wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:41 AM GMT

PDB ID	:	2XZL
Title	:	Upf1-RNA complex
Authors	:	Chakrabarti, S.; Jayachandran, U.; Bonneau, F.; Fiorini, F.; Basquin, C.; Domcke, S.; Le Hir, H.; Conti, E.
Deposited on	:	2010-11-26
Resolution	:	2.40 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

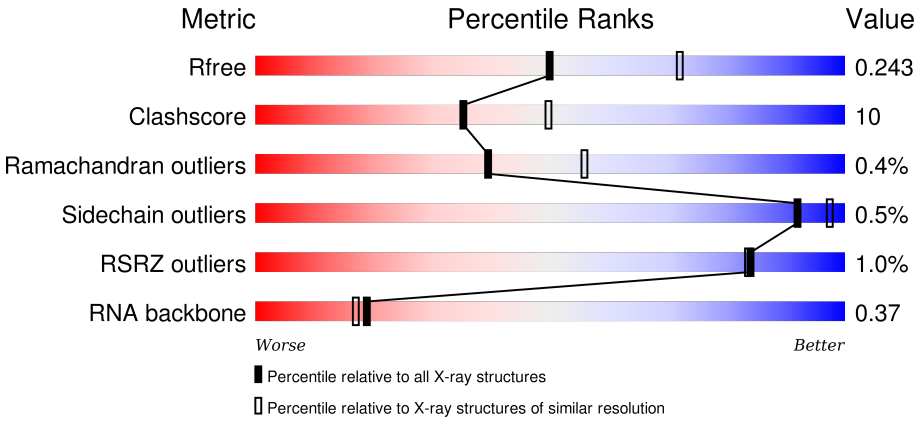
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)
RNA backbone	2183	1073 (2.90-1.90)

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 >=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 <=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	802	% 77% 17% 6%
2	B	9	67% 22% 11%

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
7	1PE	A	1858	-	-	-	X
7	1PE	A	1859	-	-	-	X
7	1PE	A	1860	-	-	-	X
7	1PE	A	1862	-	-	-	X
7	1PE	A	1863	-	-	-	X
7	1PE	A	1864	-	-	-	X
7	1PE	A	1866	-	-	-	X
7	1PE	A	1867	-	-	-	X

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 6041 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 ATP-DEPENDENT HELICASE NAM7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	756	Total	C	N	O	S	0	0	0
			5649	3626	979	1015	29			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	49	GLY	-	EXPRESSION TAG	UNP P30771
A	50	ALA	-	EXPRESSION TAG	UNP P30771
A	51	ALA	-	EXPRESSION TAG	UNP P30771
A	52	SER	-	EXPRESSION TAG	UNP P30771
A	53	MET	-	EXPRESSION TAG	UNP P30771

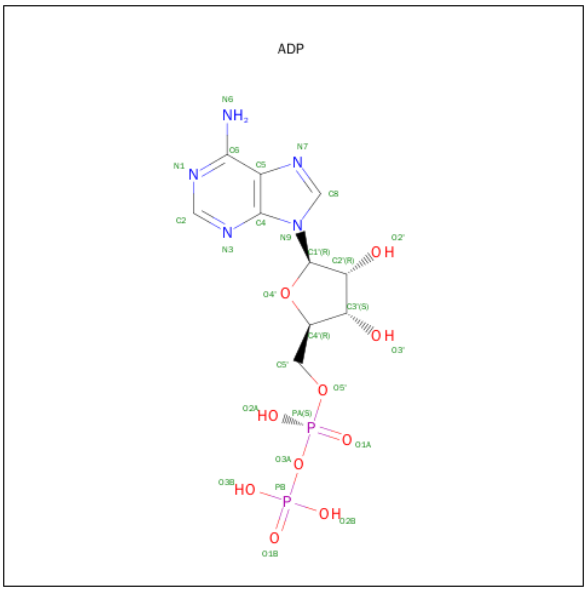
- Molecule 2 is a RNA chain called 5- R(*UP*UP*UP*UP*UP*UP*UP*UP*U) -3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	9	Total	C	N	O	P	0	0	0
			164	72	16	67	9			

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

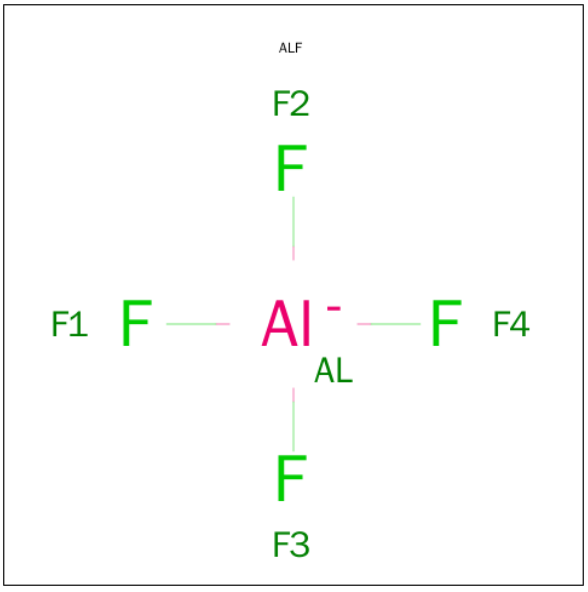
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	3	Total	Zn	0	0
			3	3		

- Molecule 4 is ADENOSINE-5'-DIPHOSPHATE (three-letter code: ADP) (formula: $C_{10}H_{15}N_5O_{10}P_2$).



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

- Molecule 5 is TETRAFLUOROALUMINATE ION (three-letter code: ALF) (formula: AlF_4).

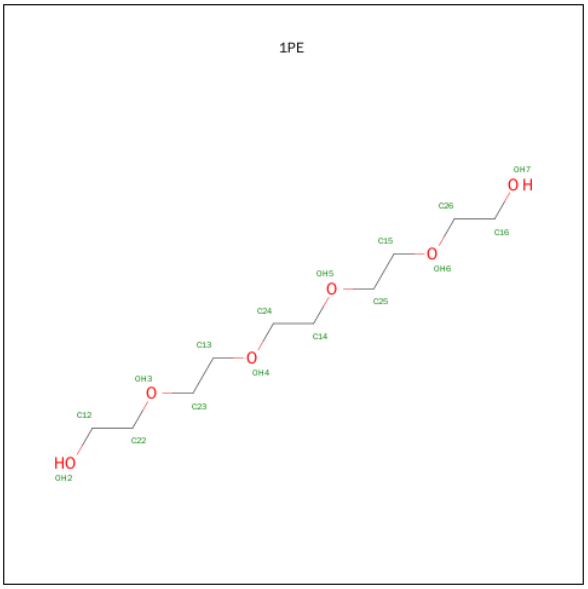


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	Al	F		0	0
			5	1	4			

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total Mg 1 1	0	0

- Molecule 7 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 10 6 4	0	0
7	A	1	Total C O 13 8 5	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 10 6 4	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0
7	A	1	Total C O 7 4 3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is water.

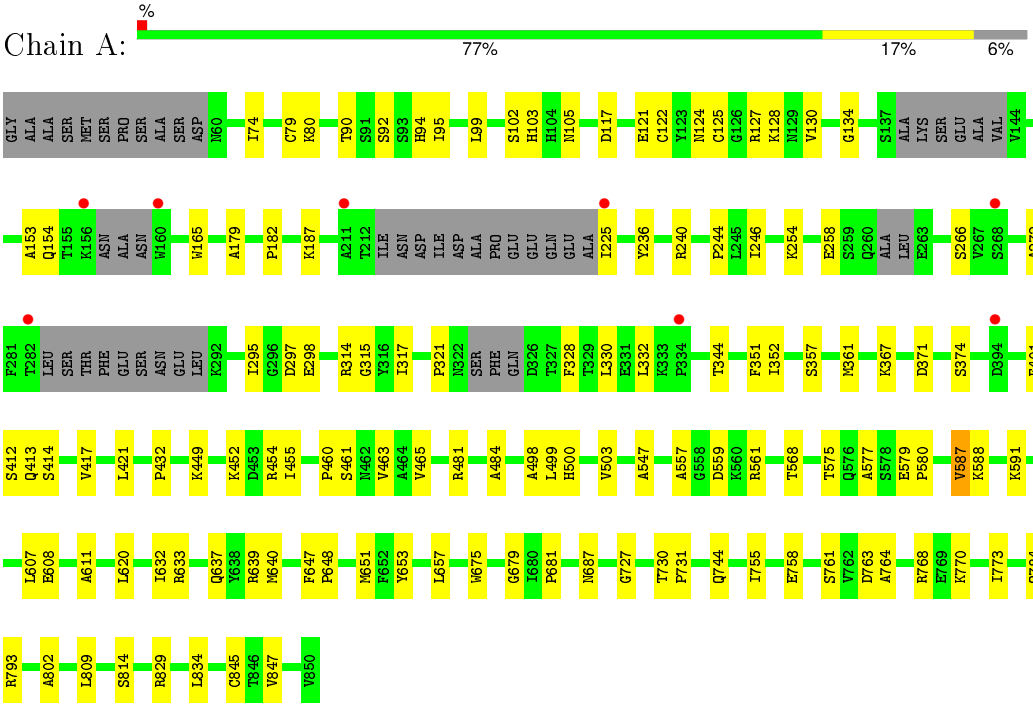
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	91	Total	O	0	0
			91	91		
8	B	5	Total	O	0	0
			5	5		

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 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: ATP-DEPENDENT HELICASE NAM7

Chain A:



- Molecule 2: 5- R(*UP*UP*UP*UP*UP*UP*UP*UP*U) -3

Chain B:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	64.14Å 114.12Å 65.73Å 90.00° 110.24° 90.00°	Depositor
Resolution (Å)	41.88 – 2.40 41.88 – 2.40	Depositor EDS
% Data completeness (in resolution range)	97.9 (41.88-2.40) 97.9 (41.88-2.40)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.68 (at 2.39Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.197 , 0.243 0.198 , 0.243	Depositor DCC
R_{free} test set	1726 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	19.5	Xtriage
Anisotropy	0.205	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 46.5	EDS
Estimated twinning fraction	0.030 for l,-k,h	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 33992 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6041	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.60% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: ALF, ZN, MG, ADP, 1PE

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	0.21	0/5763	0.41	0/7839
2	B	0.24	0/179	0.67	0/275
All	All	0.22	0/5942	0.42	0/8114

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	5649	0	5459	111	0
2	B	164	0	80	2	0
3	A	3	0	0	0	0
4	A	27	0	12	0	0
5	A	5	0	0	0	0
6	A	1	0	0	0	0
7	A	96	0	124	22	0
8	A	91	0	0	0	0
8	B	5	0	0	1	0
All	All	6041	0	5675	114	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 114 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:637:GLN:NE2	1:A:640:MET:H	1.47	1.13
1:A:637:GLN:HE22	1:A:640:MET:H	0.97	0.93
1:A:632:ILE:HD11	7:A:1864:1PE:H222	1.56	0.86
1:A:637:GLN:HE22	1:A:640:MET:N	1.73	0.84
1:A:834:LEU:HB3	7:A:1862:1PE:H122	1.64	0.80

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	742/802 (92%)	726 (98%)	13 (2%)	3 (0%)	39 56

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	117	ASP
1	A	124	ASN
1	A	587	VAL

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	564/703 (80%)	561 (100%)	3 (0%)	92 97

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	295	ILE
1	A	575	THR
1	A	788	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	419	HIS
1	A	637	GLN
1	A	658	GLN
1	A	659	ASN
1	A	744	GLN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	7/9 (77%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	B	7	U

There are no RNA pucker outliers to report.

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 (i)

Of 18 ligands modelled in this entry, 4 are monoatomic - leaving 14 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	$\#\|Z\| > 2$
4	ADP	A	1854	5,6	22,29,29	3.29	7 (31%)	27,45,45	3.18	9 (33%)
5	ALF	A	1855	8,4	0,4,4	0.00	-	0,6,6	0.00	-
7	1PE	A	1857	-	9,9,15	0.70	0	8,8,14	1.34	0
7	1PE	A	1858	-	12,12,15	0.69	0	11,11,14	1.44	0
7	1PE	A	1859	-	6,6,15	0.65	0	5,5,14	1.28	0
7	1PE	A	1860	-	6,6,15	0.61	0	5,5,14	1.45	0
7	1PE	A	1861	-	9,9,15	0.66	0	8,8,14	1.44	0
7	1PE	A	1862	-	6,6,15	0.69	0	5,5,14	1.23	0
7	1PE	A	1863	-	6,6,15	0.66	0	5,5,14	1.30	0
7	1PE	A	1864	-	6,6,15	0.65	0	5,5,14	1.38	0
7	1PE	A	1865	-	6,6,15	0.64	0	5,5,14	1.23	0
7	1PE	A	1866	-	6,6,15	0.67	0	5,5,14	1.27	0
7	1PE	A	1867	-	6,6,15	0.67	0	5,5,14	1.29	0
7	1PE	A	1868	-	6,6,15	0.66	0	5,5,14	1.29	0

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
4	ADP	A	1854	5,6	-	0/12/32/32	0/3/3/3
5	ALF	A	1855	8,4	-	0/0/0/0	0/0/0/0
7	1PE	A	1857	-	-	0/7/7/13	0/0/0/0
7	1PE	A	1858	-	-	0/10/10/13	0/0/0/0
7	1PE	A	1859	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1860	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1861	-	-	0/7/7/13	0/0/0/0
7	1PE	A	1862	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1863	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1864	-	-	0/4/4/13	0/0/0/0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	1PE	A	1865	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1866	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1867	-	-	0/4/4/13	0/0/0/0
7	1PE	A	1868	-	-	0/4/4/13	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1854	ADP	O3'-C3'	-2.00	1.38	1.43
4	A	1854	ADP	C5-N7	2.54	1.48	1.39
4	A	1854	ADP	C6-N6	2.94	1.44	1.34
4	A	1854	ADP	C2-N3	4.53	1.40	1.32
4	A	1854	ADP	C2-N1	5.22	1.43	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1854	ADP	N3-C2-N1	-11.57	120.03	128.89
4	A	1854	ADP	C4'-O4'-C1'	-4.55	104.72	109.72
4	A	1854	ADP	PA-O3A-PB	-3.43	121.18	132.67
4	A	1854	ADP	O2'-C2'-C3'	2.28	119.25	111.83
4	A	1854	ADP	O3'-C3'-C2'	2.36	119.51	111.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1858	1PE	1	0
7	A	1859	1PE	2	0
7	A	1860	1PE	1	0
7	A	1862	1PE	6	0
7	A	1863	1PE	3	0
7	A	1864	1PE	1	0
7	A	1865	1PE	2	0
7	A	1866	1PE	4	0
7	A	1867	1PE	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

6.1 Protein, DNA and RNA chains

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^{th} 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	756/802 (94%)	-0.14	8 (1%) 82 82	6, 20, 43, 66	0
2	B	9/9 (100%)	-0.18	0 100 100	17, 25, 44, 75	0
All	All	765/811 (94%)	-0.14	8 (1%) 84 83	6, 20, 43, 75	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	225	ILE	3.7
1	A	282	THR	2.9
1	A	394	ASP	2.9
1	A	160	TRP	2.7
1	A	334	PRO	2.6

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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^{th} 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^2)	Q<0.9
7	1PE	A	1864	7/16	0.88	0.40	13.39	21,28,37,45	0
7	1PE	A	1860	7/16	0.88	0.27	7.99	26,33,45,47	0
7	1PE	A	1859	7/16	0.91	0.28	6.95	22,29,35,39	0
7	1PE	A	1862	7/16	0.82	0.27	5.87	26,27,33,40	0
7	1PE	A	1863	7/16	0.91	0.25	5.42	18,28,32,46	0
7	1PE	A	1866	7/16	0.88	0.26	5.18	28,34,41,46	0
7	1PE	A	1867	7/16	0.89	0.25	4.37	34,36,44,44	0
7	1PE	A	1858	13/16	0.88	0.19	3.35	21,32,42,50	0
7	1PE	A	1868	7/16	0.92	0.16	1.10	18,22,30,42	0
7	1PE	A	1861	10/16	0.89	0.17	0.90	31,35,43,59	0
7	1PE	A	1865	7/16	0.95	0.16	-0.16	30,37,41,45	0
4	ADP	A	1854	27/27	0.97	0.14	-0.22	6,10,15,16	0
5	ALF	A	1855	5/5	0.99	0.15	-0.54	7,8,9,10	0
7	1PE	A	1857	10/16	0.95	0.11	-1.18	20,26,33,36	0
3	ZN	A	1852	1/1	0.97	0.10	-1.19	19,19,19,19	0
3	ZN	A	1851	1/1	1.00	0.09	-1.36	13,13,13,13	0
3	ZN	A	1853	1/1	0.97	0.08	-2.24	43,43,43,43	0
6	MG	A	1856	1/1	0.89	0.30	-	12,12,12,12	0

6.5 Other polymers (i)

There are no such residues in this entry.