



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 24, 2017 – 10:54 PM EST

PDB ID : 5KIY
Title : p97 ND1-A232E in complex with VIMP
Authors : Tang, W.K.; Xia, D.
Deposited on : 2016-06-17
Resolution : 2.79 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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-20030736
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-20030736

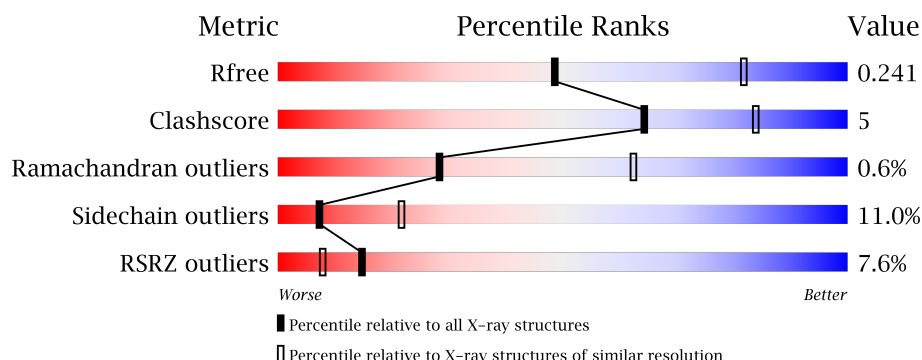
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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	468	
2	B	81	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 3887 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 Transitional endoplasmic reticulum ATPase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	454	Total	C	N	O	S	0	0	0
			3563	2232	634	679	18			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	232	GLU	ALA	engineered mutation	UNP P55072
A	461	ARG	-	expression tag	UNP P55072
A	462	SER	-	expression tag	UNP P55072
A	463	HIS	-	expression tag	UNP P55072
A	464	HIS	-	expression tag	UNP P55072
A	465	HIS	-	expression tag	UNP P55072
A	466	HIS	-	expression tag	UNP P55072
A	467	HIS	-	expression tag	UNP P55072
A	468	HIS	-	expression tag	UNP P55072

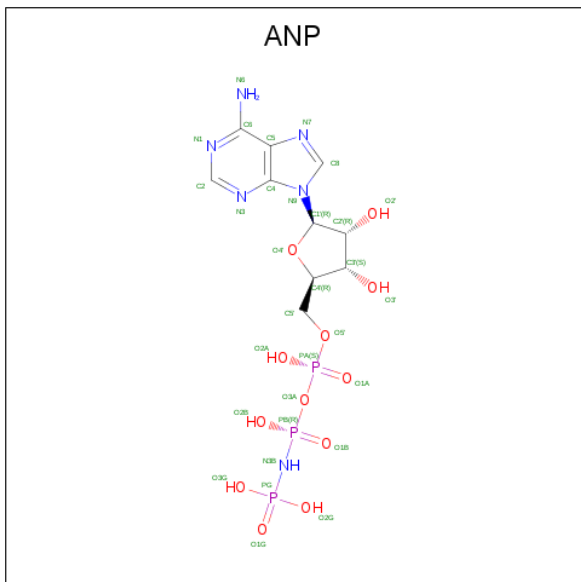
- Molecule 2 is a protein called Selenoprotein S.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	33	Total	C	N	O	S	0	0	0
			271	168	52	50	1			

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	42	MET	-	initiating methionine	UNP Q9BQE4
B	43	HIS	-	expression tag	UNP Q9BQE4
B	44	HIS	-	expression tag	UNP Q9BQE4
B	45	HIS	-	expression tag	UNP Q9BQE4
B	46	HIS	-	expression tag	UNP Q9BQE4
B	47	HIS	-	expression tag	UNP Q9BQE4
B	48	HIS	-	expression tag	UNP Q9BQE4

- Molecule 3 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

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

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

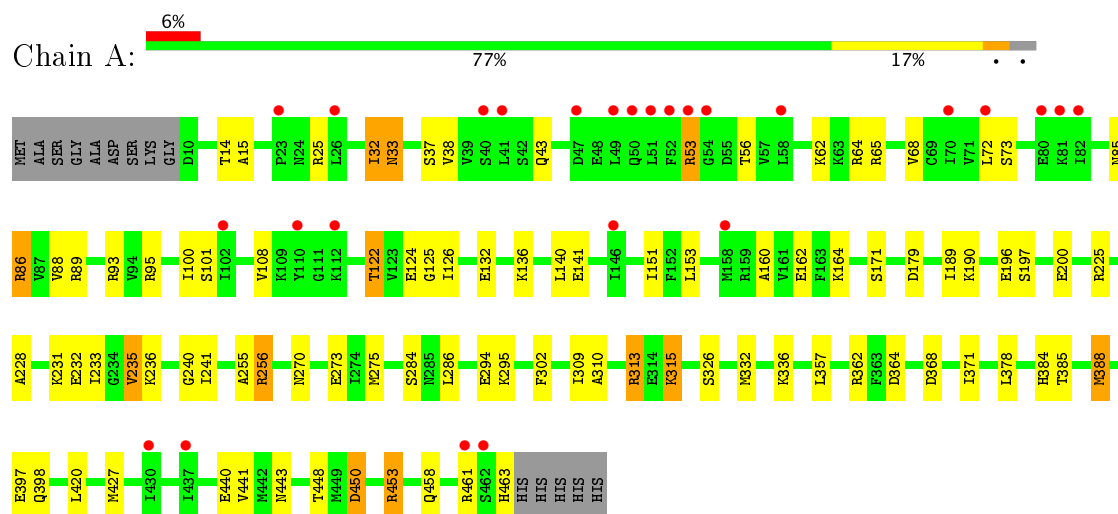
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	21	Total	O	0	0
			21	21		

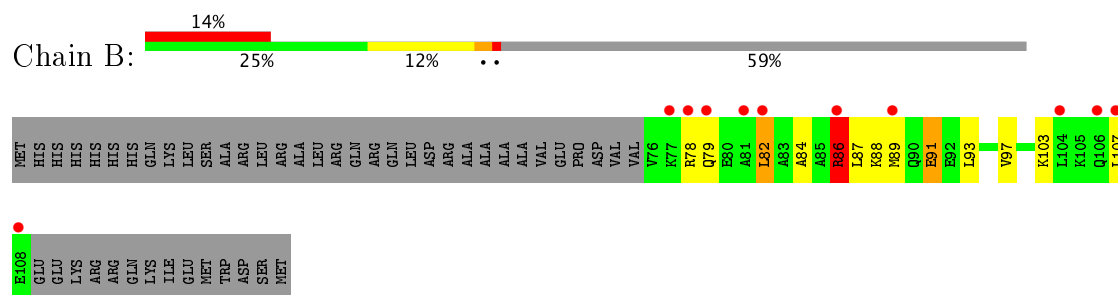
3 Residue-property plots [i](#)

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: Transitional endoplasmic reticulum ATPase



- Molecule 2: Selenoprotein S



4 Data and refinement statistics

Property	Value	Source
Space group	P 6 2 2	Depositor
Cell constants a, b, c, α , β , γ	145.00Å 145.00Å 119.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	125.57 – 2.79 47.46 – 2.79	Depositor EDS
% Data completeness (in resolution range)	98.2 (125.57-2.79) 98.3 (47.46-2.79)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.193 , 0.256 0.191 , 0.241	Depositor DCC
R_{free} test set	933 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	94.7	Xtriage
Anisotropy	0.036	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 74.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3887	wwPDB-VP
Average B, all atoms (Å ²)	126.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% 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: MG, ANP

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	0.85	0/3617	1.02	14/4885 (0.3%)
2	B	0.97	0/271	1.18	3/357 (0.8%)
All	All	0.86	0/3888	1.04	17/5242 (0.3%)

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

There are no bond length outliers.

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	453	ARG	NE-CZ-NH1	8.79	124.69	120.30
2	B	82	LEU	CA-CB-CG	7.18	131.83	115.30
1	A	368	ASP	CB-CG-OD1	6.87	124.48	118.30
2	B	86	ARG	NE-CZ-NH1	6.65	123.62	120.30
1	A	256	ARG	NE-CZ-NH1	6.62	123.61	120.30
2	B	89	MET	CG-SD-CE	6.09	109.94	100.20
1	A	388	MET	CG-SD-CE	-5.95	90.68	100.20
1	A	95	ARG	NE-CZ-NH1	5.83	123.21	120.30
1	A	53	ARG	NE-CZ-NH1	5.81	123.20	120.30
1	A	295	LYS	CA-CB-CG	5.79	126.13	113.40
1	A	241	ILE	CG1-CB-CG2	-5.40	99.52	111.40
1	A	313	ARG	NE-CZ-NH1	5.30	122.95	120.30
1	A	364	ASP	CB-CG-OD1	5.24	123.02	118.30
1	A	86	ARG	NE-CZ-NH1	5.11	122.85	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	450	ASP	CB-CG-OD2	-5.10	113.71	118.30
1	A	64	ARG	NE-CZ-NH1	5.01	122.81	120.30
1	A	89	ARG	NE-CZ-NH2	-5.01	117.80	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	124	GLU	Peptide

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	3563	0	3615	29	0
2	B	271	0	291	10	0
3	A	31	0	13	0	0
4	A	1	0	0	0	0
5	A	21	0	0	0	0
All	All	3887	0	3919	37	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 5.

All (37) 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)
2:B:82:LEU:O	2:B:86:ARG:HG2	1.79	0.81
1:A:153:LEU:HD11	1:A:160:ALA:HB1	1.68	0.75
2:B:84:ALA:O	2:B:88:LYS:HG3	1.91	0.69
1:A:56:THR:HG21	1:A:108:VAL:HG11	1.80	0.63
1:A:85:ASN:OD1	1:A:88:VAL:HG23	2.00	0.62
1:A:315:LYS:HA	1:A:315:LYS:CE	2.31	0.61
2:B:87:LEU:O	2:B:91:GLU:HG3	2.04	0.57
2:B:93:LEU:O	2:B:97:VAL:HG23	2.04	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:315:LYS:HE2	1:A:315:LYS:HA	1.87	0.54
1:A:62:LYS:O	1:A:65:ARG:HG2	2.08	0.54
2:B:79:GLN:OE1	2:B:82:LEU:CD1	2.57	0.53
1:A:255:ALA:HB2	1:A:302:PHE:CZ	2.45	0.52
1:A:33:ASN:HD22	2:B:93:LEU:HD13	1.75	0.52
2:B:79:GLN:OE1	2:B:82:LEU:HD11	2.11	0.51
2:B:84:ALA:HA	2:B:87:LEU:HD12	1.95	0.48
1:A:197:SER:O	1:A:200:GLU:HG2	2.13	0.48
2:B:87:LEU:O	2:B:91:GLU:CG	2.62	0.47
1:A:100:ILE:HD12	1:A:100:ILE:C	2.35	0.47
1:A:275:MET:HG2	1:A:309:ILE:HG22	1.95	0.47
1:A:179:ASP:OD2	2:B:78:ARG:NH2	2.49	0.45
1:A:378:LEU:HD13	1:A:397:GLU:HA	2.00	0.43
1:A:151:ILE:HD11	1:A:164:LYS:HD2	2.01	0.43
1:A:228:ALA:HA	1:A:231:LYS:HG2	2.02	0.42
1:A:65:ARG:HG3	1:A:93:ARG:HG2	2.02	0.42
1:A:153:LEU:CD1	1:A:160:ALA:HB1	2.45	0.42
1:A:371:ILE:N	1:A:371:ILE:HD12	2.35	0.42
1:A:310:ALA:CB	1:A:357:LEU:HD11	2.50	0.41
1:A:398:GLN:NE2	1:A:398:GLN:HA	2.35	0.41
1:A:233:ILE:HG23	1:A:235:VAL:CG1	2.51	0.41
1:A:122:THR:HG21	1:A:162:GLU:H	1.85	0.41
1:A:14:THR:O	1:A:14:THR:HG22	2.21	0.41
1:A:164:LYS:HD3	1:A:189:ILE:HD13	2.02	0.40
1:A:38:VAL:HG21	1:A:72:LEU:HD12	2.04	0.40
1:A:125:GLY:O	1:A:126:ILE:HG23	2.21	0.40
1:A:136:LYS:O	1:A:140:LEU:HB2	2.21	0.40
1:A:240:GLY:HA3	1:A:362:ARG:O	2.22	0.40
1:A:385:THR:HA	1:A:388:MET:HG3	2.03	0.40

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	452/468 (97%)	435 (96%)	14 (3%)	3 (1%)	25	59
2	B	31/81 (38%)	31 (100%)	0	0	100	100
All	All	483/549 (88%)	466 (96%)	14 (3%)	3 (1%)	28	62

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	15	ALA
1	A	32	ILE
1	A	313	ARG

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	391/401 (98%)	349 (89%)	42 (11%)	8	22
2	B	28/70 (40%)	24 (86%)	4 (14%)	4	11
All	All	419/471 (89%)	373 (89%)	46 (11%)	7	22

All (46) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	ARG
1	A	32	ILE
1	A	33	ASN
1	A	37	SER
1	A	43	GLN
1	A	53	ARG
1	A	68	VAL
1	A	73	SER
1	A	86	ARG
1	A	101	SER
1	A	122	THR
1	A	132	GLU
1	A	141	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	171	SER
1	A	190	LYS
1	A	196	GLU
1	A	225	ARG
1	A	232	GLU
1	A	235	VAL
1	A	236	LYS
1	A	256	ARG
1	A	270	ASN
1	A	273	GLU
1	A	284	SER
1	A	286	LEU
1	A	294	GLU
1	A	315	LYS
1	A	326	SER
1	A	332	MET
1	A	336	LYS
1	A	384	HIS
1	A	420	LEU
1	A	427	MET
1	A	440	GLU
1	A	441	VAL
1	A	443	ASN
1	A	448	THR
1	A	450	ASP
1	A	453	ARG
1	A	458	GLN
1	A	461	ARG
1	A	463	HIS
2	B	86	ARG
2	B	91	GLU
2	B	103	LYS
2	B	107	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	91	ASN
1	A	183	HIS
1	A	285	ASN
1	A	296	ASN
1	A	398	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	443	ASN
1	A	458	GLN
2	B	90	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](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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
3	ANP	A	800	4	29,33,33	1.58	8 (27%)	28,52,52	2.10	8 (28%)

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
3	ANP	A	800	4	-	0/13/38/38	0/3/3/3

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	800	ANP	PG-O2G	-2.85	1.48	1.56
3	A	800	ANP	PG-O3G	-2.71	1.49	1.56
3	A	800	ANP	PB-O2B	-2.42	1.50	1.56
3	A	800	ANP	C5-N7	-2.14	1.32	1.39
3	A	800	ANP	C5-C4	2.35	1.45	1.40
3	A	800	ANP	PB-N3B	2.41	1.69	1.63
3	A	800	ANP	PB-O1B	2.52	1.49	1.46
3	A	800	ANP	PG-N3B	3.65	1.73	1.63

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	800	ANP	N3-C2-N1	-5.75	123.85	128.86
3	A	800	ANP	PA-O3A-PB	-2.56	123.34	132.38
3	A	800	ANP	O1B-PB-N3B	-2.23	108.45	111.79
3	A	800	ANP	C2'-C3'-C4'	-2.01	98.70	102.62
3	A	800	ANP	O2A-PA-O1A	2.26	123.99	112.28
3	A	800	ANP	N6-C6-N1	2.43	123.58	118.77
3	A	800	ANP	O2B-PB-O1B	4.52	119.26	109.87
3	A	800	ANP	O3G-PG-O2G	4.86	121.30	107.69

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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, 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	454/468 (97%)	0.15	26 (5%) 24 16	76, 118, 172, 233	0
2	B	33/81 (40%)	1.39	11 (33%) 0 0	165, 184, 210, 228	0
All	All	487/549 (88%)	0.23	37 (7%) 15 8	76, 119, 184, 233	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	78	ARG	7.8
1	A	461	ARG	5.1
1	A	52	PHE	5.1
1	A	49	LEU	4.8
1	A	51	LEU	4.2
2	B	82	LEU	4.0
1	A	72	LEU	4.0
2	B	79	GLN	3.9
1	A	80	GLU	3.8
2	B	77	LYS	3.8
1	A	26	LEU	3.6
2	B	89	MET	3.6
2	B	108	GLU	3.6
1	A	23	PRO	3.4
2	B	86	ARG	3.4
1	A	41	LEU	3.3
1	A	53	ARG	3.3
1	A	54	GLY	3.2
1	A	102	ILE	3.0
2	B	106	GLN	3.0
1	A	47	ASP	2.9
1	A	158	MET	2.8
1	A	70	ILE	2.6
2	B	104	LEU	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	462	SER	2.6
2	B	81	ALA	2.6
1	A	50	GLN	2.4
1	A	58	LEU	2.4
1	A	110	TYR	2.4
1	A	437	ILE	2.2
1	A	112	LYS	2.2
1	A	81	LYS	2.1
1	A	40	SER	2.1
1	A	82	ILE	2.1
1	A	430	ILE	2.0
1	A	146	ILE	2.0
2	B	107	LEU	2.0

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
3	ANP	A	800	31/31	0.97	0.22	1.31	78,84,93,99	0
4	MG	A	801	1/1	0.99	0.23	-	92,92,92,92	0

6.5 Other polymers [i](#)

There are no such residues in this entry.