



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 04:31 PM GMT

PDB ID : 4F93  
Title : Brr2 Helicase Region S1087L, Mg-ATP  
Authors : Santos, K.F.; Jovin, S.M.; Weber, G.; Pena, V.; Luehrmann, R.; Wahl, M.C.  
Deposited on : 2012-05-18  
Resolution : 2.92 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

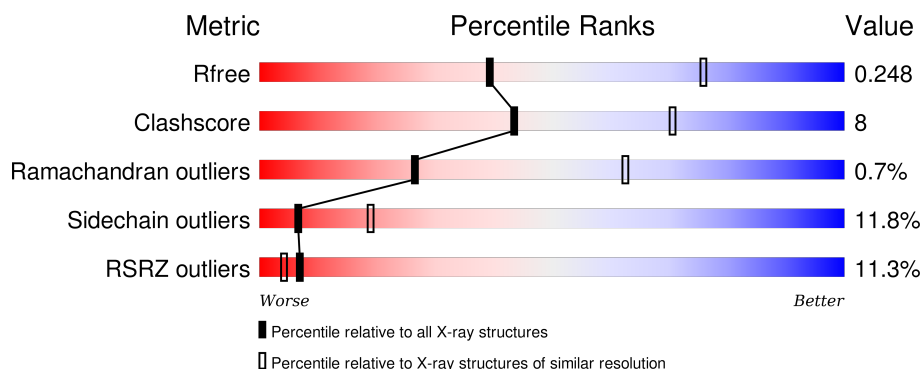
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.92 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))
$R_{free}$	91344	1643 (2.94-2.90)
Clashscore	102246	1871 (2.94-2.90)
Ramachandran outliers	100387	1824 (2.94-2.90)
Sidechain outliers	100360	1826 (2.94-2.90)
RSRZ outliers	91569	1650 (2.94-2.90)

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  $\geq 3$ , 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	B	1724	<div> <div>11%</div> <div>71%</div> <div>25%</div> <div>.</div> </div>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 13987 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
			Total	C	N	O	S			
1	B	1723	13866	8863	2373	2558	72	0	1	0

There is a discrepancy between the modelled and reference sequences:

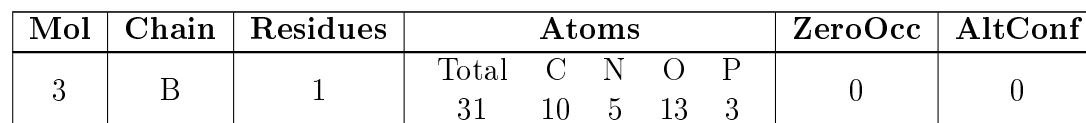
Chain	Residue	Modelled	Actual	Comment	Reference
B	1087	LEU	SER	ENGINEERED MUTATION	UNP O75643

- Molecule 2 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula:  $C_{10}H_{15}N_5O_{10}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
2	B	1	27	10	5	10	2	0	0

- Molecule 3 is ADENOSINE-5'-TRIPHOSPHATE (three-letter code: ATP) (formula:  $C_{10}H_{16}N_5O_{13}P_3$ ).



- | Mol | Chain | Residues | Atoms           | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 4   | B     | 1        | Total Mg<br>1 1 | 0       | 0       |

- SAN
- 
- NS(=O)(=O)c1ccc(N)cc1

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	B	1	Total	C	N	O	S	0	0
			11	6	2	2	1		

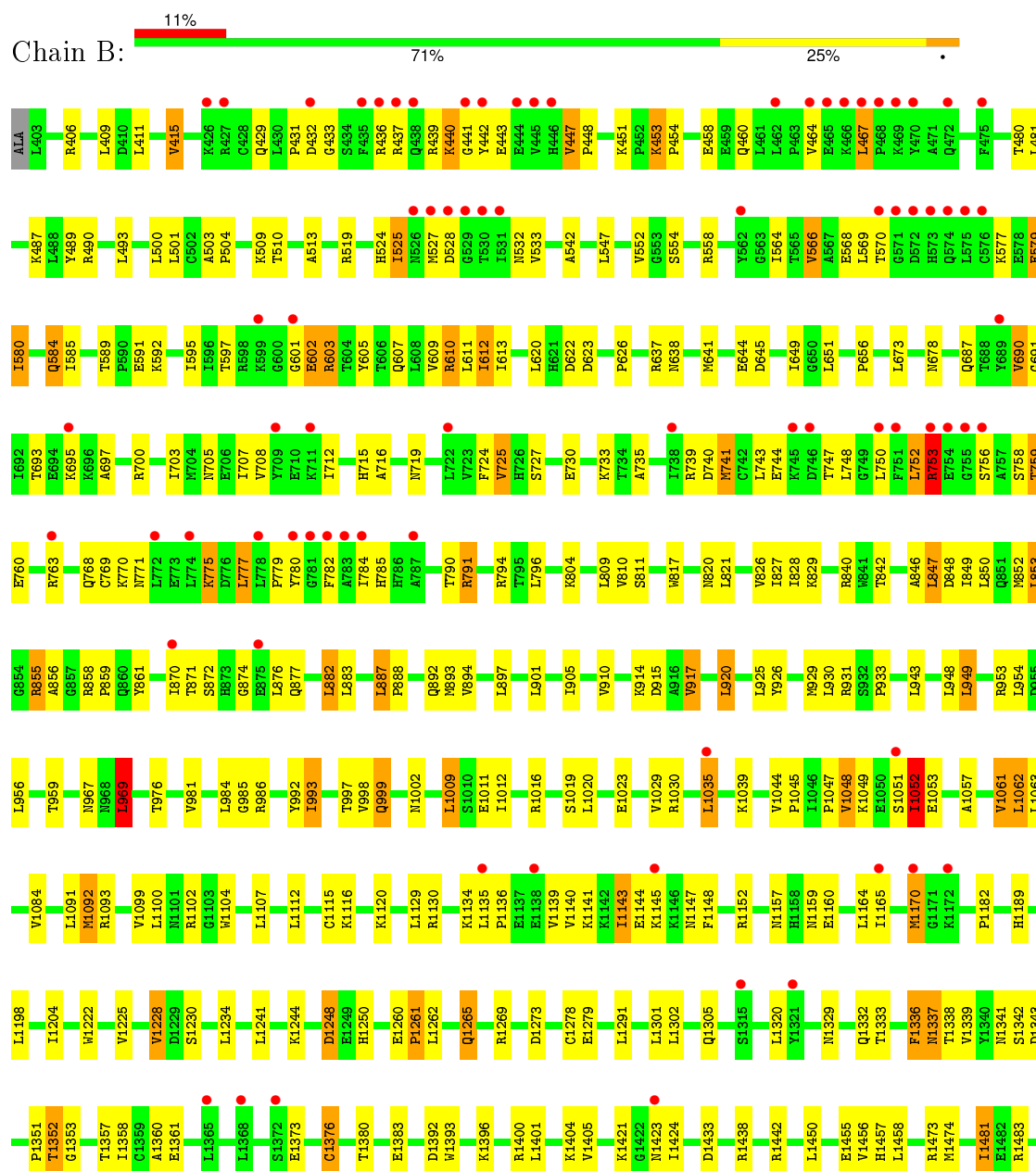
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	51	Total	O	0	0
			51	51		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: U5 small nuclear ribonucleoprotein 200 kDa helicase



K2091	L2092	D2093	F2094	A2098	T2099	G2100	A2101	H2102	N2103	Y2104	T2105	L2106	Y2107	F2108	A2112	Y2113	N2114	G2115	C2116	D2117	Q2118	E2119	Y2120	K2121	F2122	S2123	V2124	D2125	P2015	N2016	I2017	E2018	L2019	S2020	Y2021	E2022	V2023	V2024	D2025	K2026	D2027	S2028	I2029	G2032	G2033	P2034	V2035	V2036	V2037	L2038	V2039	Q2040	L2041	E2042	R2043	E2044	E2045	E2046	V2047	I2052	L2055	F2056	P2057	R2060	V2066	V2067	I2068	K2072	L2076	I2079	L2082	T2083	L2084	Q2085	Q2086	K2087	A2088	K2089	V2090	M1948	V1949	T1950	M1953	V1954	S1955	K1956	Y1959	L1960	K1961	Q1962	L1963	P1964	H1965	F1966	T1967	S1968	E1969	H1970	I1971	K1972	R1973	C1974	T1975	D1976	K1977	G1978	V1979	E1980	S1981	V1982	F1983	D1984	I1985	M1986	E1987	M1988	E1989	D1990	E1991	E1992	E1993	M1994	A1995	L1996	L1997	D1998	L1999	D2006	V2007	A2008	Q2009	F2010	C2011	R2012	R2013	V2014	V1842	L1845	I1846	I1849	A1852	A1853	E1854	Y1855	I1858	F1859	R1860	R1861	L1868	R1869	Q1870	L1871	A1872	Q1873	K1878	L1879	P1882	K1883	F1884	P1887	H1888	T1891	H1892	L1893	L1894	L1895	L1904	S1905	A1906	E1907	L1908	D1911	A1919	L1922	D1928	L1936	S1937	P1938	E1944	L1945	Q1947	H1730	C1731	D1734	E1739	I1740	T1744	K1748	Q1749	D1753	T1756	R1763	Y1771	M1772	L1773	Q1774	G1775	I1776	S1777	H1778	H1779	H1780	H1784	L1785	L1796	E1797	K1800	C1801	I1802	S1803	I1804	M1808	M1814	I1818	A1819	I1824	M1825	I1829	E1830	L1831	M1834	T1840	K1841	I1598	K1603	L1604	S1607	T1608	L1609	K1610	E1611	L1614	M1627	F1636	S1637	V1644	S1649	V1656	A1657	A1658	D1665	K1672	Y1676	D1683	V1684	L1685	Q1686	M1687	V1688	G1689	M1692	R1693	P1694	L1695	Q1696	E1699	G1700	R1701	K1715	L1718	F1589	L1722	E1725	E1594	K1595	R1486	A1489	L1490	S1491	S1492	S1493	L1494	K1498	A1501	C1506	T1509	P1516	M1517	V1518	R1519	P1520	V1521	I1527	T1535	Q1536	T1537	R1538	L1539	M1542	P1545	H1553	I1560	T1569	L1577	C1580	A1581	A1582	D1583	I1584	Q1585	R1586	Q1587	R1588	L1590	T1593	E1594	K1595
-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	147.12Å 154.59Å 143.28Å 90.00° 120.62° 90.00°	Depositor
Resolution (Å)	47.36 – 2.92 48.20 – 2.92	Depositor EDS
% Data completeness (in resolution range)	98.1 (47.36-2.92) 98.2 (48.20-2.92)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.98 (at 2.91Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.197 , 0.248 0.195 , 0.248	Depositor DCC
$R_{free}$ test set	2925 reflections (5.26%)	DCC
Wilson B-factor (Å <sup>2</sup> )	88.6	Xtriage
Anisotropy	0.476	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 88.2	EDS
Estimated twinning fraction	0.000 for k,h,-1/2*h-1/2*k-l 0.000 for -k,-h,-1/2*h+1/2*k-l 0.003 for h,-k,-h-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	0 of 58516 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13987	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	134.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.44% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, ATP, ADP, SAN

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	B	0.31	0/14161	0.52	3/19188 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	969	LEU	CA-CB-CG	6.01	129.13	115.30
1	B	1609	LEU	CA-CB-CG	5.41	127.75	115.30
1	B	753	ARG	NE-CZ-NH2	5.12	122.86	120.30

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	B	13866	0	14008	226	0
2	B	27	0	12	2	0
3	B	31	0	12	0	0
4	B	1	0	0	0	0
5	B	11	0	8	0	0
6	B	51	0	0	8	0
All	All	13987	0	14040	226	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 8.

The worst 5 of 226 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:993:ILE:HD11	1:B:998:VAL:HG23	1.60	0.84
1:B:1739:GLU:HG3	1:B:1744:THR:HB	1.61	0.82
1:B:1052:ILE:HG23	1:B:1053:GLU:H	1.56	0.70
1:B:1099:VAL:HG23	1:B:1104:TRP:HE3	1.57	0.70
1:B:933:PRO:HG3	1:B:943:LEU:HD22	1.74	0.69

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	B	1722/1724 (100%)	1603 (93%)	107 (6%)	12 (1%)	26 62

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	756	SER
1	B	1049	LYS
1	B	758	SER
1	B	1052	ILE
1	B	1584	ILE

### 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	B	1543/1542 (100%)	1361 (88%)	182 (12%)	6 19

5 of 182 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1052	ILE
1	B	1204	ILE
1	B	1878	LYS
1	B	1062	LEU
1	B	1129	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	820	ASN
1	B	1655	ASN
1	B	1862	HIS

### 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](#)

Of 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	ADP	B	3001	-	22,29,29	1.06	2 (9%)	27,45,45	1.87	4 (14%)
3	ATP	B	3002	4	24,33,33	1.00	2 (8%)	31,52,52	1.96	4 (12%)
5	SAN	B	3004	-	11,11,11	3.30	1 (9%)	16,16,16	2.35	3 (18%)

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	ADP	B	3001	-	-	0/12/32/32	0/3/3/3
3	ATP	B	3002	4	-	0/18/38/38	0/3/3/3
5	SAN	B	3004	-	-	0/6/6/6	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	3004	SAN	C4-S	-10.81	1.60	1.77
3	B	3002	ATP	O4'-C1'	2.07	1.43	1.41
2	B	3001	ADP	O4'-C1'	2.31	1.44	1.41
3	B	3002	ATP	C5-C4	3.00	1.47	1.40
2	B	3001	ADP	C5-C4	3.15	1.47	1.40

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	3004	SAN	O2-S-O1	-8.02	107.53	118.80
3	B	3002	ATP	N3-C2-N1	-7.62	123.06	128.89
2	B	3001	ADP	N3-C2-N1	-6.93	123.59	128.89
3	B	3002	ATP	PA-O3A-PB	-3.48	122.97	132.73
2	B	3001	ADP	C4-C5-N7	-3.32	106.43	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	3001	ADP	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	B	1723/1724 (99%)	0.57	194 (11%) <b>7</b> <b>4</b>	66, 127, 213, 267	0

The worst 5 of 194 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	2084	LEU	8.6
1	B	1963	LEU	8.1
1	B	2019	LEU	7.5
1	B	1985	ILE	7.4
1	B	751	PHE	7.2

### 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, 95<sup>th</sup> 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(Å <sup>2</sup> )	Q<0.9
2	ADP	B	3001	27/27	0.90	0.24	0.56	116,145,168,183	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
5	SAN	B	3004	11/11	0.95	0.20	0.05	96,102,121,132	0
3	ATP	B	3002	31/31	0.93	0.20	-0.09	87,120,199,217	0
4	MG	B	3003	1/1	0.89	0.56	-	164,164,164,164	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.