



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 3, 2017 – 07:03 PM EDT

PDB ID : 2VF7
Title : Crystal structure of UvrA2 from *Deinococcus radiodurans*
Authors : Timmins, J.; Gordon, E.; Caria, S.; Leonard, G.; Kuo, M.S.; Monchois, V.;
McSweeney, S.
Deposited on : unknown
Resolution : 2.30 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-20029824
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-20029824

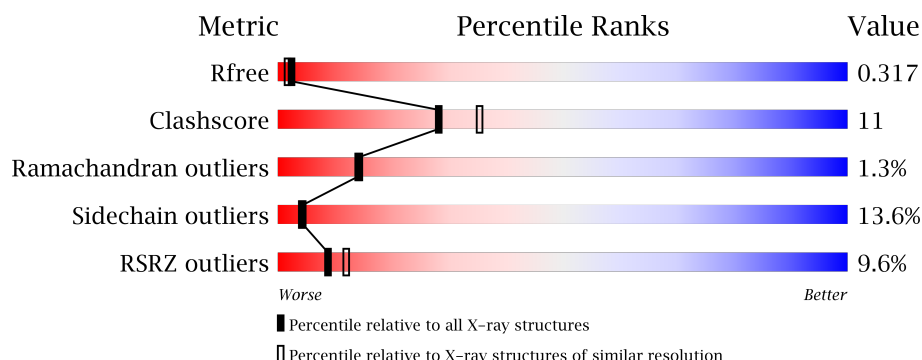
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	4130 (2.30-2.30)
Clashscore	112137	4751 (2.30-2.30)
Ramachandran outliers	110173	4705 (2.30-2.30)
Sidechain outliers	110143	4704 (2.30-2.30)
RSRZ outliers	101464	4156 (2.30-2.30)

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	842	<div> <div>8%</div> <div>71%</div> <div>20%</div> <div>5%</div> <div>.</div> </div>
1	B	842	<div> <div>7%</div> <div>70%</div> <div>22%</div> <div>.</div> <div>.</div> </div>
1	C	842	<div> <div>12%</div> <div>68%</div> <div>23%</div> <div>5%</div> <div>.</div> </div>

2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 19508 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.

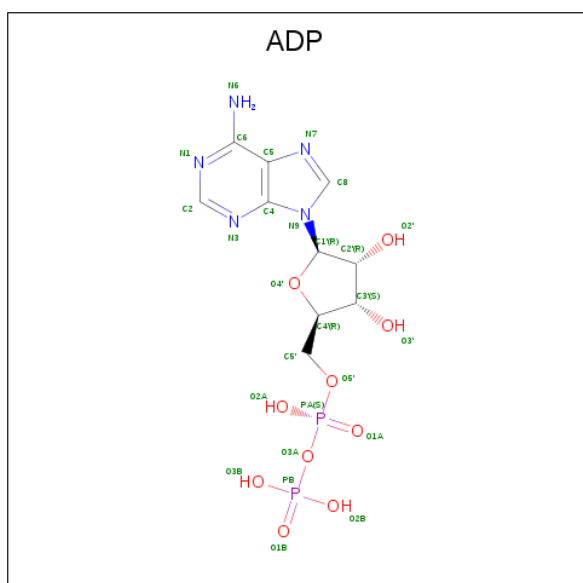
- Molecule 1 is a protein called EXCINUCLEASE ABC, SUBUNIT A..

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	815	Total	C	N	O	S	0	23	0
			6245	3918	1131	1177	19			
1	B	813	Total	C	N	O	S	0	26	0
			6240	3912	1134	1175	19			
1	C	814	Total	C	N	O	S	0	23	0
			6257	3930	1134	1174	19			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	746	ARG	GLN	engineered mutation	UNP Q9RYW8
B	746	ARG	GLN	engineered mutation	UNP Q9RYW8
C	746	ARG	GLN	engineered mutation	UNP Q9RYW8

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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Zn	0	0
			2	2		
3	A	2	Total	Zn	0	0
			2	2		
3	C	2	Total	Zn	0	0
			2	2		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	1	Total	Mg	0	0
			1	1		

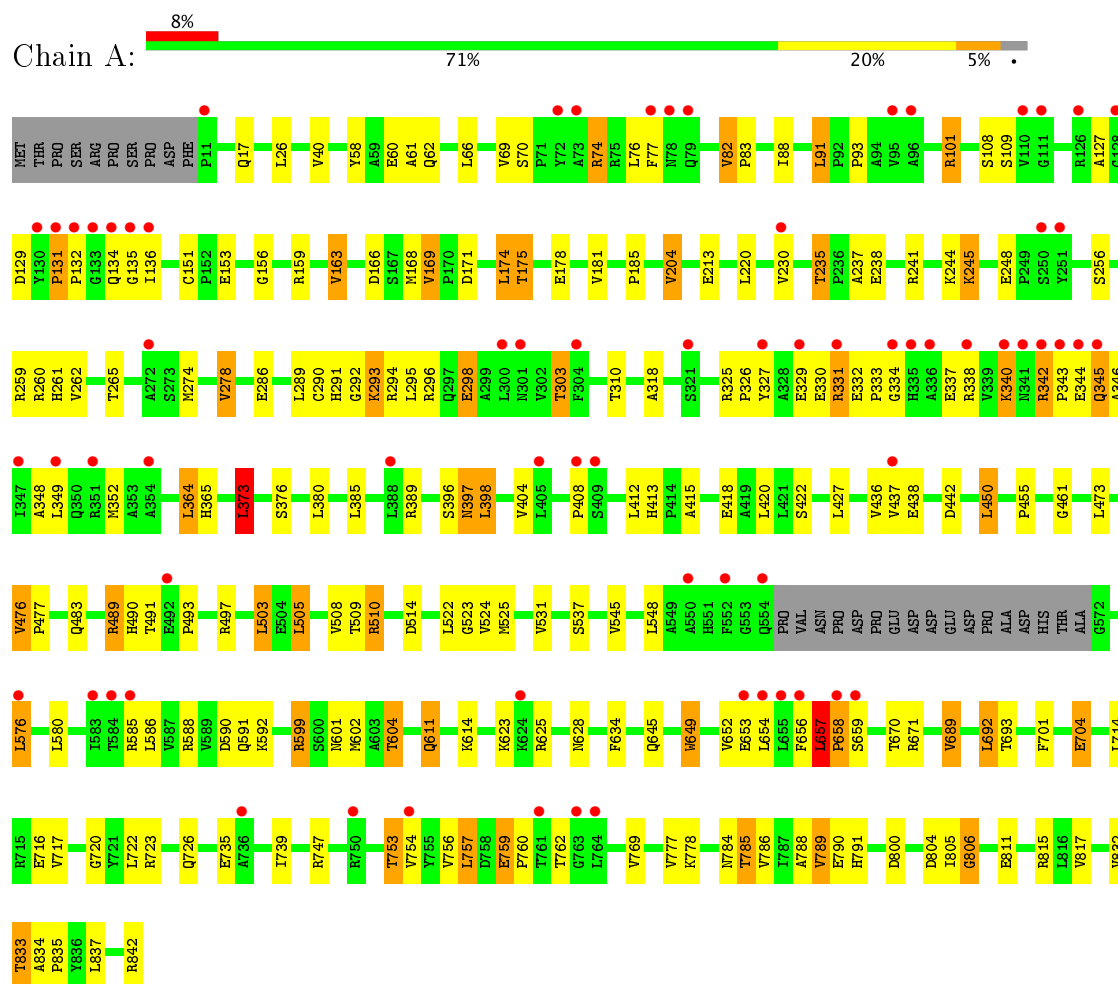
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	178	Total	O	0	0
			178	178		
5	B	232	Total	O	0	0
			232	232		
5	C	187	Total	O	0	0
			187	187		

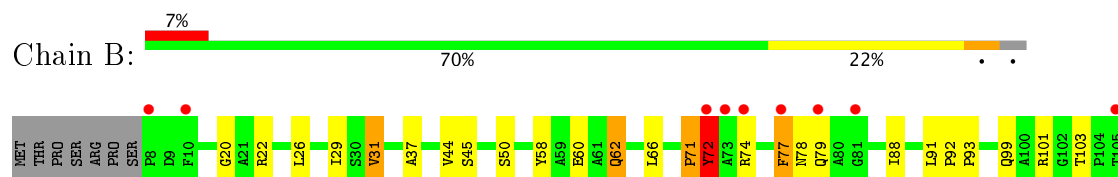
3 Residue-property plots [i](#)

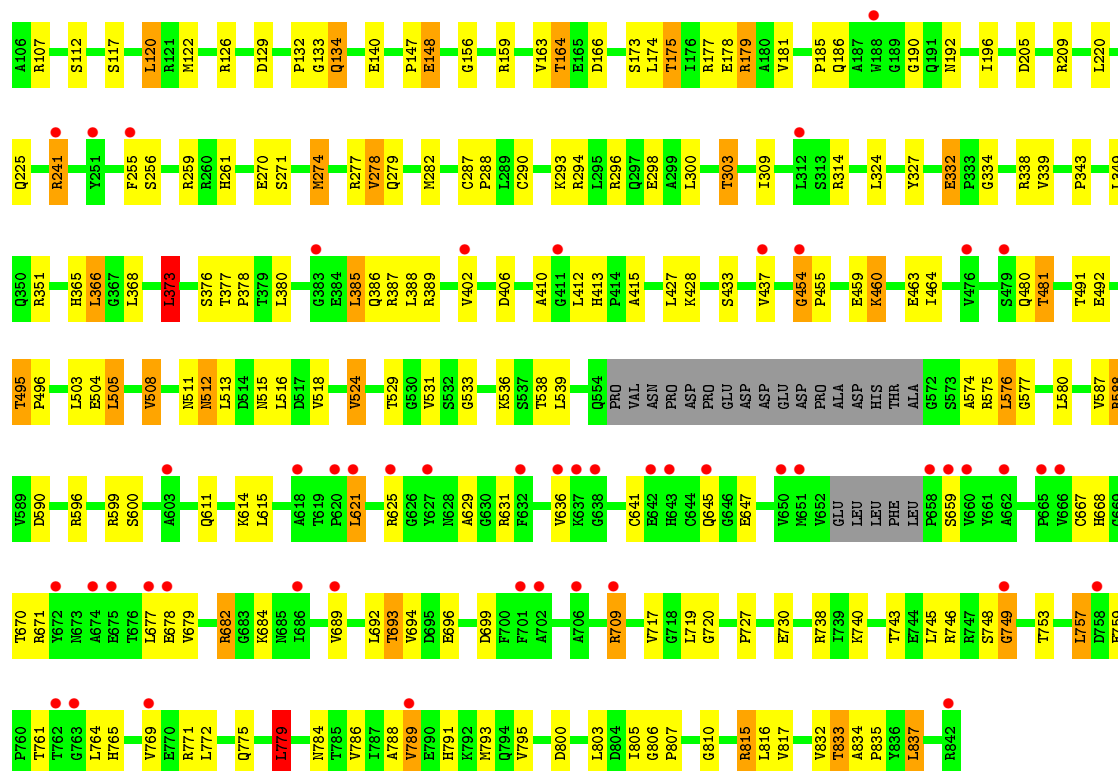
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

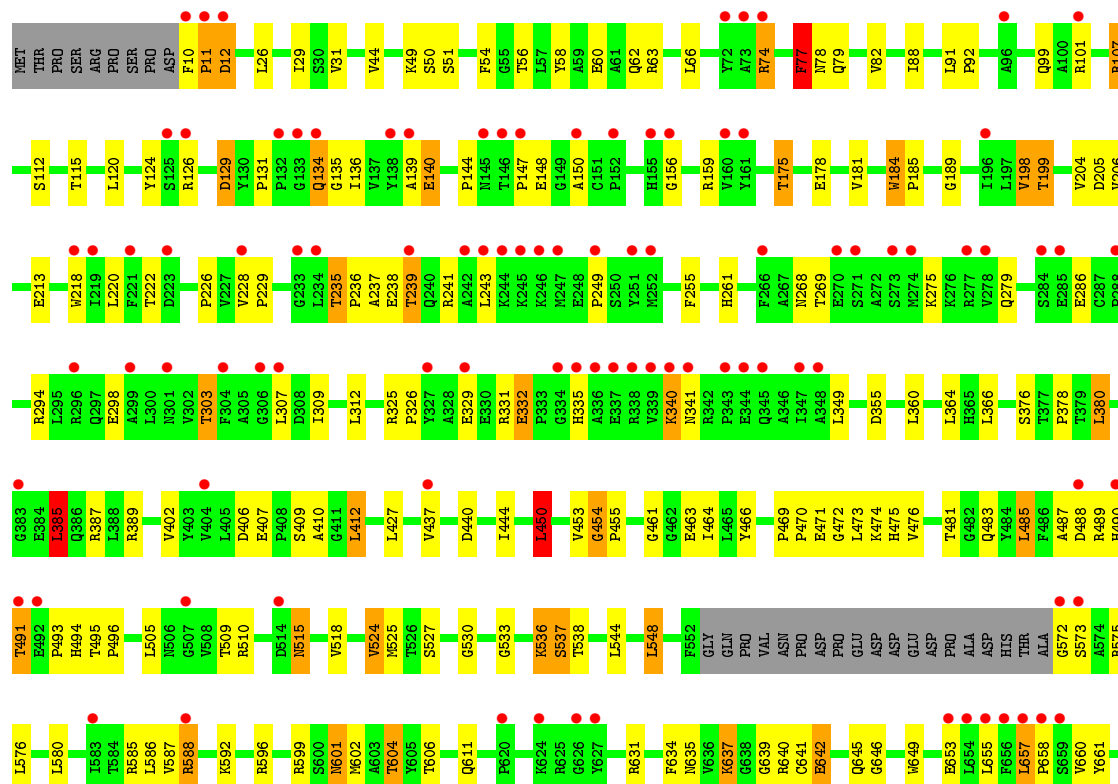


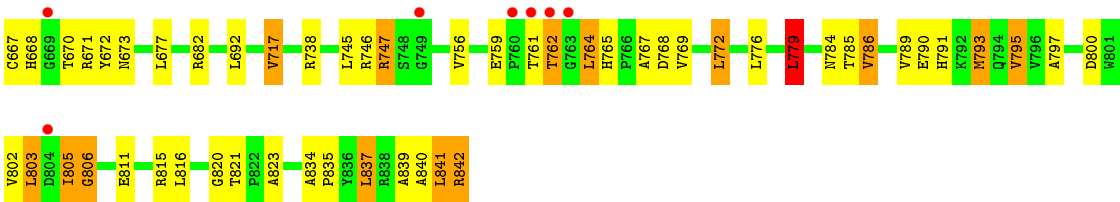
- Molecule 1: EXCINUCLEASE ABC, SUBUNIT A.





• Molecule 1: EXCINUCLEASE ABC, SUBUNIT A.





4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	271.71Å 111.67Å 103.29Å 90.00° 100.80° 90.00°	Depositor
Resolution (Å)	40.00 – 2.30 38.19 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.0 (40.00-2.30) 95.0 (38.19-2.30)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.18 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.219 , 0.295 0.259 , 0.317	Depositor DCC
R_{free} test set	6451 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	49.0	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 41.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	19508	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.47% 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 ⓘ

5.1 Standard geometry ⓘ

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

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	0.62	7/6375 (0.1%)	0.72	2/8662 (0.0%)
1	B	0.69	8/6376 (0.1%)	0.75	7/8661 (0.1%)
1	C	0.65	7/6395 (0.1%)	0.72	7/8689 (0.1%)
All	All	0.65	22/19146 (0.1%)	0.73	16/26012 (0.1%)

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	1
1	B	0	1
All	All	0	2

The worst 5 of 22 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	134	GLN	C-N	19.60	1.68	1.33
1	B	72	TYR	CE1-CZ	18.37	1.62	1.38
1	B	72	TYR	CG-CD1	15.35	1.59	1.39
1	A	330	GLU	CD-OE2	14.72	1.41	1.25
1	C	332	GLU	CD-OE1	10.20	1.36	1.25

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	72	TYR	CB-CG-CD2	-7.80	116.32	121.00
1	B	779	LEU	CA-CB-CG	7.79	133.23	115.30
1	B	385	LEU	CA-CB-CG	7.67	132.93	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	72	TYR	CD1-CG-CD2	7.00	125.61	117.90
1	C	485	LEU	CA-CB-CG	6.81	130.96	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	490	HIS	Peptide
1	B	682	ARG	Mainchain

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	6245	0	6115	131	0
1	B	6240	0	6100	136	0
1	C	6257	0	6158	146	0
2	A	54	0	24	1	0
2	B	54	0	24	1	0
2	C	54	0	24	6	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
4	C	1	0	0	0	0
5	A	178	0	0	12	0
5	B	232	0	0	19	0
5	C	187	0	0	22	0
All	All	19508	0	18445	415	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 11.

The worst 5 of 415 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:657:LEU:CD2	1:A:657:LEU:CG	1.76	1.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:134:GLN:C	1:C:135:GLY:N	1.68	1.43
1:C:184:TRP:HB2	1:C:185:PRO:HD2	1.27	1.14
1:C:657:LEU:HB3	1:C:658:PRO:HD2	1.39	1.04
2:C:902:ADP:H8	5:C:1116:HOH:O	1.48	0.95

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	811/842 (96%)	748 (92%)	51 (6%)	12 (2%)	12	11
1	B	809/842 (96%)	748 (92%)	51 (6%)	10 (1%)	15	16
1	C	812/842 (96%)	756 (93%)	46 (6%)	10 (1%)	15	16
All	All	2432/2526 (96%)	2252 (93%)	148 (6%)	32 (1%)	14	14

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	101	ARG
1	A	345	GLN
1	A	806	GLY
1	B	72	TYR
1	B	749	GLY

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	658/683 (96%)	567 (86%)	91 (14%)	4	4
1	B	657/683 (96%)	571 (87%)	86 (13%)	5	5
1	C	659/683 (96%)	566 (86%)	93 (14%)	4	4
All	All	1974/2049 (96%)	1704 (86%)	270 (14%)	4	4

5 of 270 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	373	LEU
1	B	631	ARG
1	C	682	ARG
1	B	389	ARG
1	B	512	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	512	ASN
1	B	773[A]	GLN
1	C	673	ASN
1	B	551	HIS
1	B	635	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 carbohydrates in this entry.

5.6 Ligand geometry

Of 13 ligands modelled in this entry, 7 are monoatomic - leaving 6 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	901	-	25,29,29	1.17	3 (12%)	24,45,45	1.77	1 (4%)
2	ADP	A	902	-	25,29,29	0.99	1 (4%)	24,45,45	1.69	2 (8%)
2	ADP	B	901	-	25,29,29	1.12	2 (8%)	24,45,45	1.69	3 (12%)
2	ADP	B	902	-	25,29,29	1.10	1 (4%)	24,45,45	1.63	2 (8%)
2	ADP	C	901	-	25,29,29	1.06	2 (8%)	24,45,45	1.83	1 (4%)
2	ADP	C	902	4	25,29,29	1.13	2 (8%)	24,45,45	1.61	2 (8%)

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	901	-	-	0/12/32/32	0/3/3/3
2	ADP	A	902	-	-	0/12/32/32	0/3/3/3
2	ADP	B	901	-	-	0/12/32/32	0/3/3/3
2	ADP	B	902	-	-	0/12/32/32	0/3/3/3
2	ADP	C	901	-	-	0/12/32/32	0/3/3/3
2	ADP	C	902	4	-	0/12/32/32	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	901	ADP	C2-N3	2.04	1.35	1.32
2	B	901	ADP	C2-N3	2.06	1.35	1.32
2	C	901	ADP	C2-N3	2.14	1.35	1.32
2	C	902	ADP	C2-N3	2.30	1.36	1.32
2	A	901	ADP	O4'-C1'	2.33	1.44	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	901	ADP	N3-C2-N1	-7.06	122.71	128.86
2	A	902	ADP	N3-C2-N1	-7.05	122.72	128.86
2	A	901	ADP	N3-C2-N1	-6.79	122.94	128.86
2	B	901	ADP	N3-C2-N1	-6.28	123.39	128.86
2	B	902	ADP	N3-C2-N1	-6.05	123.59	128.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	901	ADP	1	0
2	B	901	ADP	1	0
2	C	901	ADP	3	0
2	C	902	ADP	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	C	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	134:GLN	C	135:GLY	N	1.68

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	815/842 (96%)	0.78	70 (8%) 11 16	43, 57, 70, 79	6 (0%)
1	B	813/842 (96%)	0.74	60 (7%) 15 21	42, 56, 70, 88	5 (0%)
1	C	814/842 (96%)	0.83	104 (12%) 4 6	42, 56, 69, 81	8 (0%)
All	All	2442/2526 (96%)	0.78	234 (9%) 9 12	42, 56, 70, 88	19 (0%)

The worst 5 of 234 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	336	ALA	9.3
1	A	654	LEU	8.0
1	C	654	LEU	7.4
1	C	656	PHE	6.8
1	C	491	THR	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
4	MG	C	905	1/1	0.97	0.30	1.79	42,42,42,42	0
3	ZN	C	903	1/1	0.72	0.17	0.03	70,70,70,70	1
3	ZN	A	903	1/1	0.99	0.14	-0.33	59,59,59,59	1
2	ADP	C	901	27/27	0.97	0.12	-0.40	48,55,56,57	1
2	ADP	A	901	27/27	0.97	0.12	-0.53	43,55,56,57	0
2	ADP	B	902	27/27	0.96	0.14	-0.90	46,50,58,59	0
2	ADP	C	902	27/27	0.96	0.14	-1.00	39,45,56,57	0
3	ZN	B	903	1/1	0.98	0.09	-1.02	71,71,71,71	0
3	ZN	B	904	1/1	0.88	0.14	-1.08	59,59,59,59	1
2	ADP	B	901	27/27	0.97	0.15	-1.39	35,40,42,43	0
2	ADP	A	902	27/27	0.98	0.12	-1.96	36,39,45,47	0
3	ZN	C	904	1/1	0.98	0.06	-3.30	66,66,66,66	1
3	ZN	A	904	1/1	1.00	0.08	-4.61	55,55,55,55	0

6.5 Other polymers [i](#)

There are no such residues in this entry.