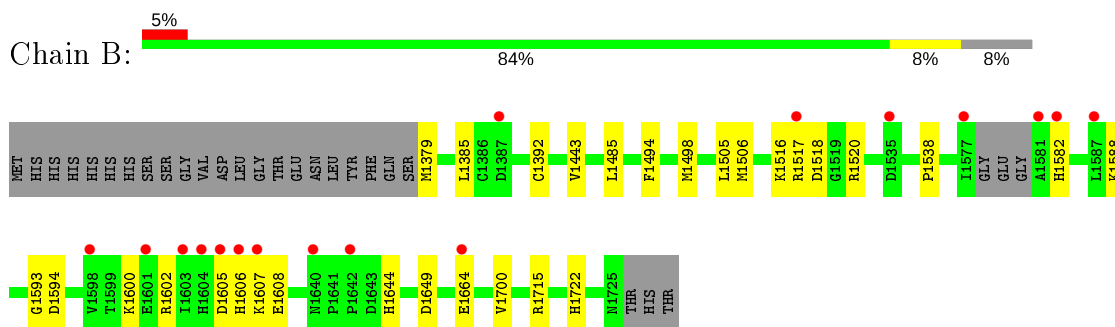
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.37Å　　93.59Å　　93.58Å 90.00°　　107.85°　　90.00°	Depositor
Resolution (Å)	89.08　–　1.66 89.08　–　1.66	Depositor EDS
% Data completeness (in resolution range)	99.0 (89.08-1.66) 99.0 (89.08-1.66)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.45 (at 1.66Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.180　,　0.205 0.195　,　0.218	Depositor DCC
R_{free} test set	5558 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	23.5	Xtriage
Anisotropy	0.064	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 49.1	EDS
L-test for twinning[2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6283	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.79% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GX4, MN, CL

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	0.91	2/2837 (0.1%)	0.92	5/3849 (0.1%)
1	B	0.83	0/2872	0.86	1/3895 (0.0%)
All	All	0.87	2/5709 (0.0%)	0.89	6/7744 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1467	GLU	CD-OE2	-10.24	1.14	1.25
1	A	1602	ARG	C-O	5.47	1.33	1.23

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1474	ARG	NE-CZ-NH1	8.88	124.74	120.30
1	B	1602	ARG	NE-CZ-NH2	-6.27	117.17	120.30
1	A	1474	ARG	NE-CZ-NH2	-5.83	117.38	120.30
1	A	1540	MET	CG-SD-CE	-5.82	90.89	100.20
1	A	1661	ARG	CG-CD-NE	-5.26	100.75	111.80
1	A	1659	ARG	NE-CZ-NH1	5.20	122.90	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2764	0	2646	32	0
1	B	2798	0	2685	26	0
2	A	18	0	0	0	0
3	A	4	0	0	0	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	375	0	0	9	3
5	B	321	0	0	9	1
All	All	6283	0	5331	58	3

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 5.

All (58) 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:1518:ASP:HB3	5:B:2143:HOH:O	1.84	0.78
1:A:1582:HIS:CE1	5:A:1930:HOH:O	2.38	0.77
1:A:1639:GLU:O	5:A:1902:HOH:O	2.07	0.72
1:B:1644:HIS:NE2	1:B:1649:ASP:OD2	2.24	0.70
1:B:1518:ASP:OD1	5:B:1901:HOH:O	2.09	0.68
1:A:1601:GLU:OE1	5:A:1903:HOH:O	2.11	0.68
1:A:1582:HIS:O	1:A:1582:HIS:CG	2.46	0.67
1:B:1494:PHE:CZ	1:B:1498:MET:HE3	2.32	0.65
1:A:1406[A]:ARG:NH1	5:A:1907:HOH:O	2.30	0.64
1:B:1517[B]:ARG:HH11	1:B:1517[B]:ARG:CG	2.10	0.64
1:B:1517[B]:ARG:HB2	1:B:1517[B]:ARG:HH11	1.62	0.64
1:B:1664[B]:GLU:HA	1:B:1664[B]:GLU:OE1	1.98	0.63
1:B:1715:ARG:HG3	5:B:2131:HOH:O	1.99	0.63
1:B:1516:LYS:HE3	5:B:1930:HOH:O	1.99	0.62
1:A:1631:LYS:HE2	1:A:1666:TYR:OH	2.00	0.62
1:A:1407:GLU:HG3	1:A:1410:LYS:HE2	1.81	0.62
1:B:1506:MET:CE	1:B:1538:PRO:HG3	2.31	0.60
1:B:1517[B]:ARG:CB	1:B:1517[B]:ARG:HH11	2.15	0.60
1:A:1385:LEU:HD12	1:A:1392[B]:CYS:SG	2.41	0.59
1:A:1406[A]:ARG:NH2	5:A:1909:HOH:O	2.35	0.58
1:A:1502:PHE:CE2	1:A:1506:MET:HE2	2.39	0.58
1:A:1407:GLU:CB	1:A:1410:LYS:HZ3	2.17	0.57
1:B:1715:ARG:HD3	5:B:2187:HOH:O	2.07	0.55
1:A:1407:GLU:CG	1:A:1410:LYS:NZ	2.71	0.54
1:A:1597:GLU:OE1	5:A:1904:HOH:O	2.18	0.54

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1385:LEU:HG	1:A:1392[B]:CYS:SG	2.48	0.54
1:A:1407:GLU:HG3	1:A:1410:LYS:CE	2.38	0.53
1:A:1407:GLU:CG	1:A:1410:LYS:HZ3	2.21	0.53
1:A:1385:LEU:CD1	1:A:1392[B]:CYS:SG	2.97	0.53
1:A:1407:GLU:HG3	1:A:1410:LYS:NZ	2.24	0.53
1:A:1624:LYS:NZ	5:A:1901:HOH:O	1.65	0.53
1:A:1583:ASP:C	1:A:1583:ASP:OD1	2.48	0.52
1:B:1443:VAL:HG21	1:B:1485:LEU:HD11	1.92	0.51
1:B:1505:LEU:HD22	1:B:1700:VAL:HG11	1.91	0.51
1:B:1517[B]:ARG:NH1	1:B:1517[B]:ARG:HG3	2.24	0.51
1:B:1379:MET:N	5:B:1911:HOH:O	2.45	0.49
1:A:1644:HIS:NE2	1:A:1649:ASP:OD2	2.45	0.49
1:A:1506:MET:HE1	1:A:1538:PRO:HD3	1.94	0.49
1:B:1517[B]:ARG:HH11	1:B:1517[B]:ARG:HG3	1.76	0.47
1:B:1600:LYS:HE3	5:B:1954:HOH:O	2.14	0.47
1:A:1407:GLU:O	1:A:1410:LYS:HG2	2.15	0.47
1:A:1583:ASP:O	1:A:1583:ASP:OD1	2.33	0.46
1:B:1520:ARG:NE	5:B:1918:HOH:O	2.48	0.46
1:A:1641:PRO:HD2	5:A:1982:HOH:O	2.15	0.46
1:A:1506:MET:CE	1:A:1538:PRO:HD3	2.46	0.46
1:B:1385:LEU:HD12	1:B:1392:CYS:SG	2.56	0.45
1:B:1517[B]:ARG:NH1	1:B:1517[B]:ARG:CG	2.72	0.45
1:A:1582:HIS:CE1	1:A:1586:VAL:HG23	2.52	0.45
1:A:1649:ASP:OD1	5:A:1905:HOH:O	2.21	0.44
1:A:1392[A]:CYS:SG	1:A:1416:LEU:HD22	2.57	0.44
1:A:1407:GLU:OE1	1:A:1410:LYS:NZ	2.45	0.44
1:A:1505:LEU:HD22	1:A:1700:VAL:HG11	1.98	0.44
1:B:1644:HIS:CE1	1:B:1649:ASP:OD2	2.71	0.44
1:A:1582:HIS:O	1:A:1582:HIS:ND1	2.50	0.43
1:B:1606:HIS:HB2	1:B:1608:GLU:CD	2.39	0.42
1:B:1722:HIS:ND1	5:B:1907:HOH:O	2.37	0.42
1:B:1506:MET:HE2	1:B:1538:PRO:HG3	1.99	0.42
1:B:1593:GLY:O	1:B:1594:ASP:HB2	2.20	0.41

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:A:2211:HOH:O	5:A:2230:HOH:O[2_648]	1.63	0.57
5:A:1929:HOH:O	5:A:2027:HOH:O[2_648]	1.88	0.32
5:A:2192:HOH:O	5:B:1918:HOH:O[2_557]	2.09	0.11

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	342/372 (92%)	336 (98%)	4 (1%)	2 (1%)	25	8
1	B	342/372 (92%)	332 (97%)	10 (3%)	0	100	100
All	All	684/744 (92%)	668 (98%)	14 (2%)	2 (0%)	41	22

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1581	ALA
1	A	1580	GLY

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	299/331 (90%)	287 (96%)	12 (4%)	31	9
1	B	302/331 (91%)	298 (99%)	4 (1%)	69	50
All	All	601/662 (91%)	585 (97%)	16 (3%)	47	19

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1406[A]	ARG
1	A	1406[B]	ARG
1	A	1410	LYS
1	A	1577	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1579	GLU
1	A	1582	HIS
1	A	1583	ASP
1	A	1584	GLU
1	A	1603	ILE
1	A	1607	LYS
1	A	1644	HIS
1	A	1717	THR
1	B	1582	HIS
1	B	1588	LYS
1	B	1605	ASP
1	B	1607	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1650	GLN
1	B	1656	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)

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 8 ligands modelled in this entry, 7 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

Full wwPDB X-ray Structure Validation Report

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	#$|Z| > 2$
2	GX4	A	1801	-	20,20,20	1.98	4 (20%)	28,28,28	1.48	4 (14%)

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
2	GX4	A	1801	-	-	4/12/24/24	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1801	GX4	C10-N1	5.55	1.42	1.34
2	A	1801	GX4	C9-N	3.77	1.52	1.46
2	A	1801	GX4	C8-N1	3.10	1.52	1.47
2	A	1801	GX4	C5-C	2.04	1.41	1.37

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	1801	GX4	C13-C11-C10	5.09	123.63	117.01
2	A	1801	GX4	C11-C10-N1	-2.72	115.69	118.80
2	A	1801	GX4	C1-C-C5	-2.38	119.67	122.83
2	A	1801	GX4	C2-C1-C	2.15	120.58	118.36

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1801	GX4	N1-C10-C11-C12
2	A	1801	GX4	O-C10-C11-C12
2	A	1801	GX4	C4-C3-N-C6
2	A	1801	GX4	C2-C3-N-C6

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 GX4 A 1801
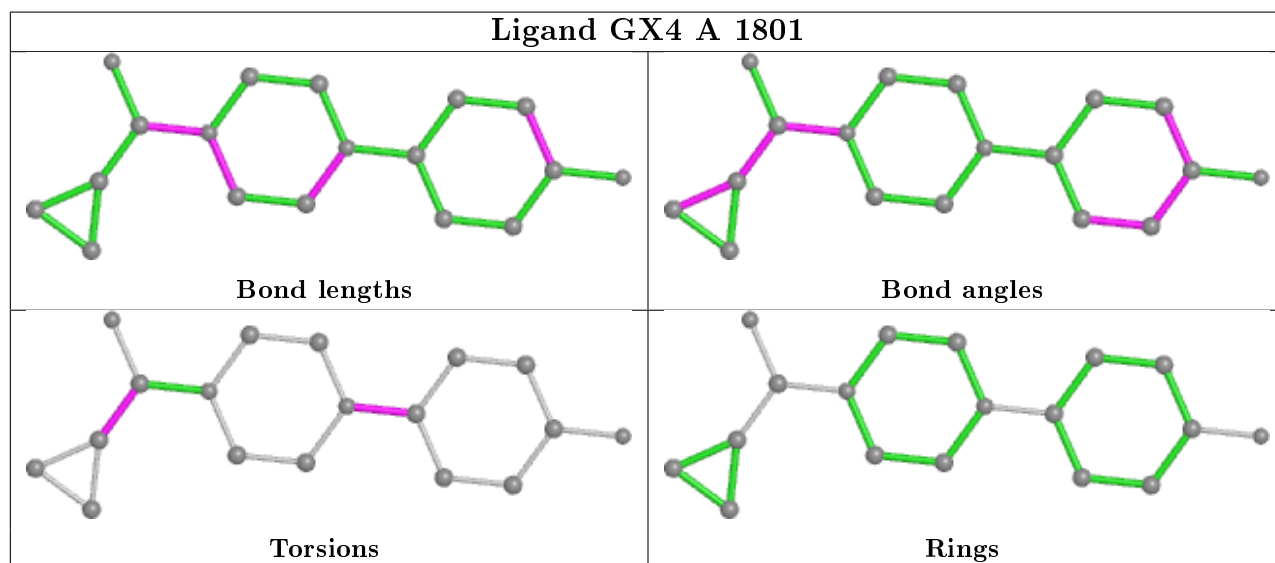
5.7 Other polymers (i)

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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	340/372 (91%)	-0.11	17 (5%)	28	27	14, 24, 59, 94	13 (3%)
1	B	344/372 (92%)	-0.09	17 (4%)	29	28	18, 30, 68, 116	2 (0%)
All	All	684/744 (91%)	-0.10	34 (4%)	28	27	14, 27, 66, 116	15 (2%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1577	ILE	11.2
1	A	1392[A]	CYS	9.9
1	A	1581	ALA	9.3
1	B	1603	ILE	7.2
1	B	1606	HIS	7.1
1	B	1601	GLU	6.6
1	A	1582	HIS	6.2
1	A	1578	GLY	5.7
1	A	1580	GLY	5.7
1	A	1564[A]	SER	5.4
1	A	1406[A]	ARG	5.4
1	B	1581	ALA	5.3
1	A	1588[A]	LYS	5.3
1	B	1605	ASP	5.3
1	B	1577	ILE	5.2
1	B	1517[A]	ARG	4.8
1	A	1583	ASP	4.3
1	A	1579	GLU	4.2
1	B	1604	HIS	4.2
1	A	1379	MET	4.2
1	B	1664[A]	GLU	3.6
1	B	1587	LEU	3.3
1	B	1582	HIS	3.1
1	B	1642	PRO	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1717	THR	2.9
1	B	1607	LYS	2.9
1	B	1535	ASP	2.8
1	A	1642	PRO	2.7
1	A	1639	GLU	2.3
1	A	1606	HIS	2.3
1	A	1641	PRO	2.2
1	B	1640	ASN	2.1
1	B	1598	VAL	2.1
1	B	1387	ASP	2.0

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

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

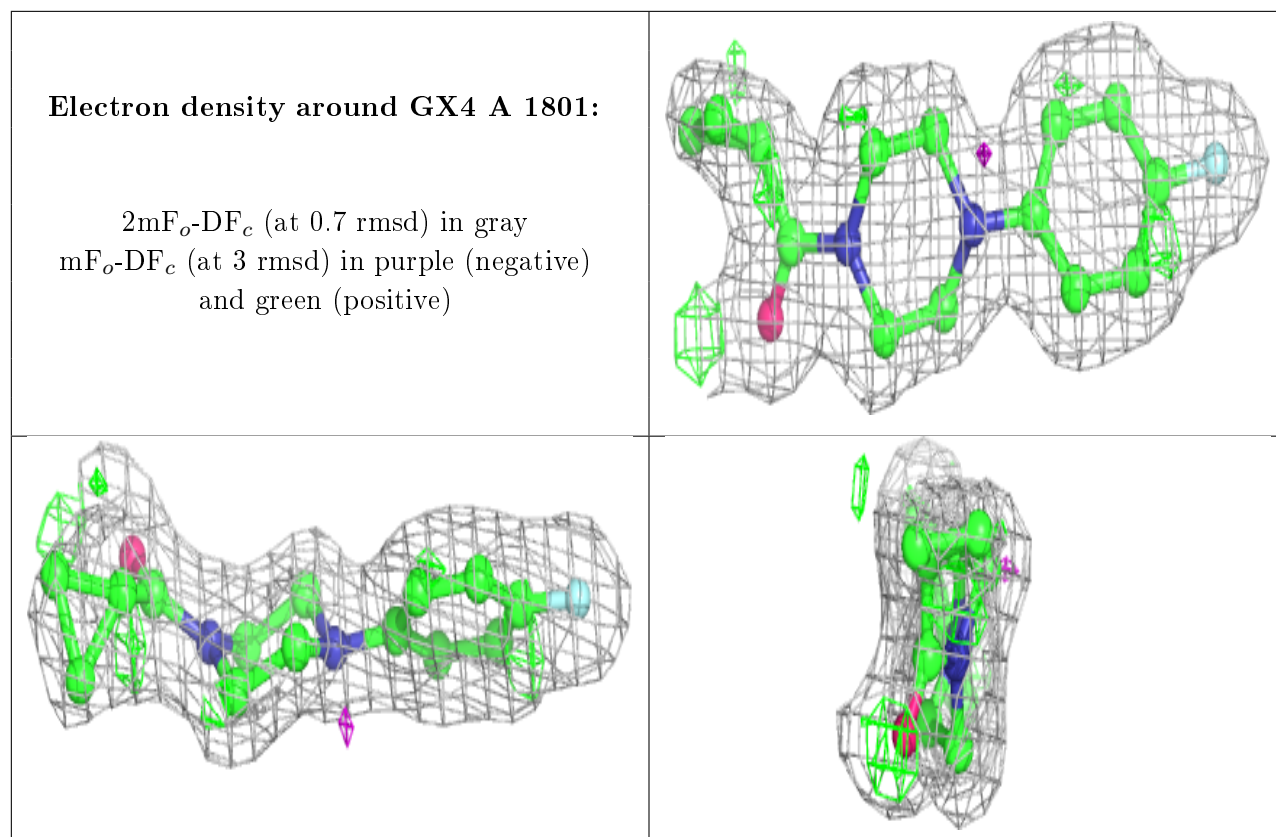## 6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th 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
4	MN	A	1806	1/1	0.49	0.13	60,60,60,60	0
3	CL	A	1805	1/1	0.92	0.05	54,54,54,54	0
2	GX4	A	1801	18/18	0.94	0.13	18,22,35,36	18
4	MN	B	1802	1/1	0.96	0.08	39,39,39,39	0
3	CL	A	1804	1/1	0.96	0.06	34,34,34,34	0
3	CL	A	1803	1/1	0.99	0.09	32,32,32,32	0
3	CL	B	1801	1/1	0.99	0.04	41,41,41,41	0
3	CL	A	1802	1/1	1.00	0.10	22,22,22,22	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 GX4 A 1801:

$2mF_o$-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.