



wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 15, 2017 – 09:45 PM EST

PDB ID : 5URM
Title : Crystal structure of human BRR2 in complex with T-1206548
Authors : Klein, M.G.; Tjhen, R.; Qin, L.
Deposited on : unknown
Resolution : 2.80 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20030345
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

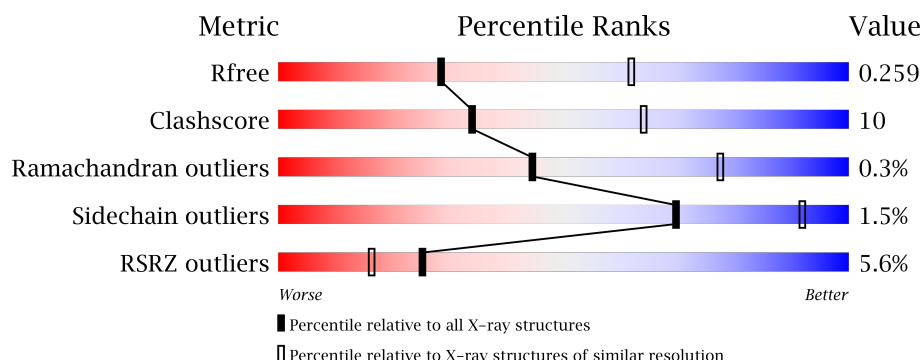
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	100719	2583 (2.80-2.80)
Clashscore	112137	3033 (2.80-2.80)
Ramachandran outliers	110173	2983 (2.80-2.80)
Sidechain outliers	110143	2985 (2.80-2.80)
RSRZ outliers	101464	2610 (2.80-2.80)

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. 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	1738	<div> <div>6%</div> <div> <div></div> <div>75%</div> <div>23%</div> <div>..</div> </div> </div>
1	B	1738	<div> <div>5%</div> <div> <div></div> <div>73%</div> <div>24%</div> <div>..</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 27571 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.

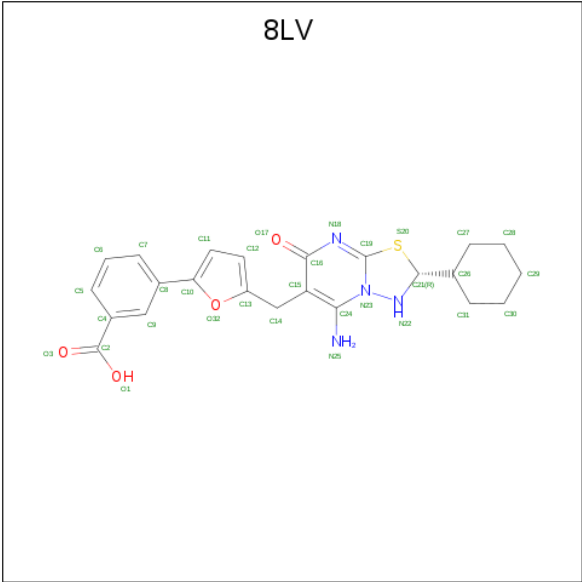
- Molecule 1 is a protein called U5 small nuclear ribonucleoprotein 200 kDa helicase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1718	Total	C	N	O	S	0	0	0
			13802	8821	2360	2550	71			
1	B	1691	Total	C	N	O	S	0	0	0
			13603	8705	2323	2504	71			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	392	GLY	-	expression tag	UNP O75643
A	393	GLY	-	expression tag	UNP O75643
A	394	SER	-	expression tag	UNP O75643
B	392	GLY	-	expression tag	UNP O75643
B	393	GLY	-	expression tag	UNP O75643
B	394	SER	-	expression tag	UNP O75643

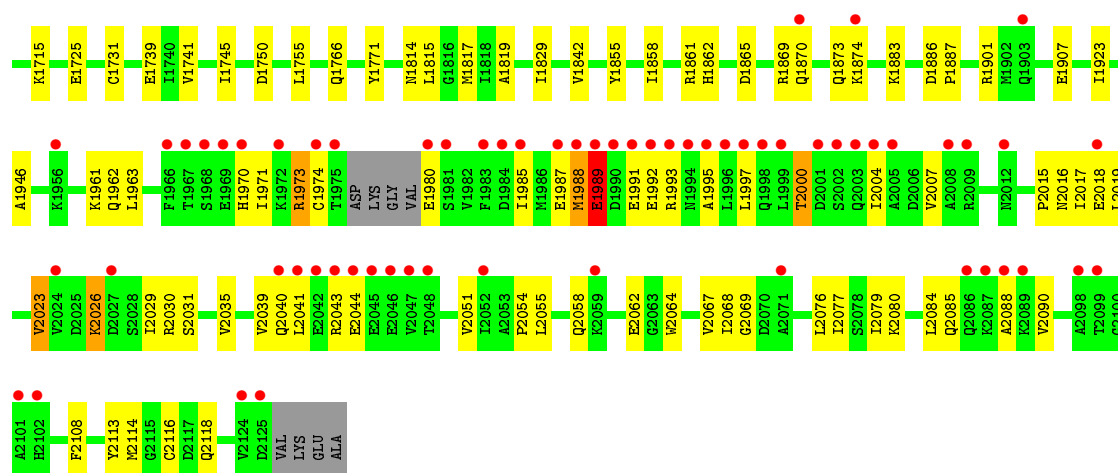
- Molecule 2 is 3-(5-([(2R)-5-amino-2-cyclohexyl-7-oxo-2,3-dihydro-7H-[1,3,4]thiadiazolo[3,2-a]pyrimidin-6-yl)methyl]furan-2-yl)benzoic acid (three-letter code: 8LV) (formula: C₂₃H₂₄N₄O₄S).



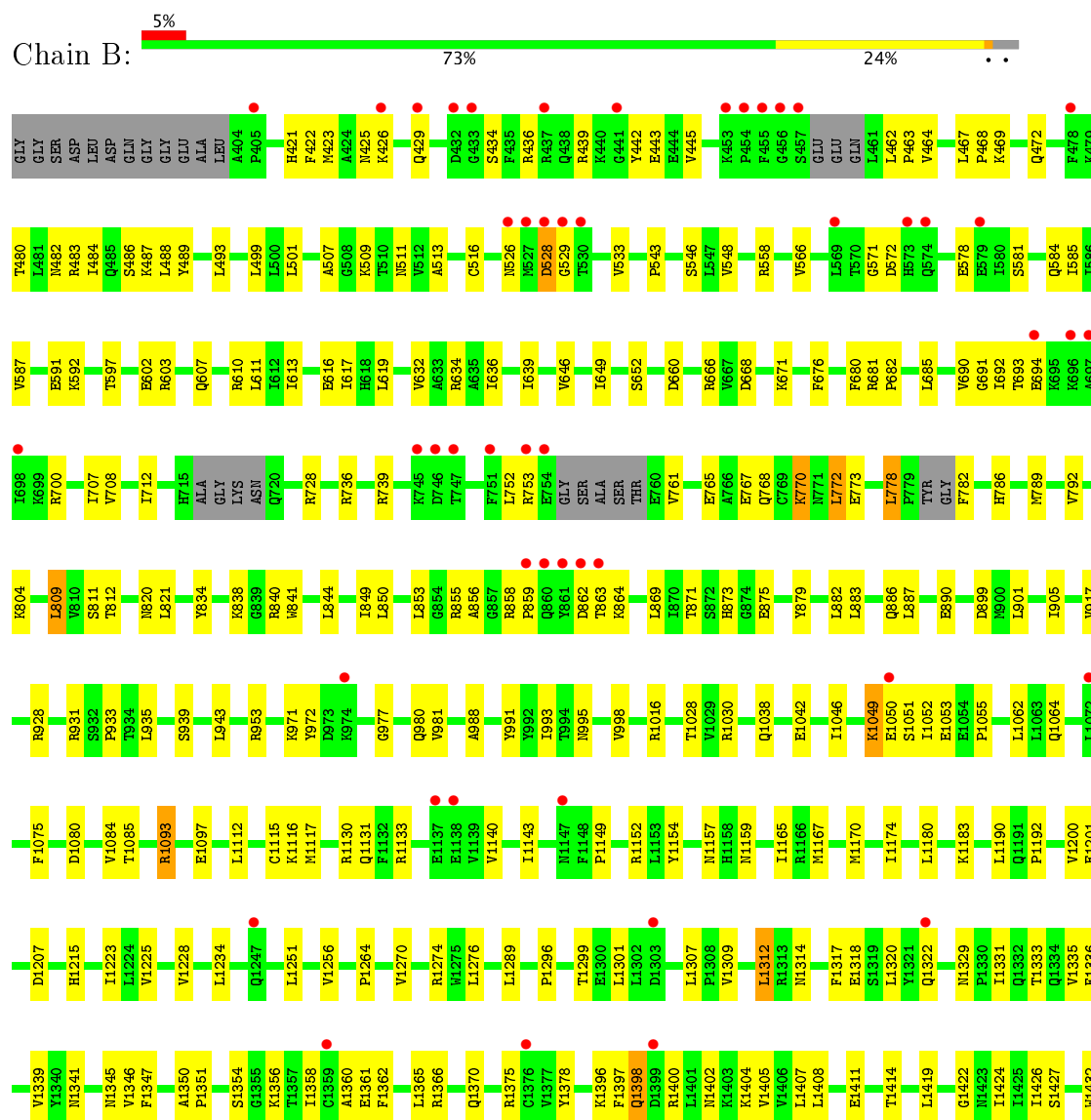
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	S	0	0
			32	23	4	4	1		
2	B	1	Total	C	N	O	S	0	0
			32	23	4	4	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	58	Total	O	0	0
			58	58		
3	B	44	Total	O	0	0
			44	44		



● Molecule 1: U5 small nuclear ribonucleoprotein 200 kDa helicase



A2098	I2017	S1777	I1584	D1433
T2099	E2018	M1953	Q1585	I1434
G2100	L2019	M1954	R1586	W1439
A2101	S2020	H1780	Q1587	M1444
H2102	Y2021	L1781	R1588	I1448
N2103	E2022	L1785	E1594	D1454
S2110	V2023	E1790	K1595	E1455
M2114	V2024	L1796	L1614	V1456
	K2026	H1965	N1615	H1457
F2122	V2035	T1802	H1621	L1458
S2123	V2036	M1814	E1628	R1473
V2124	V2037	L1815	V1632	M1474
D2125	L2038	G1816	V1656	I1481
VAL	V2039	I1817	L1660	E1482
LYS	Q2040	M1817	V1661	R1483
GLU	L2041	I1818	D1665	I1487
ALA	E2042	A1819	T1666	V1488
	R2043	ARG	A1691	L1490
	E2044	CYS	K1711	S1491
	E2045	THR	D1698	L1494
	E2046	ASP	E1699	S1495
	V2047	LYS	G1700	N1496
	T2048	GLY	A1701	A1497
	G2049	VAL	M1705	K1498
	P2050	GLU	M1705	A1501
	V2051	SER	K1711	C1506
	I2052	V1982	F1714	M1513
A2053	F1983	Y1955	K1715	H1514
P2054	D1984	R1861	L1718	P1515
L2055	I1985	H1862	H1735	H1516
	M1986	D1865	E1745	N1517
	E1987	P1876	M1747	V1518
	M1988	H1877	K1735	M1542
Q2058	D1989	P1876	E1739	V1546
K2059	E1990	H1877	I1745	H1553
	E1991	K1883	E1907	P1558
	R1993	D1886	M1921	V1559
	N1994	ALA	L1930	I1560
	LEU	LEU	S1937	R1566
	LEU	GLN	P1938	D1575
	L1999	L1999	L1945	I1576
	T2000	M1902	M1948	L1577
	D2001	Q1903		
	S2002	E1907		
	Q2003	R1921		
	T2004	L1930		
	A2005	S1937		
	D2006	P1938		
	V2007	L1945		
	A2008	M1948		
	C2011			
	N2012			
	R2013			
	N2016			
	K2091			
	A2096			
	P2097			

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	103.86Å 105.87Å 135.28Å 71.14° 70.48° 89.83°	Depositor
Resolution (Å)	29.48 – 2.80 29.48 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.2 (29.48-2.80) 87.0 (29.48-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.42 (at 2.80Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.212 , 0.260 0.211 , 0.259	Depositor DCC
R_{free} test set	6202 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	67.3	Xtriage
Anisotropy	0.430	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 56.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.017 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	27571	wwPDB-VP
Average B, all atoms (Å ²)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.64% 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: 8LV

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	0.32	0/14095	0.55	3/19100 (0.0%)
1	B	0.31	0/13889	0.55	5/18815 (0.0%)
All	All	0.31	0/27984	0.55	8/37915 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	6
1	B	0	7
All	All	0	13

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	529	GLY	N-CA-C	-9.25	89.98	113.10
1	A	772	LEU	CA-CB-CG	6.37	129.96	115.30
1	A	821	LEU	CA-CB-CG	6.35	129.91	115.30
1	A	1312	LEU	CA-CB-CG	6.34	129.87	115.30
1	B	809	LEU	CA-CB-CG	5.86	128.78	115.30

There are no chirality outliers.

5 of 13 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1049	LYS	Peptide

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	A	1988	MET	Peptide
1	A	1989	GLU	Peptide
1	A	694	GLU	Peptide
1	A	758	SER	Peptide

5.2 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	13802	0	13922	279	0
1	B	13603	0	13739	260	0
2	A	32	0	0	3	0
2	B	32	0	0	2	0
3	A	58	0	0	0	0
3	B	44	0	0	2	0
All	All	27571	0	27661	529	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 10.

The worst 5 of 529 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:2040:GLN:NE2	1:B:2088:ALA:O	1.99	0.95
1:A:742:CYS:HB3	1:A:749:GLY:HA2	1.53	0.89
1:B:572:ASP:OD1	1:B:1274:ARG:NH1	2.09	0.83
1:A:666:ARG:HH11	1:B:1595:LYS:HD3	1.44	0.82
1:A:421:HIS:NE2	1:A:875:GLU:OE2	2.15	0.79

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	1714/1738 (99%)	1633 (95%)	75 (4%)	6 (0%)	38	72
1	B	1677/1738 (96%)	1603 (96%)	70 (4%)	4 (0%)	51	83
All	All	3391/3476 (98%)	3236 (95%)	145 (4%)	10 (0%)	44	77

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1989	GLU
1	B	1050	GLU
1	A	1050	GLU
1	A	1665	ASP
1	B	1665	ASP

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	1534/1551 (99%)	1509 (98%)	25 (2%)	68	91
1	B	1514/1551 (98%)	1493 (99%)	21 (1%)	71	92
All	All	3048/3102 (98%)	3002 (98%)	46 (2%)	70	92

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1842	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	426	LYS
1	B	1699	GLU
1	A	1862	HIS
1	A	2000	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	532	ASN
1	A	2118	GLN
1	A	1690	HIS
1	A	526	ASN
1	A	2040	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](#)

2 ligands are modelled in this entry.

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	8LV	A	2501	-	25,36,36	1.75	4 (16%)	24,52,52	2.13	7 (29%)
2	8LV	B	2501	-	25,36,36	1.88	5 (20%)	24,52,52	1.94	5 (20%)

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	8LV	A	2501	-	-	0/5/32/32	0/3/5/5
2	8LV	B	2501	-	-	0/5/32/32	0/3/5/5

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2501	8LV	C24-C15	-6.60	1.34	1.42
2	A	2501	8LV	C24-C15	-6.34	1.34	1.42
2	B	2501	8LV	C14-C13	-3.81	1.48	1.51
2	A	2501	8LV	C14-C13	-2.68	1.49	1.51
2	B	2501	8LV	C14-C15	-2.46	1.49	1.52

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2501	8LV	C14-C15-C24	-3.08	118.32	122.28
2	B	2501	8LV	C7-C8-C10	-2.99	116.55	120.44
2	B	2501	8LV	C14-C15-C24	-2.63	118.90	122.28
2	A	2501	8LV	C7-C8-C10	-2.47	117.23	120.44
2	A	2501	8LV	C15-C24-N25	-2.05	119.16	120.64

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2501	8LV	3	0
2	B	2501	8LV	2	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	1718/1738 (98%)	0.05	99 (5%) 24 15	41, 71, 126, 176	0
1	B	1691/1738 (97%)	0.10	92 (5%) 26 17	46, 76, 122, 160	0
All	All	3409/3476 (98%)	0.07	191 (5%) 25 16	41, 74, 123, 176	0

The worst 5 of 191 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1975	THR	10.3
1	A	1994	ASN	7.5
1	B	1990	ASP	7.4
1	A	745	LYS	7.2
1	A	1995	ALA	6.7

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	8LV	B	2501	32/32	0.87	0.27	1.33	64,88,112,117	0
2	8LV	A	2501	32/32	0.90	0.21	0.55	58,75,81,83	0

6.5 Other polymers [i](#)

There are no such residues in this entry.