Full wwPDB X-ray Structure Validation Report ⓘ

May 17, 2020 – 07:10 am BST

PDB ID	:	5RAP
Title	:	PanDDA analysis group deposition – Crystal Structure of JMJD1B in complex with FM000707a
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	:	1.90 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 following versions of software and data (see references ⓘ) were used in the production of this report:

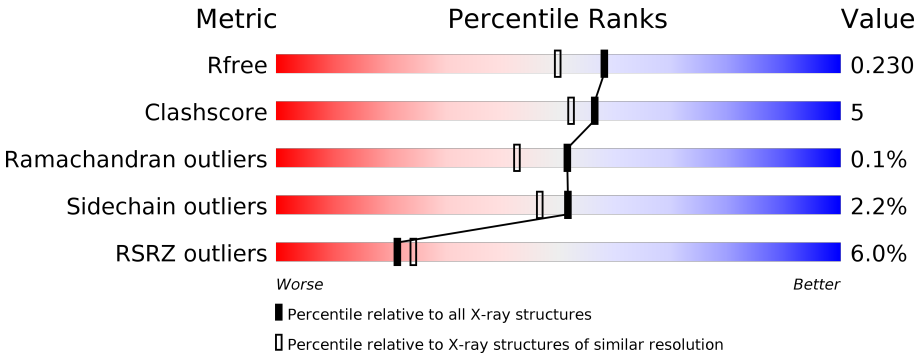
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	7% 84% 7% 9%
1	B	372	4% 83% 9% 8%

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
2	S64	A	1801	-	-	X	-

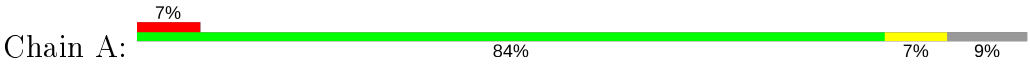
ENTRY-COMPOSITION INFOmissingINFO

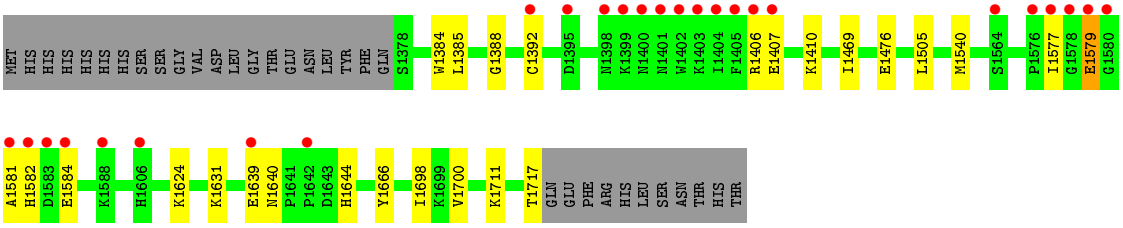
2 Residue-property plots (i)

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

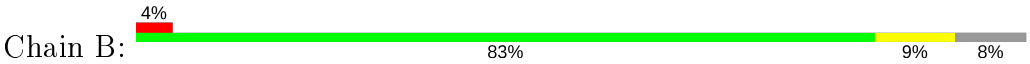
Chain A:

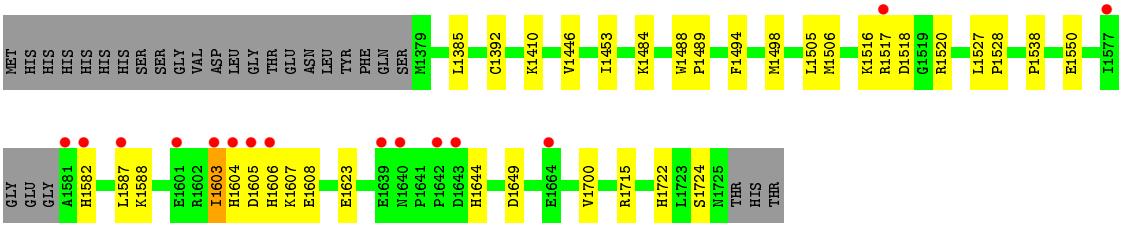




- Molecule 1: Lysine-specific demethylase 3B

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	57.49Å 93.60Å 93.19Å 90.00° 107.69° 90.00°	Depositor
Resolution (Å)	54.83 – 1.90 54.77 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.0 (54.83-1.90) 99.0 (54.77-1.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.73 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.178 , 0.215 0.194 , 0.230	Depositor DCC
R_{free} test set	3648 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	27.4	Xtriage
Anisotropy	0.140	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 50.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.033 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6278	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.87% 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

4.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: S64, 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	0.76	0/2833	0.83	0/3844
1	B	0.73	0/2872	0.81	0/3895
All	All	0.74	0/5705	0.82	0/7739

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	2760	0	2642	17	0
1	B	2798	0	2685	28	0
2	A	17	0	0	10	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	371	0	0	8	3
5	B	325	0	0	7	1
All	All	6278	0	5327	55	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 (55) 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:A:1801:S64:C13	2:A:1801:S64:C14	1.78	1.56
2:A:1801:S64:C07	2:A:1801:S64:N08	1.74	1.49
2:A:1801:S64:C15	2:A:1801:S64:C11	1.89	1.49
1:A:1624:LYS:NZ	5:A:1901:HOH:O	1.78	0.94
1:A:1406[A]:ARG:NH1	5:A:1902:HOH:O	2.11	0.82
2:A:1801:S64:C15	2:A:1801:S64:C13	2.57	0.81
2:A:1801:S64:C15	2:A:1801:S64:S12	2.72	0.78
1:A:1406[A]:ARG:NH2	5:A:1904:HOH:O	2.17	0.77
1:B:1517[B]:ARG:HH11	1:B:1517[B]:ARG:CG	2.03	0.72
1:B:1385:LEU:HD12	1:B:1392:CYS:SG	2.31	0.71
1:A:1476:GLU:CB	5:A:2254:HOH:O	2.39	0.69
1:B:1517[B]:ARG:HH11	1:B:1517[B]:ARG:HG3	1.58	0.69
2:A:1801:S64:S12	2:A:1801:S64:C14	2.83	0.67
1:B:1517[B]:ARG:NH1	1:B:1517[B]:ARG:HG3	2.09	0.66
1:A:1711:LYS:CB	5:A:2127:HOH:O	2.43	0.66
1:B:1623:GLU:OE1	5:B:1901:HOH:O	2.14	0.64
1:A:1639:GLU:O	5:A:1903:HOH:O	2.14	0.63
1:A:1385:LEU:HD12	1:A:1392[B]:CYS:SG	2.38	0.63
1:B:1506:MET:CE	1:B:1538:PRO:HG3	2.28	0.63
2:A:1801:S64:C11	2:A:1801:S64:C14	2.68	0.62
1:B:1518:ASP:HB3	5:B:2139:HOH:O	1.98	0.62
1:B:1644:HIS:NE2	1:B:1649:ASP:OD2	2.33	0.61
2:A:1801:S64:C15	2:A:1801:S64:C10	2.73	0.61
2:A:1801:S64:C11	2:A:1801:S64:C07	2.74	0.59
1:B:1505:LEU:HD22	1:B:1700:VAL:HG11	1.83	0.58
1:B:1550:GLU:HG3	5:B:2046:HOH:O	2.03	0.58
1:B:1722:HIS:ND1	5:B:1904:HOH:O	2.32	0.58
1:A:1385:LEU:CD1	1:A:1392[B]:CYS:SG	2.96	0.54
1:B:1410:LYS:HE2	5:B:2185:HOH:O	2.06	0.54
1:A:1631:LYS:HE2	1:A:1666:TYR:OH	2.09	0.52
1:A:1505:LEU:HD22	1:A:1700:VAL:HG11	1.92	0.51
1:B:1494:PHE:CZ	1:B:1498:MET:HE3	2.47	0.50
1:B:1520:ARG:NE	5:B:1918:HOH:O	2.46	0.49
1:A:1407:GLU:OE2	5:A:1905:HOH:O	2.20	0.48
1:A:1640:ASN:ND2	5:A:1907:HOH:O	2.31	0.48
1:A:1540:MET:HE2	1:A:1698:ILE:HG23	1.96	0.47
1:B:1506:MET:HE3	1:B:1538:PRO:HD3	1.96	0.47
1:B:1506:MET:HE2	1:B:1538:PRO:HG3	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:1517[B]:ARG:HB2	1:B:1517[B]:ARG:HH11	1.80	0.47
2:A:1801:S64:C07	2:A:1801:S64:C09	2.87	0.47
1:B:1517[B]:ARG:CB	1:B:1517[B]:ARG:HH11	2.26	0.47
1:A:1582:HIS:O	1:A:1582:HIS:CG	2.68	0.46
1:A:1385:LEU:HG	1:A:1392[B]:CYS:SG	2.56	0.46
1:B:1603:ILE:HG13	1:B:1603:ILE:O	2.16	0.45
1:B:1488:TRP:HA	1:B:1489:PRO:C	2.37	0.43
1:B:1715:ARG:HD3	5:B:2195:HOH:O	2.17	0.43
1:A:1469:ILE:HD12	1:A:1579:GLU:HB2	2.01	0.43
1:B:1446:VAL:HB	1:B:1484:LYS:HG2	2.01	0.43
1:B:1506:MET:CE	1:B:1538:PRO:CG	2.97	0.42
1:B:1446:VAL:HG22	1:B:1453:ILE:HD13	2.01	0.42
1:A:1384:TRP:HB3	1:A:1388:GLY:HA2	2.02	0.41
1:B:1527:LEU:HA	1:B:1528:PRO:HD3	1.97	0.41
1:B:1587:LEU:HD13	1:B:1603:ILE:HD11	2.03	0.41
1:B:1506:MET:O	1:B:1516:LYS:HE2	2.21	0.40
1:B:1606:HIS:HB2	1:B:1608:GLU:OE2	2.22	0.40

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:2247:HOH:O	5:B:1983:HOH:O[2_657]	1.82	0.38
5:A:1952:HOH:O	5:A:2079:HOH:O[2_648]	2.11	0.09
5:A:2177:HOH:O	5:A:2249:HOH:O[2_648]	2.19	0.01

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	342/372 (92%)	334 (98%)	7 (2%)	1 (0%)	41 31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	342/372 (92%)	322 (94%)	20 (6%)	0	100 100
All	All	684/744 (92%)	656 (96%)	27 (4%)	1 (0%)	51 43

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1581	ALA

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	298/331 (90%)	292 (98%)	6 (2%)	55 51
1	B	302/331 (91%)	295 (98%)	7 (2%)	50 45
All	All	600/662 (91%)	587 (98%)	13 (2%)	52 47

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

Mol	Chain	Res	Type
1	A	1410	LYS
1	A	1577	ILE
1	A	1579	GLU
1	A	1584	GLU
1	A	1644	HIS
1	A	1717	THR
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 (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1640	ASN
1	A	1650	GLN
1	B	1656	GLN

4.3.3 RNA

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains

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

4.5 Carbohydrates

There are no carbohydrates in this entry.

4.6 Ligand geometry

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

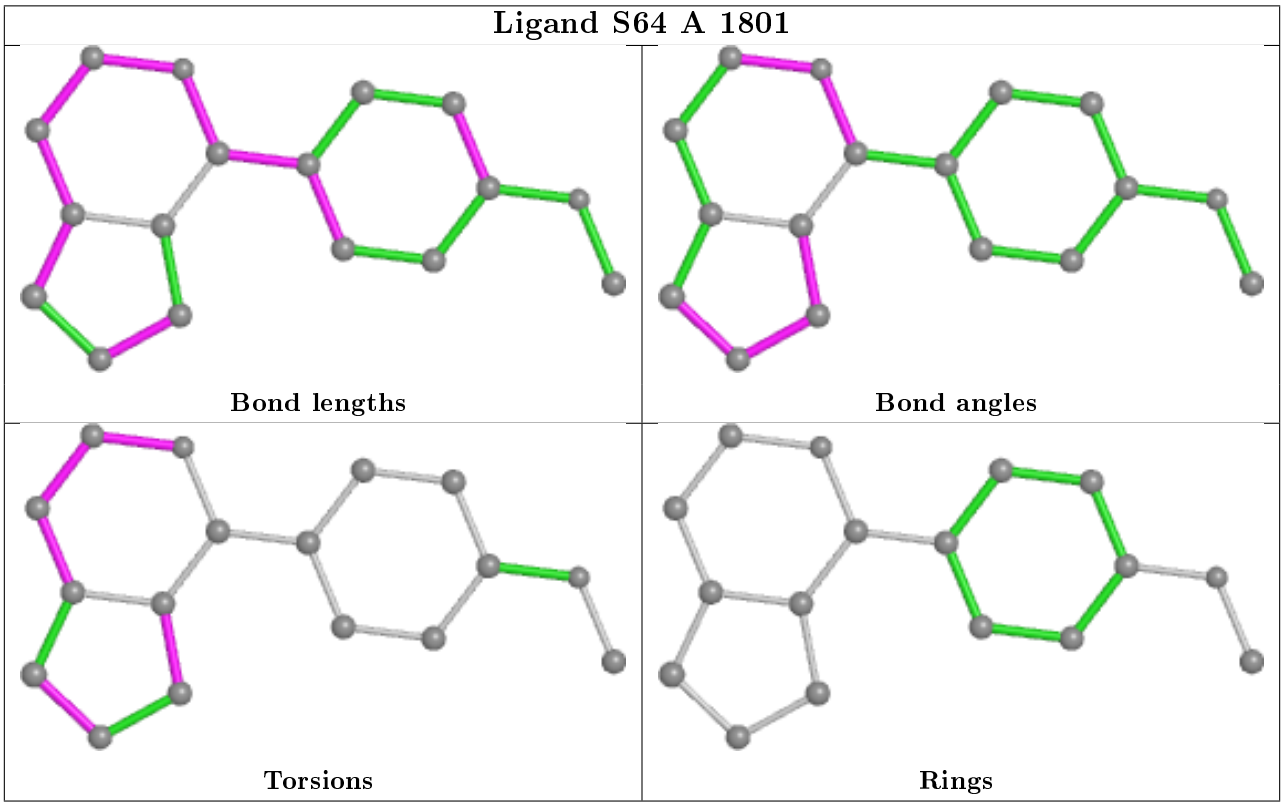
There are no torsion outliers.

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

There are no such residues in this entry.

4.8 Polymer linkage issues

There are no chain breaks in this entry.

5 Fit of model and data

5.1 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	340/372 (91%)	-0.01	26 (7%)	13	15	17, 29, 64, 101	24 (7%)
1	B	344/372 (92%)	-0.10	15 (4%)	34	37	20, 35, 71, 129	2 (0%)
All	All	684/744 (91%)	-0.05	41 (5%)	21	24	17, 32, 69, 129	26 (3%)

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1402	TRP	10.0
1	A	1405	PHE	9.2
1	A	1577	ILE	8.5
1	A	1578	GLY	8.0
1	B	1577	ILE	7.9
1	A	1580	GLY	7.9
1	A	1404	ILE	7.6
1	A	1392[A]	CYS	7.5
1	B	1517[A]	ARG	7.3
1	A	1564[A]	SER	7.2
1	A	1406[A]	ARG	7.1
1	A	1581	ALA	6.8
1	B	1605	ASP	5.7
1	B	1601	GLU	5.6
1	A	1401	ASN	5.3
1	A	1576	PRO	5.3
1	A	1582	HIS	5.3
1	B	1603	ILE	5.3
1	A	1579	GLU	4.9
1	A	1399	LYS	4.7
1	A	1400	ASN	4.6
1	B	1581	ALA	4.6
1	B	1604	HIS	4.5
1	A	1403	LYS	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1407	GLU	4.4
1	A	1588[A]	LYS	3.9
1	A	1398	ASN	3.7
1	A	1583	ASP	3.6
1	A	1395	ASP	3.4
1	B	1639	GLU	3.2
1	A	1642	PRO	3.1
1	B	1642	PRO	2.9
1	B	1640	ASN	2.7
1	B	1587	LEU	2.6
1	A	1584	GLU	2.6
1	A	1606	HIS	2.6
1	B	1643	ASP	2.5
1	B	1606	HIS	2.3
1	A	1639	GLU	2.3
1	B	1582	HIS	2.2
1	B	1664[A]	GLU	2.2

5.2 Non-standard residues in protein, DNA, RNA chains

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

5.3 Carbohydrates

There are no carbohydrates in this entry.

5.4 Ligands

LIGAND-RSR INFOmissingINFO

5.5 Other polymers

There are no such residues in this entry.