



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 28, 2020 – 02:27 PM BST

PDB ID : 6VR4
Title : Virion-packaged DNA-dependent RNA polymerase of crAss-like phage phi14:2
Authors : Leiman, P.G.; Sokolova, M.L.
Deposited on : 2020-02-06
Resolution : 3.50 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
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

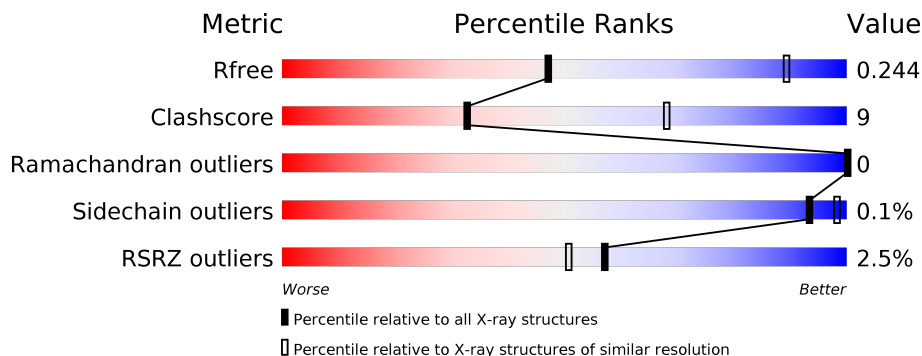
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	2194	 2% 78% 21%
1	B	2194	 3% 78% 21%

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
3	NA	A	2212	-	-	-	X
3	NA	B	2211	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 34715 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 DNA-dependent RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	N	O	S				Se
1	A	2166	17344	11014	2859	3433	1	37	0	0	0
1	B	2166	17344	11014	2859	3433	1	37	0	0	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MSE	-	expression tag	UNP S0A2C3
A	-12	GLY	-	expression tag	UNP S0A2C3
A	-11	SER	-	expression tag	UNP S0A2C3
A	-10	SER	-	expression tag	UNP S0A2C3
A	-9	HIS	-	expression tag	UNP S0A2C3
A	-8	HIS	-	expression tag	UNP S0A2C3
A	-7	HIS	-	expression tag	UNP S0A2C3
A	-6	HIS	-	expression tag	UNP S0A2C3
A	-5	HIS	-	expression tag	UNP S0A2C3
A	-4	HIS	-	expression tag	UNP S0A2C3
A	-3	SER	-	expression tag	UNP S0A2C3
A	-2	GLN	-	expression tag	UNP S0A2C3
A	-1	ASP	-	expression tag	UNP S0A2C3
A	0	PRO	-	expression tag	UNP S0A2C3
B	-13	MSE	-	expression tag	UNP S0A2C3
B	-12	GLY	-	expression tag	UNP S0A2C3
B	-11	SER	-	expression tag	UNP S0A2C3
B	-10	SER	-	expression tag	UNP S0A2C3
B	-9	HIS	-	expression tag	UNP S0A2C3
B	-8	HIS	-	expression tag	UNP S0A2C3
B	-7	HIS	-	expression tag	UNP S0A2C3
B	-6	HIS	-	expression tag	UNP S0A2C3
B	-5	HIS	-	expression tag	UNP S0A2C3
B	-4	HIS	-	expression tag	UNP S0A2C3
B	-3	SER	-	expression tag	UNP S0A2C3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	GLN	-	expression tag	UNP S0A2C3
B	-1	ASP	-	expression tag	UNP S0A2C3
B	0	PRO	-	expression tag	UNP S0A2C3

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	10	Total Cl 10 10	0	0
2	A	11	Total Cl 11 11	0	0

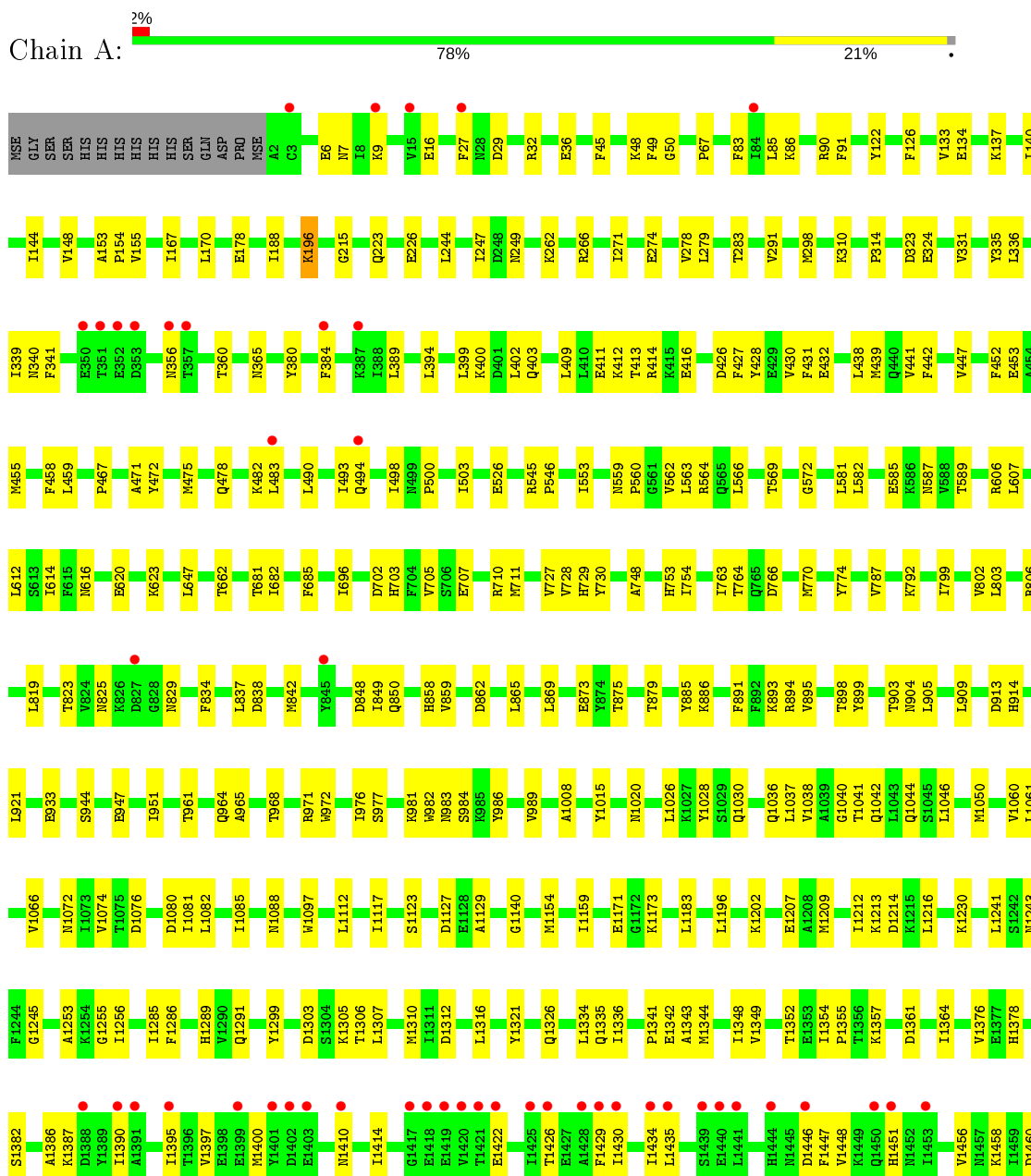
- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

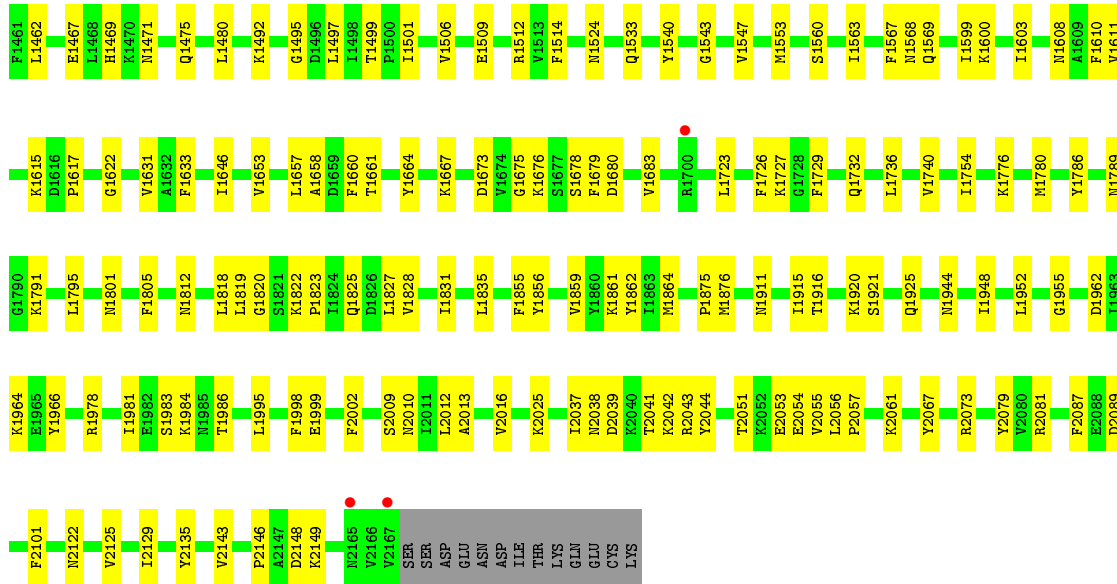
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Na 2 2	0	0
3	A	4	Total Na 4 4	0	0

3 Residue-property plots [i](#)

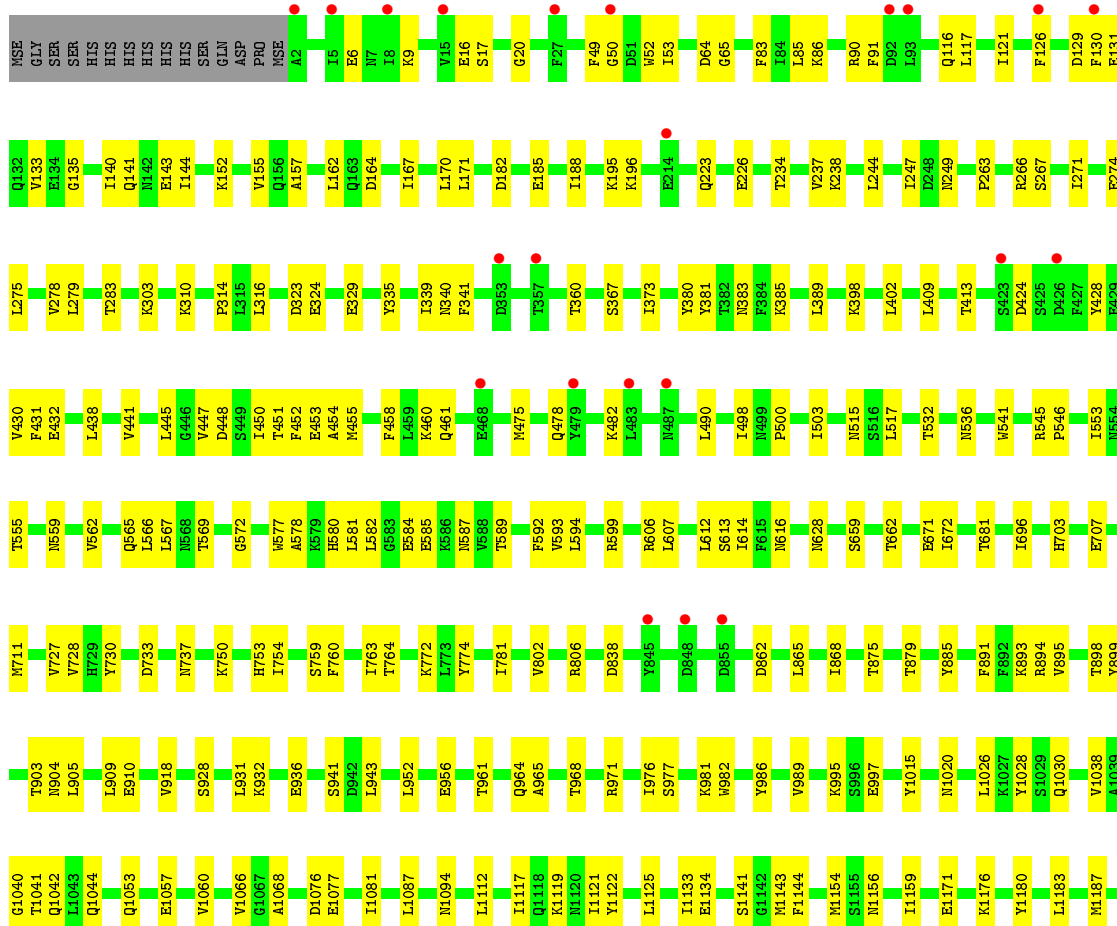
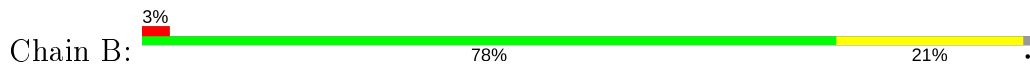
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DNA-dependent RNA polymerase





• Molecule 1: DNA-dependent RNA polymerase



Y2044	V1859	L1707	M1553	I1480	P1355	S1192
V2055	I1863	I1711	L1954	H1433	T1396	I1396
L2056	M1864	I1718	V1555	I1434	K1357	M1195
P2057	F1867	I1725	R1559	L1435	D1361	L1196
V2062	A1868	V1722	M1568	S1439	I1364	K1202
Y2067	P1869	L1723	L1572	E1440	M1367	I1206
V2079	D1893	T1725	T1582	L1441	M1370	E1207
D2089	M1911	F1726	T1586	H1444	M1370	M1209
F2101	I1915	F1729	K1586	M1445	L1371	M1209
L2118	T1916	L1736	I1599	V1448	T1379	I1212
R2122	K1920	V1740	I1603	H1451	K1380	L1216
V2125	S1921	Q1741	I1603	N1452	K1381	Y1217
I2136	F1924	L1742	L1607	I1453	S1382	K1230
V2143	Q1925	D1746	M1608	D1460	L1385	G1235
F2146	Y1929	I1754	A1609	M1471	A1386	M1243
K2149	Y1944	I1761	F1610	Q1475	K1387	M1243
S2155	I1948	I1764	V1611	L1480	I1390	A1253
L2156	Y1966	Y1768	K1615	L1483	A1391	L1259
L2159	I1981	Y1779	D1616	Y1484	E1394	I1285
L2160	E1982	M1780	P1617	Y1494	I1395	F1286
V2167	S1983	M1786	G1622	V1487	T1396	M1287
SER	K1984	Y1786	V1631	L1497	F1399	S1288
SER	M1985	L1795	A1632	I1498	M1400	H1289
ASP	T1986	L1799	F1633	T1499	F1401	V1290
G1U	E1989	K1800	I1637	T1499	D1402	Q1291
ASP	E2000	M1801	R1641	V1506	E1403	D1312
ASN	Q2001	F1805	H1641	V1509	F1405	L1319
ASP	F2002	I1809	P1642	E1509	D1406	I1319
ILE	S2009	L1812	D1643	R1512	H1407	G1320
THR	L2012	M1819	I1646	F1514	F1409	Y1321
LYS	V2016	L1818	I1646	G1522	M1410	Q1326
G1U	S2017	L1819	K1667	E1523	I1411	S1329
G1U	Q2018	G1820	S1678	M1524	I1414	L1334
CYS	I2021	L1827	F1679	L1525	A1415	L1334
LYS	I2027	V1828	D1680	K1526	N1416	Q1335
	V2032	I1831	I1681	D1529	G1417	I1339
	F2033	L1835	I1682	P1530	E1419	L1340
	M2038	F1855	V1683	T1420	F1420	F1341
	D2039		S1689	T1421	T1421	M1344
			F1702	L1422	E1422	I1348
			S1703	Y1540	A1428	V1349
				G1543	F1429	I1354

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	266.44Å 297.18Å 92.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.93 – 3.50 49.60 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.7 (31.93-3.50) 99.7 (49.60-3.50)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.51 (at 3.48Å)	Xtrriage
Refinement program	PHENIX 1.18.1_3865	Depositor
R, R_{free}	0.192 , 0.239 0.197 , 0.244	Depositor DCC
R_{free} test set	4647 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	96.4	Xtrriage
Anisotropy	0.039	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 54.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	34715	wwPDB-VP
Average B, all atoms (Å ²)	93.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: NA, 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	0.26	0/17630	0.44	0/23740
1	B	0.26	0/17630	0.45	0/23740
All	All	0.26	0/35260	0.45	0/47480

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 [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	17344	0	17141	301	0
1	B	17344	0	17141	298	0
2	A	11	0	0	0	0
2	B	10	0	0	0	0
3	A	4	0	0	0	0
3	B	2	0	0	0	0
All	All	34715	0	34282	597	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 9.

The worst 5 of 597 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:1286:PHE:HB2	1:B:1349:VAL:HB	1.51	0.91
1:A:707:GLU:OE1	1:A:710:ARG:NH2	2.12	0.81
1:A:1286:PHE:HB2	1:A:1349:VAL:HB	1.62	0.81
1:A:1015:TYR:HB3	1:A:1026:LEU:HB2	1.64	0.78
1:B:121:ILE:HG12	1:B:143:GLU:HG2	1.66	0.78

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	2164/2194 (99%)	2067 (96%)	97 (4%)	0	100	100
1	B	2164/2194 (99%)	2071 (96%)	93 (4%)	0	100	100
All	All	4328/4388 (99%)	4138 (96%)	190 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	1946/1934 (101%)	1944 (100%)	2 (0%)	93	98
1	B	1946/1934 (101%)	1946 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3892/3868 (101%)	3890 (100%)	2 (0%)	93 98

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	196	LYS
1	A	2025	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	825	ASN
1	A	850	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 27 ligands modelled in this entry, 27 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.

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	2129/2194 (97%)	-0.05	50 (2%) 60 54	52, 88, 155, 319	0
1	B	2129/2194 (97%)	-0.05	56 (2%) 56 49	48, 82, 155, 452	0
All	All	4258/4388 (97%)	-0.05	106 (2%) 57 51	48, 85, 155, 452	0

The worst 5 of 106 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1410	ASN	8.3
1	B	1419	GLU	7.7
1	B	1418	GLU	5.7
1	A	1421	THR	5.5
1	A	1418	GLU	5.3

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
3	NA	A	2212	1/1	0.71	0.89	68,68,68,68	0
2	CL	A	2204	1/1	0.71	0.35	83,83,83,83	0
3	NA	A	2213	1/1	0.73	0.22	54,54,54,54	0
2	CL	B	2201	1/1	0.77	0.33	65,65,65,65	0
3	NA	B	2211	1/1	0.77	0.71	70,70,70,70	0
2	CL	A	2205	1/1	0.78	0.23	95,95,95,95	0
2	CL	A	2207	1/1	0.81	0.34	80,80,80,80	0
2	CL	B	2209	1/1	0.82	0.33	81,81,81,81	0
2	CL	A	2211	1/1	0.83	0.88	106,106,106,106	0
2	CL	A	2210	1/1	0.83	0.13	94,94,94,94	0
3	NA	B	2212	1/1	0.84	0.15	64,64,64,64	0
2	CL	B	2204	1/1	0.85	0.24	74,74,74,74	0
2	CL	A	2206	1/1	0.87	0.13	79,79,79,79	0
2	CL	B	2203	1/1	0.87	0.21	73,73,73,73	0
2	CL	B	2202	1/1	0.87	0.30	73,73,73,73	0
2	CL	B	2206	1/1	0.88	0.17	59,59,59,59	0
2	CL	B	2207	1/1	0.89	0.49	94,94,94,94	0
2	CL	A	2208	1/1	0.89	0.33	80,80,80,80	0
3	NA	A	2215	1/1	0.89	0.60	91,91,91,91	0
2	CL	A	2209	1/1	0.89	0.40	78,78,78,78	0
2	CL	A	2201	1/1	0.91	0.17	67,67,67,67	0
2	CL	A	2203	1/1	0.91	0.35	72,72,72,72	0
2	CL	B	2208	1/1	0.92	0.18	80,80,80,80	0
2	CL	B	2210	1/1	0.92	0.55	81,81,81,81	0
2	CL	B	2205	1/1	0.93	0.35	76,76,76,76	0
3	NA	A	2214	1/1	0.93	0.32	49,49,49,49	0
2	CL	A	2202	1/1	0.94	0.41	96,96,96,96	0

6.5 Other polymers [i](#)

There are no such residues in this entry.