



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 01:54 PM BST

PDB ID : 4EZ4
Title : free KDM6B structure
Authors : Cheng, Z.J.; Patel, D.J.
Deposited on : 2012-05-02
Resolution : 2.99 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 : 2.13.1
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) : 2.13.1

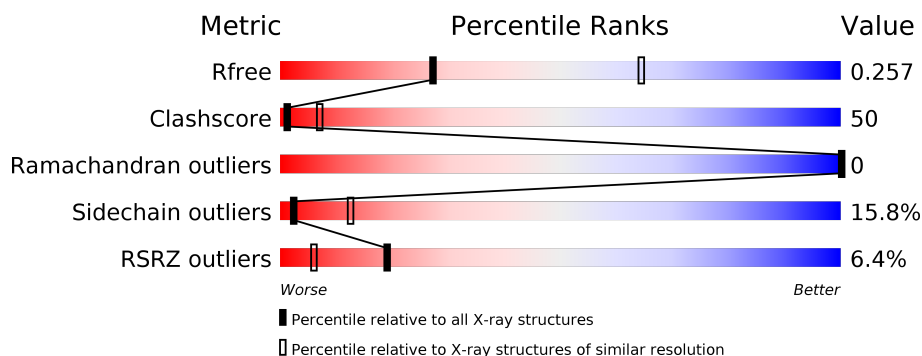
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.99 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 ≥ 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	486	
1	B	486	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7268 atoms, of which 32 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 6B.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	448	Total	C	H	N	O	S	0	1	0
			3620	2293	16	625	667	19			
1	B	447	Total	C	H	N	O	S	0	1	0
			3612	2290	16	627	660	19			

There are 52 discrepancies between the modelled and reference sequences:

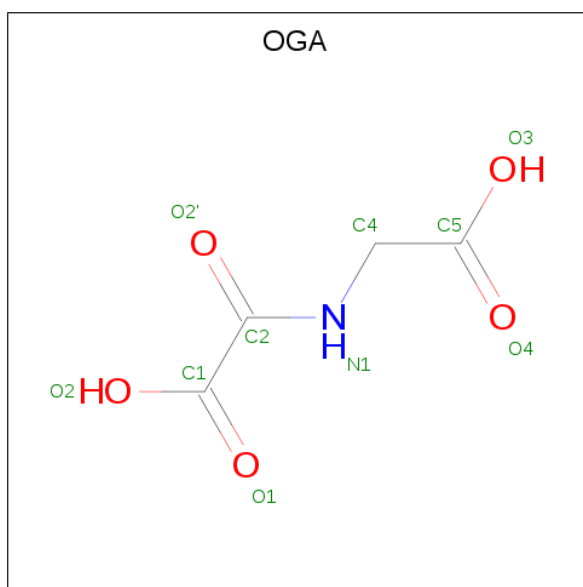
Chain	Residue	Modelled	Actual	Comment	Reference
A	1297	LEU	-	SEE REMARK 999	UNP Q5NCY0
A	1298	GLU	-	SEE REMARK 999	UNP Q5NCY0
A	1299	VAL	-	SEE REMARK 999	UNP Q5NCY0
A	1300	LEU	-	SEE REMARK 999	UNP Q5NCY0
A	1301	PHE	-	SEE REMARK 999	UNP Q5NCY0
A	1302	GLN	-	SEE REMARK 999	UNP Q5NCY0
A	1303	GLY	-	SEE REMARK 999	UNP Q5NCY0
A	1304	PRO	-	SEE REMARK 999	UNP Q5NCY0
A	1305	THR	-	SEE REMARK 999	UNP Q5NCY0
A	1306	LYS	-	SEE REMARK 999	UNP Q5NCY0
A	1307	ALA	-	SEE REMARK 999	UNP Q5NCY0
A	1308	ALA	-	SEE REMARK 999	UNP Q5NCY0
A	1309	ARG	-	SEE REMARK 999	UNP Q5NCY0
A	1310	LYS	-	SEE REMARK 999	UNP Q5NCY0
A	1311	SER	-	SEE REMARK 999	UNP Q5NCY0
A	1312	ALA	-	SEE REMARK 999	UNP Q5NCY0
A	1313	PRO	-	SEE REMARK 999	UNP Q5NCY0
A	1314	ALA	-	SEE REMARK 999	UNP Q5NCY0
A	1315	THR	-	SEE REMARK 999	UNP Q5NCY0
A	1316	GLY	-	SEE REMARK 999	UNP Q5NCY0
A	1317	GLY	-	SEE REMARK 999	UNP Q5NCY0
A	1318	GLY	-	SEE REMARK 999	UNP Q5NCY0
A	1319	SER	-	SEE REMARK 999	UNP Q5NCY0
A	1320	SER	-	SEE REMARK 999	UNP Q5NCY0
A	1321	GLY	-	SEE REMARK 999	UNP Q5NCY0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1322	SER	-	SEE REMARK 999	UNP Q5NCY0
B	1297	LEU	-	SEE REMARK 999	UNP Q5NCY0
B	1298	GLU	-	SEE REMARK 999	UNP Q5NCY0
B	1299	VAL	-	SEE REMARK 999	UNP Q5NCY0
B	1300	LEU	-	SEE REMARK 999	UNP Q5NCY0
B	1301	PHE	-	SEE REMARK 999	UNP Q5NCY0
B	1302	GLN	-	SEE REMARK 999	UNP Q5NCY0
B	1303	GLY	-	SEE REMARK 999	UNP Q5NCY0
B	1304	PRO	-	SEE REMARK 999	UNP Q5NCY0
B	1305	THR	-	SEE REMARK 999	UNP Q5NCY0
B	1306	LYS	-	SEE REMARK 999	UNP Q5NCY0
B	1307	ALA	-	SEE REMARK 999	UNP Q5NCY0
B	1308	ALA	-	SEE REMARK 999	UNP Q5NCY0
B	1309	ARG	-	SEE REMARK 999	UNP Q5NCY0
B	1310	LYS	-	SEE REMARK 999	UNP Q5NCY0
B	1311	SER	-	SEE REMARK 999	UNP Q5NCY0
B	1312	ALA	-	SEE REMARK 999	UNP Q5NCY0
B	1313	PRO	-	SEE REMARK 999	UNP Q5NCY0
B	1314	ALA	-	SEE REMARK 999	UNP Q5NCY0
B	1315	THR	-	SEE REMARK 999	UNP Q5NCY0
B	1316	GLY	-	SEE REMARK 999	UNP Q5NCY0
B	1317	GLY	-	SEE REMARK 999	UNP Q5NCY0
B	1318	GLY	-	SEE REMARK 999	UNP Q5NCY0
B	1319	SER	-	SEE REMARK 999	UNP Q5NCY0
B	1320	SER	-	SEE REMARK 999	UNP Q5NCY0
B	1321	GLY	-	SEE REMARK 999	UNP Q5NCY0
B	1322	SER	-	SEE REMARK 999	UNP Q5NCY0

- Molecule 2 is N-OXALYLGLYCINE (three-letter code: OGA) (formula: C₄H₅NO₅).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			10	4	1	5		
2	B	1	Total	C	N	O	0	0
			10	4	1	5		

- Molecule 3 is NICKEL (II) ION (three-letter code: NI) (formula: Ni).

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

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

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

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	5	Total	O	0	0
			5	5		

Continued on next page...

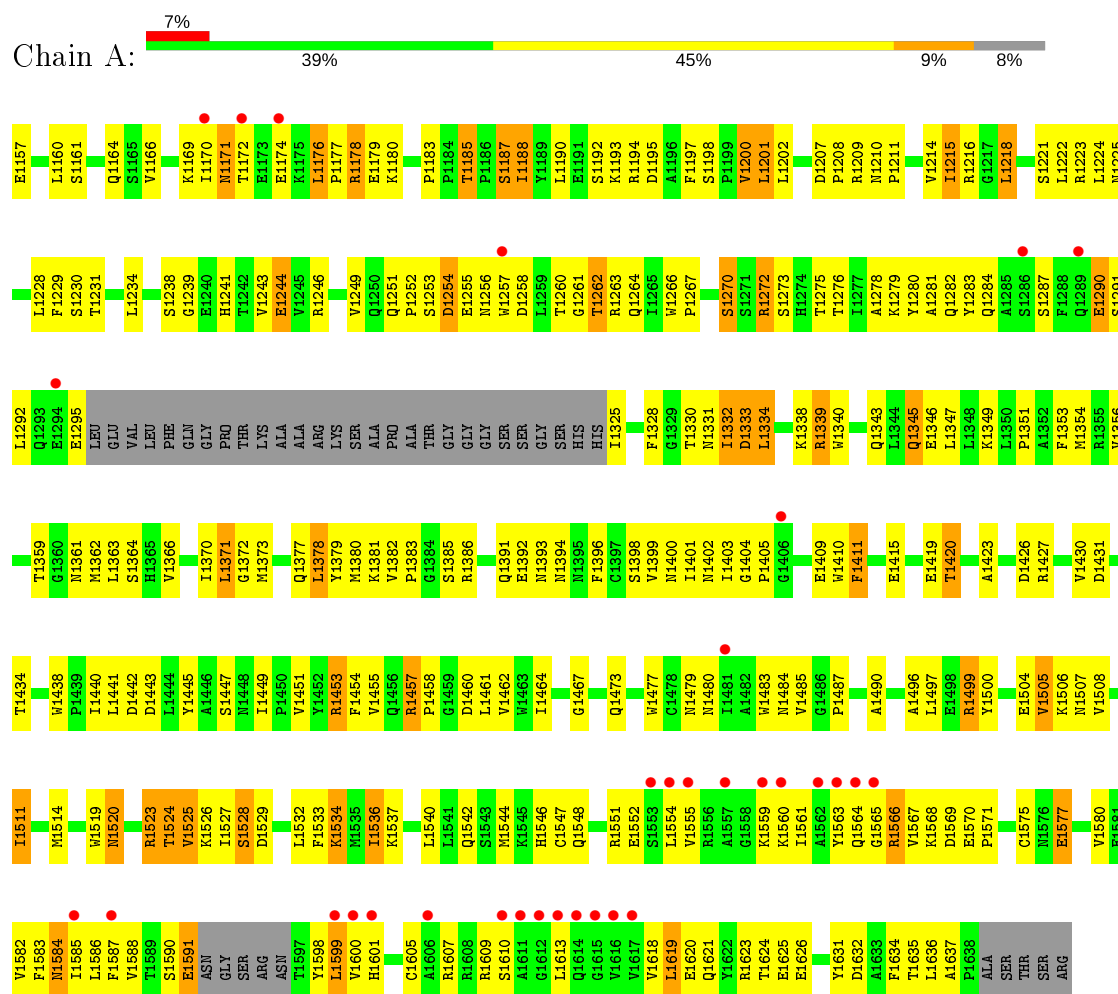
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	7	Total	O	0	0
			7	7		

3 Residue-property plots

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: Lysine-specific demethylase 6B



E1591	ASN	T1524	D1443	I1370	LEU	L1234
G1591	GLY	V1525	L1444	L1371	PHE	S1238
S1591	SER	K1526	Y1445	G1372	GLN	
ARG	ARG	I1527	A1446	M1373	GLY	
ASN	ASN	S1528	S1447	N1374	PRO	
T1597		D1529	N1448	T1375	THR	H1241
Y1598			I1449	V1376	LYS	T1242
L1599		L1532	F1450	Q1377	ALA	V1243
V1600		F1533	V1451	L1378	ALA	E1244
K1534		M1534	Y1452	Y1379	ARG	V1245
M1535		K1535	R1453	M1380	LYS	R1246
I1536		I1536	F1454	K1381	SER	
K1537		K1537	V1455	V1382	ALA	V1249
			Q1456	P1383	PRO	Q1250
			R1457	G1384	ALA	Q1251
L1540		L1540	D1460	S1385	THR	P1252
L1541		L1541	L1461	R1386	GLY	S1253
Q1542		Q1542	V1462		GLY	D1254
S1543		S1543	V1463	Q1391	GLY	E1255
M1544		M1544	I1464	E1392	GLY	N1256
K1545		K1545		M1393	SER	N1257
H1546		H1546	G1467	N1394	SER	D1258
C1547		C1547		M1395	GLY	L1259
Q1548		Q1548		F1396	SER	T1260
				C1397	G1261	G1261
R1551		R1551	V1472	S1398	T1262	T1262
E1552		E1552	Q1473			
S1553		S1553	A1474	V1399		R1263
L1554		L1554	T1475	M1400		Q1264
V1555		V1555	G1476	I1401		I1265
R1556		R1556	W1477	N1402		W1266
A1557		A1557		I1403		P1267
C1558		C1558	H1480	G1404		C1268
K1559		K1559	I1481	P1405		E1269
K1560		K1560		G1406		S1270
I1561		I1561	N1484			S1271
A1562		A1562	V1485	D1336		R1272
Y1563		Y1563	G1486	A1337		S1273
Q1564		Q1564	P1487	K1338		H1274
				R1339		T1275
G1565		G1565	Q1492	W1340		T1276
R1566		R1566		E1415		I1277
V1567		V1567				A1278
K1568		K1568	A1496	E1419		K1279
D1569		D1569	L1497	T1420		Y1280
E1570		E1570	E1498			A1281
P1571		P1571	R1499	A1423		Q1282
						Y1283
C1575		C1575	W1502	D1426		Q1284
N1576		N1576	N1503	R1427		
E1577		E1577	E1504	H1428		S1287
C1578		C1578	V1505	G1429		E1290
D1579		D1579	K1506	V1430		S1291
V1580		V1580	N1507	D1431		L1292
E1581		E1581	V1508			Q1293
V1582		V1582	K1509	T1434		E1294
			S1510			GLU
L1586		L1586	I1511	W1438		LEU
F1587		F1587	P1439	P1439		GLU
V1588		V1588	L1440	L1441		VAL
T1589		T1589	W1519			
S1590		S1590	R1523			

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	57.66Å 123.71Å 82.04Å 90.00° 109.62° 90.00°	Depositor
Resolution (Å)	49.73 – 2.99 49.73 – 2.99	Depositor EDS
% Data completeness (in resolution range)	95.6 (49.73-2.99) 95.7 (49.73-2.99)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	0.13	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.69 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.190 , 0.250 0.213 , 0.257	Depositor DCC
R_{free} test set	1079 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å ²)	55.5	Xtriage
Anisotropy	0.337	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 49.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.030 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	7268	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.76 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.9848e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NI, ZN, OGA

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	0.51	0/3700	0.69	1/5035 (0.0%)
1	B	0.52	1/3693 (0.0%)	0.70	1/5024 (0.0%)
All	All	0.52	1/7393 (0.0%)	0.69	2/10059 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	1361	ASN	CG-OD1	-5.02	1.12	1.24

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	1333	ASP	CB-CG-OD2	-5.97	112.93	118.30
1	A	1334	LEU	CA-CB-CG	-5.17	103.40	115.30

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	3604	16	3524	364	0
1	B	3596	16	3517	374	0
2	A	10	0	3	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	10	0	3	1	0
3	A	1	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	5	0	0	0	0
5	B	7	0	0	0	0
All	All	7236	32	7047	713	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 50.

The worst 5 of 713 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:1193:LYS:CE	1:B:1578:CYS:HA	1.53	1.38
1:B:1568:LYS:CA	1:B:1568:LYS:HE3	1.41	1.37
1:A:1215:ILE:O	1:A:1215:ILE:HD12	1.17	1.32
1:B:1215:ILE:HD12	1:B:1215:ILE:O	1.21	1.27
1:A:1339:ARG:HH11	1:A:1339:ARG:CB	1.46	1.26

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	443/486 (91%)	420 (95%)	23 (5%)	0	100	100
1	B	440/486 (90%)	421 (96%)	19 (4%)	0	100	100
All	All	883/972 (91%)	841 (95%)	42 (5%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	398/424 (94%)	334 (84%)	64 (16%)	2	12
1	B	396/424 (93%)	335 (85%)	61 (15%)	2	13
All	All	794/848 (94%)	669 (84%)	125 (16%)	2	13

5 of 125 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1577	GLU
1	B	1188	ILE
1	B	1564	GLN
1	A	1591	GLU
1	A	1635	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1584	ASN
1	A	1621	GLN
1	B	1473	GLN
1	A	1473	GLN
1	B	1507	ASN

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

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	$\# Z > 2$
2	OGA	A	1701	3	3,9,9	0.48	0	4,11,11	0.46	0
2	OGA	B	1701	3	3,9,9	0.43	0	4,11,11	0.61	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
2	OGA	A	1701	3	-	0/3/9/9	-
2	OGA	B	1701	3	-	0/3/9/9	-

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.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1701	OGA	2	0
2	B	1701	OGA	1	0

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	448/486 (92%)	0.31	33 (7%)	14 4	24, 58, 98, 141	0
1	B	447/486 (91%)	0.21	24 (5%)	25 9	26, 57, 100, 148	0
All	All	895/972 (92%)	0.26	57 (6%)	19 6	24, 57, 100, 148	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1613	LEU	7.5
1	B	1170	ILE	6.1
1	B	1613	LEU	6.0
1	A	1615	GLY	5.7
1	B	1323	HIS	4.9

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 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, 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(\AA^2)	Q<0.9
2	OGA	B	1701	10/10	0.92	0.22	20,20,20,20	0
2	OGA	A	1701	10/10	0.93	0.27	41,60,68,86	0
3	NI	A	1702	1/1	0.94	0.20	46,46,46,46	0
3	NI	B	1702	1/1	0.97	0.21	41,41,41,41	0
4	ZN	B	1703	1/1	0.97	0.14	73,73,73,73	0
4	ZN	A	1703	1/1	0.99	0.05	67,67,67,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.