Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 05:08 pm BST

PDB ID	:	5RB7
Title	:	PanDDA analysis group deposition – Crystal Structure of JMJD1B in complex with FM001648a
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	:	1.57 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:

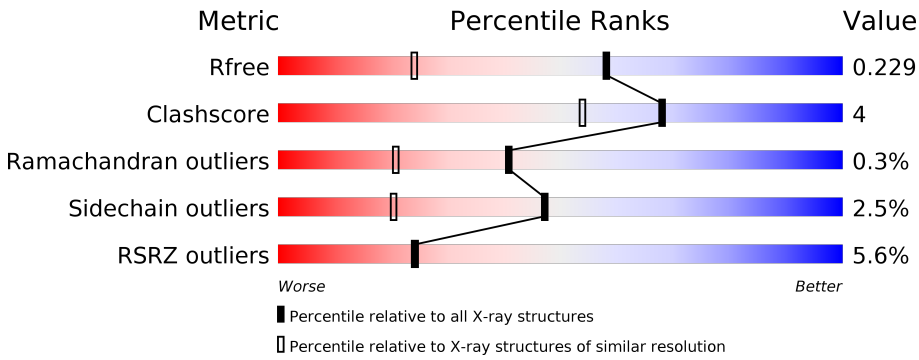
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.57 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	5534 (1.60-1.56)
Clashscore	141614	5861 (1.60-1.56)
Ramachandran outliers	138981	5708 (1.60-1.56)
Sidechain outliers	138945	5703 (1.60-1.56)
RSRZ outliers	127900	5431 (1.60-1.56)

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	6% 85% 6% • 9%
1	B	372	4% 83% 9% 8%

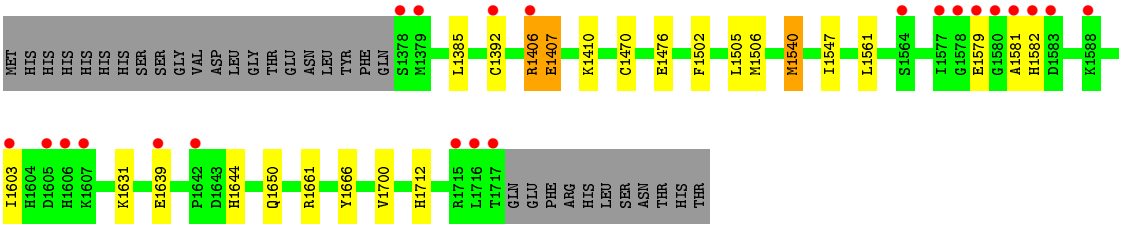
ENTRY-COMPOSITION INFOmissingINFO

2 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.

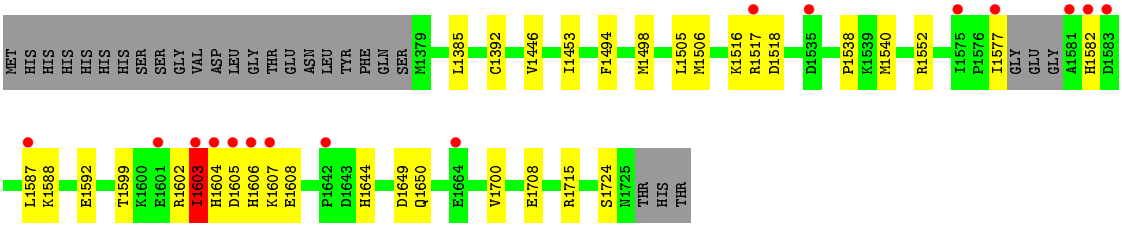
- Molecule 1: Lysine-specific demethylase 3B

Chain A:



- Molecule 1: Lysine-specific demethylase 3B

Chain B:



3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	57.89Å 93.77Å 93.77Å 90.00° 107.94° 90.00°	Depositor
Resolution (Å)	64.72 – 1.57 64.63 – 1.57	Depositor EDS
% Data completeness (in resolution range)	98.4 (64.72-1.57) 98.5 (64.63-1.57)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.57Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.193 , 0.217 0.207 , 0.229	Depositor DCC
R_{free} test set	6562 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	23.2	Xtriage
Anisotropy	0.105	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 46.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.026 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6282	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 4.83% 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.

4 Model quality (i)

4.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: S4D, 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	0.88	0/2841	0.92	3/3854 (0.1%)
1	B	0.83	1/2872 (0.0%)	0.86	0/3895
All	All	0.85	1/5713 (0.0%)	0.89	3/7749 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1708	GLU	CD-OE2	5.44	1.31	1.25

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	1540	MET	CG-SD-CE	-6.01	90.58	100.20
1	A	1470	CYS	N-CA-CB	5.68	120.82	110.60
1	A	1661	ARG	CG-CD-NE	-5.44	100.37	111.80

There are no chirality outliers.

There are no planarity outliers.

4.2 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	2768	0	2650	20	0
1	B	2798	0	2685	22	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	17	0	0	0	0
3	A	4	0	0	1	0
3	B	1	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	373	0	0	7	0
5	B	319	0	0	6	0
All	All	6282	0	5335	42	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 (42) 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:1407:GLU:HG2	5:A:2205:HOH:O	1.66	0.96
1:B:1517[B]:ARG:HG3	1:B:1517[B]:ARG:HH11	1.36	0.88
1:A:1406[A]:ARG:NH1	5:A:1902:HOH:O	2.11	0.82
1:B:1517[B]:ARG:HH11	1:B:1517[B]:ARG:CG	1.94	0.81
1:B:1715:ARG:HG3	5:B:2114:HOH:O	1.82	0.78
1:B:1518:ASP:HB3	5:B:2138:HOH:O	1.84	0.77
1:B:1516:LYS:HE3	5:B:1926:HOH:O	1.86	0.74
1:B:1517[B]:ARG:HG3	1:B:1517[B]:ARG:NH1	1.99	0.74
1:A:1476:GLU:CB	5:A:2254:HOH:O	2.37	0.72
1:A:1631:LYS:HE2	1:A:1666:TYR:OH	1.95	0.67
1:A:1407:GLU:O	1:A:1410:LYS:HG2	1.95	0.65
1:A:1582:HIS:CE1	5:A:1977:HOH:O	2.49	0.65
1:B:1715:ARG:HD3	5:B:2177:HOH:O	2.00	0.62
1:B:1517[A]:ARG:NH1	5:B:1904:HOH:O	2.30	0.59
1:A:1406[A]:ARG:NH2	5:A:1909:HOH:O	2.37	0.57
1:B:1552:ARG:NH1	1:B:1592:GLU:OE1	2.36	0.57
1:B:1505:LEU:HD22	1:B:1700:VAL:HG11	1.86	0.56
1:A:1385:LEU:HD12	1:A:1392[B]:CYS:SG	2.46	0.55
1:B:1644:HIS:NE2	1:B:1649:ASP:OD2	2.41	0.54
1:A:1385:LEU:HG	1:A:1392[B]:CYS:SG	2.50	0.52
1:A:1547:ILE:HA	1:A:1582:HIS:HE2	1.75	0.51
1:B:1385:LEU:HD12	1:B:1392:CYS:SG	2.51	0.51
1:A:1385:LEU:CD1	1:A:1392[B]:CYS:SG	2.99	0.51
1:A:1582:HIS:CG	1:A:1582:HIS:O	2.64	0.50
1:A:1502:PHE:CE2	1:A:1506:MET:HE2	2.47	0.50
1:B:1506:MET:CE	1:B:1538:PRO:HG3	2.43	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1712:HIS:NE2	3:A:1805:CL:CL	2.71	0.47
1:B:1446:VAL:HG22	1:B:1453:ILE:HD13	1.94	0.47
1:B:1517[B]:ARG:CB	1:B:1517[B]:ARG:HH11	2.28	0.46
1:B:1587:LEU:HD13	1:B:1603:ILE:HD11	1.96	0.46
1:B:1494:PHE:CZ	1:B:1498:MET:HE3	2.50	0.46
1:B:1644:HIS:CE1	1:B:1649:ASP:OD2	2.69	0.45
1:B:1602:ARG:HH11	1:B:1608:GLU:HG2	1.83	0.44
1:A:1650:GLN:NE2	5:A:1916:HOH:O	2.51	0.44
1:B:1606:HIS:HB2	1:B:1608:GLU:OE2	2.18	0.44
1:A:1547:ILE:HG23	1:A:1582:HIS:CD2	2.53	0.42
1:B:1599:THR:O	1:B:1603:ILE:HG23	2.19	0.42
1:A:1547:ILE:HA	1:A:1582:HIS:NE2	2.35	0.42
1:A:1505:LEU:HD22	1:A:1700:VAL:HG11	2.01	0.42
1:B:1650:GLN:OE1	5:B:1901:HOH:O	2.22	0.42
1:A:1639:GLU:O	5:A:1903:HOH:O	2.21	0.41
1:A:1561:LEU:C	1:A:1561:LEU:HD12	2.40	0.41

There are no symmetry-related clashes.

4.3 Torsion angles (i)

4.3.1 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	343/372 (92%)	334 (97%)	8 (2%)	1 (0%)	41 21
1	B	342/372 (92%)	331 (97%)	10 (3%)	1 (0%)	41 21
All	All	685/744 (92%)	665 (97%)	18 (3%)	2 (0%)	41 21

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1581	ALA
1	B	1603	ILE

4.3.2 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%)	292 (98%)	7 (2%)	50	23
1	B	302/331 (91%)	293 (97%)	9 (3%)	41	15
All	All	601/662 (91%)	585 (97%)	16 (3%)	47	18

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	1407	GLU
1	A	1540	MET
1	A	1579	GLU
1	A	1603	ILE
1	A	1644	HIS
1	B	1540	MET
1	B	1577	ILE
1	B	1582	HIS
1	B	1588	LYS
1	B	1603	ILE
1	B	1604	HIS
1	B	1605	ASP
1	B	1607	LYS
1	B	1724	SER

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	1650	GLN

4.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates (i)

There are no carbohydrates in this entry.

4.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 7 are monoatomic and 1 could not be matched to an existing wwPDB Chemical Component Dictionary definition at this stage - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

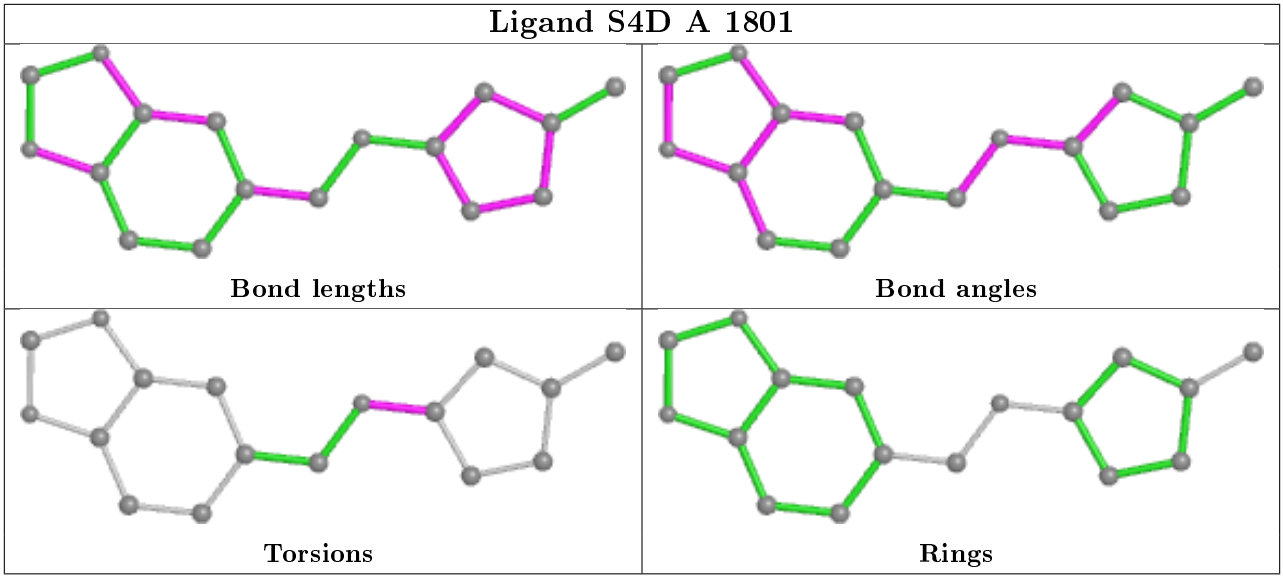
There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.



4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 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	22 (6%)	18	19	15, 26, 55, 126	8 (2%)
1	B	344/372 (92%)	-0.00	16 (4%)	31	31	17, 29, 61, 129	2 (0%)
All	All	684/744 (91%)	0.05	38 (5%)	24	24	15, 27, 59, 129	10 (1%)

All (38) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1581	ALA	12.5
1	B	1605	ASP	11.6
1	A	1578	GLY	11.6
1	A	1577	ILE	9.9
1	B	1581	ALA	9.6
1	A	1392[A]	CYS	9.6
1	A	1379[A]	MET	8.5
1	A	1716	LEU	7.6
1	A	1606	HIS	6.7
1	B	1577	ILE	6.4
1	B	1606	HIS	6.1
1	A	1564[A]	SER	6.1
1	A	1588[A]	LYS	5.6
1	A	1406[A]	ARG	5.5
1	B	1603	ILE	5.4
1	B	1517[A]	ARG	5.2
1	B	1604	HIS	5.0
1	A	1582	HIS	5.0
1	A	1580	GLY	5.0
1	A	1715	ARG	4.7
1	B	1587	LEU	4.6
1	B	1664[A]	GLU	4.2
1	A	1583	ASP	4.2
1	A	1605	ASP	4.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	1583	ASP	3.3
1	A	1579	GLU	3.1
1	B	1601	GLU	2.8
1	B	1582	HIS	2.7
1	A	1603	ILE	2.7
1	A	1607	LYS	2.6
1	B	1607	LYS	2.6
1	A	1639	GLU	2.5
1	A	1717	THR	2.4
1	B	1642	PRO	2.3
1	B	1575	ILE	2.2
1	A	1642	PRO	2.2
1	A	1378	SER	2.2
1	B	1535	ASP	2.1

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

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

5.3 Carbohydrates (i)

There are no carbohydrates in this entry.

5.4 Ligands (i)

LIGAND-RSR INFOmissingINFO

5.5 Other polymers (i)

There are no such residues in this entry.