



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 08:13 PM GMT

PDB ID : 4R7Y
Title : Crystal structure of an active MCM hexamer
Authors : Miller, J.M.; Arachea, B.T.; Epling, L.B.; Enemark, E.J.
Deposited on : 2014-08-28
Resolution : 2.70 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 (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

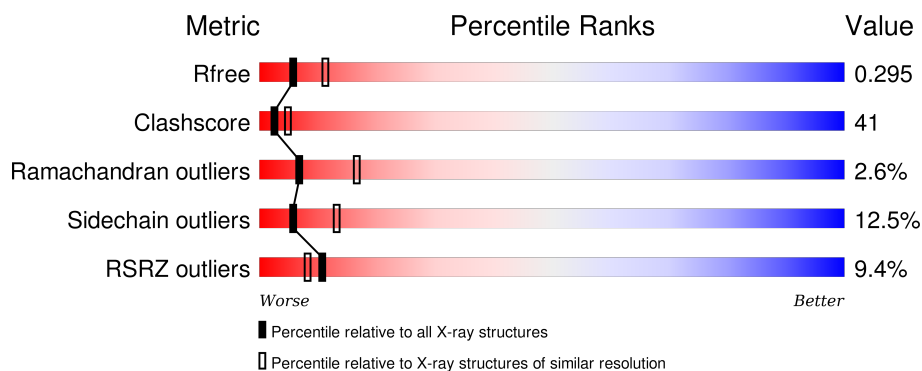
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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	613	<div> <div>10%</div> <div> <div>44%</div> <div>43%</div> <div>9%</div> <div>••</div> </div> </div>
1	B	613	<div> <div>8%</div> <div> <div>44%</div> <div>44%</div> <div>8%</div> <div>•</div> </div> </div>

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
5	CL	A	2006	-	-	X	-
5	CL	B	2006	-	-	X	-

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 9494 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.

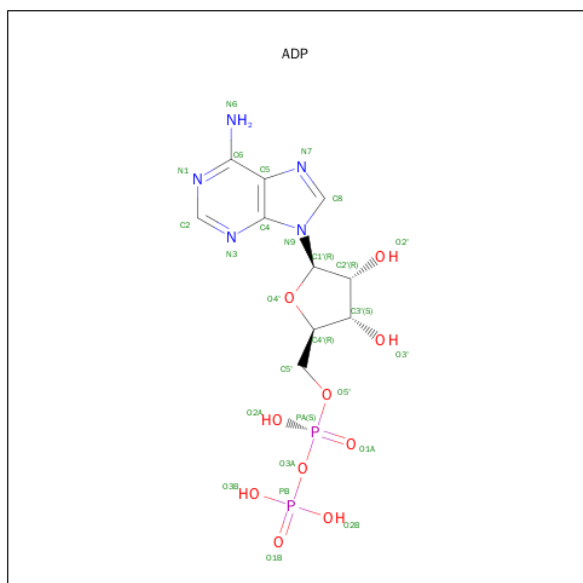
- Molecule 1 is a protein called Minichromosome maintenance protein MCM, Cell division control protein 21.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	591	Total	C	N	O	S	0	0	0
			4715	3001	822	878	14			
1	B	591	Total	C	N	O	S	0	0	0
			4715	3001	822	878	14			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	EXPRESSION TAG	UNP Q9UXG1
A	1	LEU	-	EXPRESSION TAG	UNP Q9UXG1
B	0	SER	-	EXPRESSION TAG	UNP Q9UXG1
B	1	LEU	-	EXPRESSION TAG	UNP Q9UXG1

- 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
2	A	1	Total	C	N	O	P	0	0
			27	10	5	10	2		
2	B	1	Total	C	N	O	P	0	0
			27	10	5	10	2		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

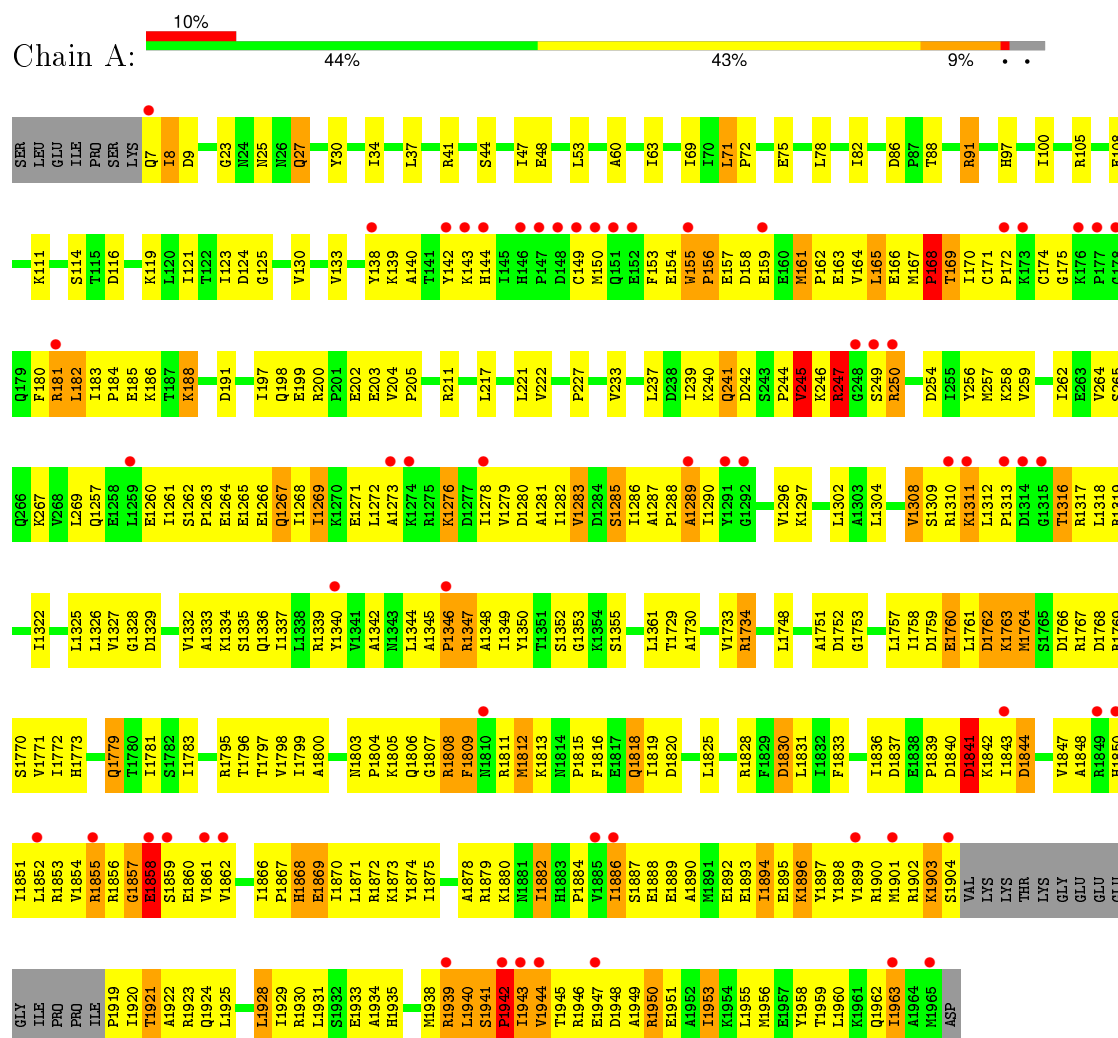
- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	3	Total	Cl	0	0
			3	3		
5	A	3	Total	Cl	0	0
			3	3		

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 ($\text{RSRZ} > 2$). 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: Minichromosome maintenance protein MCM, Cell division control protein 21



GLU	L1852	H1773	R1319	S260	K173	I103
GLY	R1853	E1774	G1320	S261	C174	I104
ILE	V1854	A1775	D1321	I262	G175	R105
PRO	R1855		I1322	E263	K176	
PRO	R1856	Q1779	L1325	K267	P177	E108
ILE	G1857	T1780	L1326		G178	L109
P1919	E1858	I1781	L1327	I1261	Q179	R110
I1920	S1859	S1782	V1327		F180	K111
T1921	E1860	I1783	G1328	E1264	R181	I112
A1922	V1861	T1789	D1329	E1265	L182	R113
R1923	V1862		V1332	E1266		
Q1924	A1863	T1796	A1333	Q1267	K186	K119
L1925	P1864		K1334	I1268	T187	L120
E1926	K1865	A1800	S1335	I1269	K188	I121
A1927	L1866		Q1336	K1270		T122
L1928	P1867	N1803	I1337	E1271	Q193	I123
R1929	H1868	P1804		L1272	I123	D124
R1930	E1869	K1805	Y1340	A1273	R200	V130
L1931	L1870		V1341	K1274	P201	T131
S1932	L1871	R1808	A1342	L1275	E202	P132
E1933	R1872	F1809	N1343	K1276	E203	
A1934	K1873	N1810	L1344	D1277	V204	
H1935	Y1874	M1811	A1345	I1278	P205	
A1936	L1875	M1812	P1346	I1279	E135	E136
R1937		K1813	R1347	D1280	I137	I138
M1938	A1878	N1814	A1348	A1287	E214	Y142
R1939	R1879	P1815	I1349	P1288	I215	K143
L1940	K1880	F1816		A1289	I216	H144
S1941	N1881	F1817	S1352	I1290	L217	I145
P1942	I1882	Q1818	G1353	V1733	V222	H146
I1943	H1883	I1819		R1734	P147	P148
T1944	P1884			D1735	D148	C149
V1945	V1885			G1746	M150	M150
R1946	L1886	P1823	T1729	A1747	Q151	Q151
E1947	S1887	T1824	V1733	L1748	F152	F152
P1948	E1888	L1825	R1734		F153	F153
A1949	E1889	L1826	D1735		P210	K139
R1950	A1890	S1827			R211	A140
E1951	M1891	R1828			E214	T141
A1952	E1892	F1829			K1287	Y142
I1953	E1893	D1830			K240	K143
K1954	I1894				Q241	H144
L1955	E1895	F1833			D242	E154
	K1896	I1836			S243	P156
I1963	Y1897	D1837			P244	E157
A1964	Y1898	E1838			V245	D158
M1965	V1899	P1839			K246	E159
ASP	R1900	D1840			R247	E160
	M1901	D1841			G248	M161
	R1902	K1842			S249	P162
	K1903	I1843			R250	E163
	S1904	D1844				
	VAL	S1845				E166
	LYS	V1847				M167
	LYS	A1848				P168
	THR	R1849				T169
	LYS	H1850				I170
	GLY	I1851				C171
	GLU					P172

4 Data and refinement statistics

Property	Value	Source
Space group	P 63	Depositor
Cell constants a, b, c, α , β , γ	118.90Å 118.90Å 199.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.83 – 2.70 49.83 – 2.70	Depositor EDS
% Data completeness (in resolution range)	89.7 (49.83-2.70) 89.7 (49.83-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.24 (at 2.69Å)	Xtriage
Refinement program	CNS 1.3	Depositor
R, R_{free}	0.263 , 0.295 0.263 , 0.295	Depositor DCC
R_{free} test set	1976 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	39.4	Xtriage
Anisotropy	0.248	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 44.8	EDS
Estimated twinning fraction	0.044 for h,-h-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 39044 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9494	wwPDB-VP
Average B, all atoms (Å ²)	61.0	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 49.42 % 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 7.4816e-05. 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, MG, ADP, 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	0.45	0/4793	0.69	0/6476
1	B	0.45	0/4793	0.70	0/6476
All	All	0.45	0/9586	0.70	0/12952

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

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	4715	0	4876	420	0
1	B	4715	0	4876	393	0
2	A	27	0	12	3	0
2	B	27	0	12	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	3	0	0	2	0
5	B	3	0	0	2	0
All	All	9494	0	9776	798	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 41.

The worst 5 of 798 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:1346:PRO:HB3	1:B:1872:ARG:HD3	1.29	1.10
1:A:156:PRO:HB2	1:A:159:GLU:O	1.53	1.07
1:B:1261:ILE:HA	1:B:1873:LYS:HZ1	1.17	1.05
1:A:1878:ALA:HA	1:A:1882:ILE:HD11	1.37	1.05
1:B:1261:ILE:HA	1:B:1873:LYS:NZ	1.73	1.04

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	587/613 (96%)	519 (88%)	53 (9%)	15 (3%)	7	16
1	B	587/613 (96%)	524 (89%)	47 (8%)	16 (3%)	6	16
All	All	1174/1226 (96%)	1043 (89%)	100 (8%)	31 (3%)	7	16

5 of 31 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	155	TRP
1	A	247	ARG
1	A	1841	ASP
1	A	1857	GLY
1	B	155	TRP

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	517/537 (96%)	444 (86%)	73 (14%)	4	10
1	B	517/537 (96%)	461 (89%)	56 (11%)	8	18
All	All	1034/1074 (96%)	905 (88%)	129 (12%)	6	13

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1892	GLU
1	A	1963	ILE
1	B	1879	ARG
1	A	1896	LYS
1	A	1940	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1814	ASN
1	A	1962	GLN
1	B	1810	ASN
1	A	1779	GLN
1	B	1779	GLN

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 10 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	ADP	A	2001	3	22,29,29	1.36	2 (9%)	27,45,45	2.45	2 (7%)
2	ADP	B	2001	3	22,29,29	1.51	3 (13%)	27,45,45	2.30	3 (11%)

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	A	2001	3	-	0/12/32/32	0/3/3/3
2	ADP	B	2001	3	-	0/12/32/32	0/3/3/3

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	2001	ADP	C8-N7	-2.56	1.29	1.34
2	A	2001	ADP	PB-O1B	2.88	1.60	1.51
2	B	2001	ADP	PB-O1B	3.05	1.61	1.51
2	A	2001	ADP	O4'-C1'	3.41	1.45	1.41
2	B	2001	ADP	O4'-C1'	3.96	1.46	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	ADP	N3-C2-N1	-9.68	121.48	128.89
2	B	2001	ADP	N3-C2-N1	-8.72	122.22	128.89

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	ADP	PA-O3A-PB	-6.70	110.20	132.67
2	B	2001	ADP	PA-O3A-PB	-6.53	110.77	132.67
2	B	2001	ADP	C4-C5-N7	-2.13	107.52	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	ADP	3	0
2	B	2001	ADP	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	591/613 (96%)	0.58	59 (9%) 9 7	10, 63, 109, 125	0
1	B	591/613 (96%)	0.51	52 (8%) 12 10	9, 61, 110, 128	0
All	All	1182/1226 (96%)	0.55	111 (9%) 11 8	9, 62, 110, 128	0

The worst 5 of 111 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	248	GLY	8.1
1	B	249	SER	7.8
1	A	249	SER	7.7
1	B	1315	GLY	6.7
1	B	146	HIS	6.5

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	ADP	A	2001	27/27	0.80	0.27	0.52	113,114,115,115	0
2	ADP	B	2001	27/27	0.81	0.27	0.35	118,121,127,128	0
4	ZN	A	2003	1/1	0.93	0.09	-1.80	163,163,163,163	0
4	ZN	B	2003	1/1	0.89	0.05	-2.28	169,169,169,169	0
5	CL	A	2006	1/1	0.99	0.22	-	41,41,41,41	0
5	CL	B	2004	1/1	0.96	0.23	-	48,48,48,48	0
5	CL	A	2004	1/1	0.97	0.25	-	48,48,48,48	0
5	CL	B	2006	1/1	0.92	0.09	-	39,39,39,39	0
5	CL	B	2005	1/1	0.96	0.17	-	53,53,53,53	0
3	MG	A	2002	1/1	0.95	0.13	-	64,64,64,64	0
5	CL	A	2005	1/1	0.96	0.21	-	59,59,59,59	0
3	MG	B	2002	1/1	0.93	0.19	-	64,64,64,64	0

6.5 Other polymers [i](#)

There are no such residues in this entry.