



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 07:55 am GMT

PDB ID : 2J0S
Title : The crystal structure of the Exon Junction Complex at 2.2 Å resolution
Authors : Bono, F.; Ebert, J.; Lorentzen, E.; Conti, E.
Deposited on : 2006-08-04
Resolution : 2.21 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

<https://www.wwpdb.org/validation/2017/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 : 1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

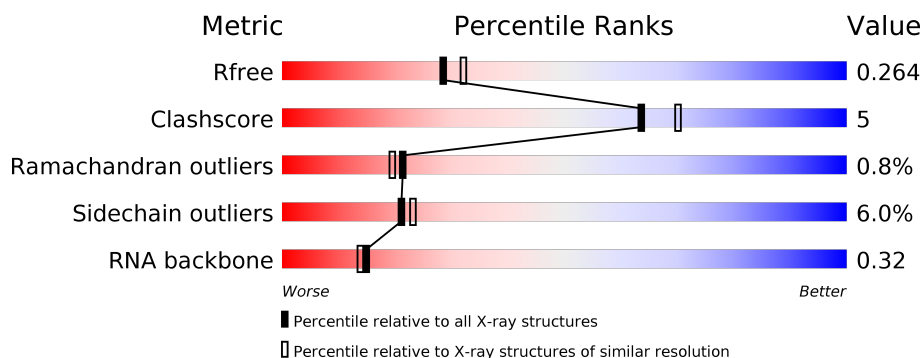
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.21 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}	111664	5192 (2.24-2.20)
Clashscore	122126	5911 (2.24-2.20)
Ramachandran outliers	120053	5828 (2.24-2.20)
Sidechain outliers	120020	5829 (2.24-2.20)
RNA backbone	2636	1043 (2.70-1.74)

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	410	82% 10% • 5%
2	C	146	85% 11% • •
3	D	89	87% 11% •
4	E	15	27% 13% 60%
5	T	150	23% 6% 71%

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 5906 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 RNA HELICASE DDX48.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	391	Total	C	N	O	S	0	8	0
			3138	1986	540	591	21			

- Molecule 2 is a protein called PROTEIN MAGO NASHI HOMOLOG.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	143	Total	C	N	O	S	0	2	0
			1192	773	199	217	3			

- Molecule 3 is a protein called RNA-BINDING PROTEIN 8A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	89	Total	C	N	O	S	0	1	0
			710	453	116	138	3			

- Molecule 4 is a RNA chain called 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP*UP*UP*UP*UP*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	6	Total	C	N	O	P	0	0	0
			121	54	12	49	6			

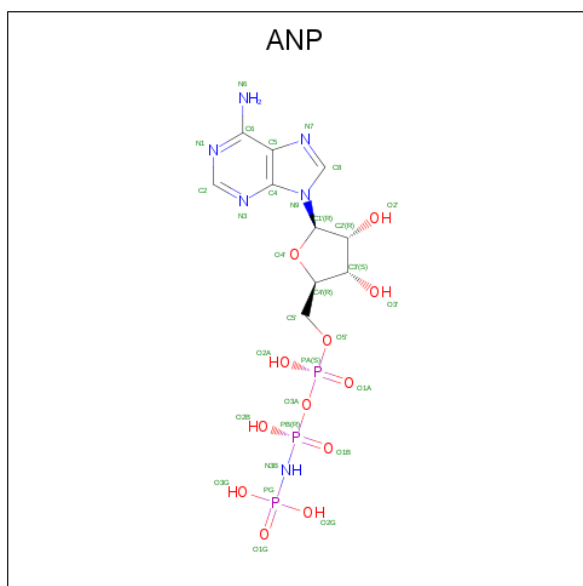
- Molecule 5 is a protein called PROTEIN CASC3.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
5	T	44	Total	C	N	O	0	0	0
			369	228	69	72			

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

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

- Molecule 7 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
7	A	1	Total	C	N	O	P	0	0
			31	10	6	12	3		

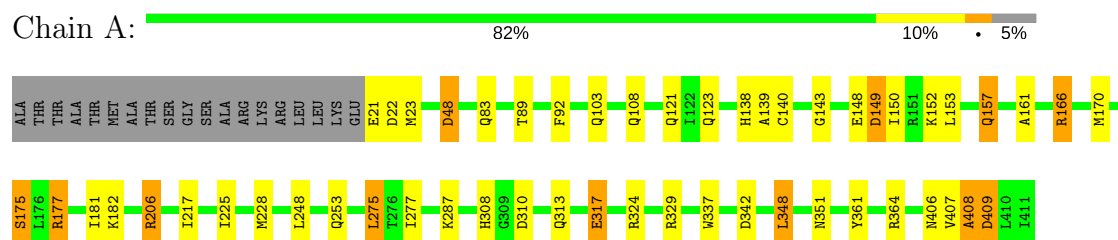
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	168	Total	O	0	0
			168	168		
8	C	117	Total	O	0	0
			117	117		
8	D	46	Total	O	0	0
			46	46		
8	E	4	Total	O	0	0
			4	4		
8	T	9	Total	O	0	0
			9	9		

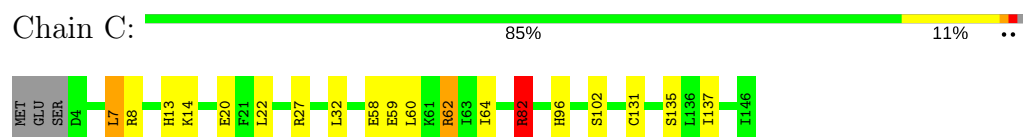
3 Residue-property plots

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

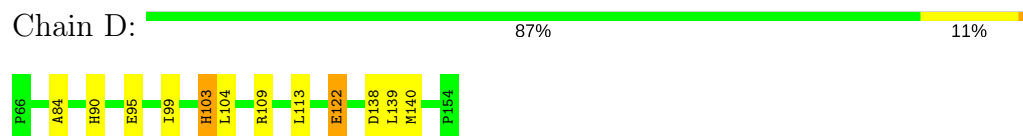
• Molecule 1: ATP-DEPENDENT RNA HELICASE DDX48



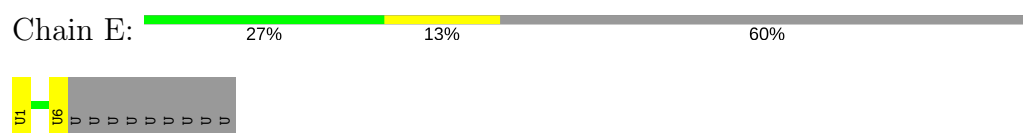
• Molecule 2: PROTEIN MAGO NASHI HOMOLOG



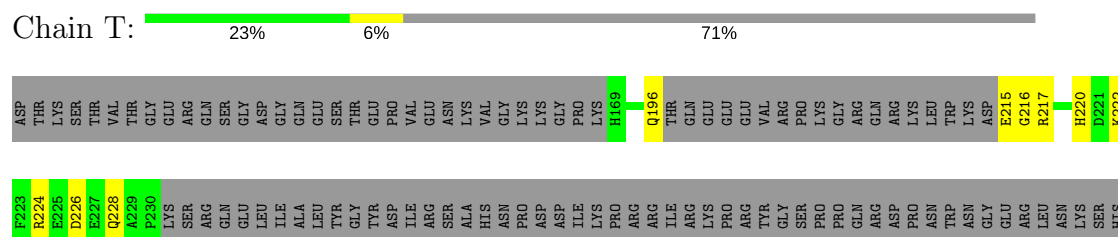
• Molecule 3: RNA-BINDING PROTEIN 8A



• Molecule 4: 5'-R(*UP*UP*UP*UP*UP*UP*UP*UP*UP*UP *UP*UP*UP*UP*U)-3'



• Molecule 5: PROTEIN CASC3



ARG
HIS
GLN
GLY

4 Data and refinement statistics

Property	Value	Source
Space group	P 64	Depositor
Cell constants a, b, c, α , β , γ	169.44Å 169.44Å 71.04Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.73 – 2.21 43.72 – 2.21	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.73-2.21) 99.7 (43.72-2.21)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.39 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.185 , 0.216 0.237 , 0.264	Depositor DCC
R_{free} test set	2931 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	35.6	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 51.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.030 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	5906	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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: 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	0.62	0/3211	0.75	2/4340 (0.0%)
2	C	0.80	1/1228 (0.1%)	0.84	3/1655 (0.2%)
3	D	0.65	0/730	0.67	0/989
4	E	1.33	1/132 (0.8%)	1.77	3/200 (1.5%)
5	T	0.66	0/379	0.72	0/510
All	All	0.69	2/5680 (0.0%)	0.80	8/7694 (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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	E	1	U	OP3-P	-9.43	1.49	1.61
2	C	131	CYS	CB-SG	-6.91	1.70	1.82

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	27	ARG	NE-CZ-NH1	7.26	123.93	120.30
1	A	48	ASP	CB-CG-OD1	6.39	124.05	118.30
4	E	1	U	O4'-C1'-N1	6.33	113.26	108.20
2	C	27	ARG	NE-CZ-NH2	-5.86	117.37	120.30
2	C	82	ARG	NE-CZ-NH2	5.72	123.16	120.30
1	A	275	LEU	CA-CB-CG	5.35	127.61	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	E	6	U	C5'-C4'-O4'	5.15	115.28	109.10
4	E	6	U	O5'-C5'-C4'	5.12	121.42	111.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	177	ARG	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	3138	0	3166	35	0
2	C	1192	0	1173	8	0
3	D	710	0	662	6	0
4	E	121	0	61	0	0
5	T	369	0	308	4	0
6	A	1	0	0	0	0
7	A	31	0	13	0	0
8	A	168	0	0	0	0
8	C	117	0	0	2	0
8	D	46	0	0	0	0
8	E	4	0	0	0	0
8	T	9	0	0	0	0
All	All	5906	0	5383	50	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 (50) 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:121:GLN:HE22	1:A:342:ASP:H	1.29	0.80
1:A:121:GLN:HE21	1:A:324:ARG:HH22	1.35	0.72

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:275:LEU:HD23	1:A:348[B]:LEU:CD2	2.18	0.72
2:C:102[A]:SER:OG	8:C:2087:HOH:O	2.08	0.70
1:A:275:LEU:HD23	1:A:348[B]:LEU:HD21	1.74	0.68
1:A:150:ILE:HG23	1:A:175:SER:OG	1.95	0.67
1:A:108:GLN:OE1	1:A:182:LYS:NZ	2.18	0.66
3:D:104:LEU:HD11	3:D:113:LEU:HD13	1.79	0.65
1:A:275:LEU:HD23	1:A:348[B]:LEU:HD23	1.87	0.56
1:A:408:ALA:HB3	1:A:409:ASP:HB2	1.88	0.56
1:A:408:ALA:N	1:A:409:ASP:HB2	2.21	0.55
2:C:14:LYS:O	2:C:82:ARG:NH1	2.41	0.54
1:A:121:GLN:HE22	1:A:342:ASP:N	2.03	0.53
2:C:7:LEU:HD12	2:C:8:ARG:N	2.24	0.53
1:A:408:ALA:HB3	1:A:409:ASP:CB	2.40	0.52
2:C:137:ILE:HD11	8:C:2039:HOH:O	2.09	0.52
3:D:84:ALA:HA	3:D:140:MET:CE	2.40	0.51
1:A:351:ASN:ND2	1:A:364:ARG:HD3	2.26	0.50
1:A:206:ARG:HD3	5:T:222:LYS:HG3	1.92	0.50
1:A:121:GLN:NE2	1:A:342:ASP:H	2.06	0.50
1:A:408:ALA:CA	1:A:409:ASP:HB2	2.41	0.49
2:C:59:GLU:OE2	2:C:62:ARG:NH1	2.48	0.47
1:A:140[B]:CYS:SG	1:A:166:ARG:HG3	2.55	0.47
2:C:135:SER:OG	3:D:103:HIS:HD2	1.98	0.47
1:A:138:HIS:HB2	1:A:157:GLN:HG2	1.96	0.46
5:T:224:ARG:O	5:T:228:GLN:HB2	2.16	0.46
1:A:139:ALA:HA	1:A:161:ALA:O	2.16	0.46
1:A:149:ASP:HB3	1:A:170:MET:HE1	1.96	0.46
3:D:84:ALA:HA	3:D:140:MET:HE1	1.97	0.46
1:A:351:ASN:HD21	1:A:364:ARG:HD3	1.81	0.46
1:A:108:GLN:HA	1:A:181:ILE:HA	1.99	0.45
3:D:122:GLU:HG3	3:D:122:GLU:O	2.17	0.45
1:A:275:LEU:HA	1:A:348[B]:LEU:HD23	1.99	0.44
3:D:90:HIS:CD2	3:D:99:ILE:HD12	2.52	0.44
1:A:21:GLU:HA	1:A:22:ASP:C	2.38	0.44
1:A:206:ARG:HA	5:T:220:HIS:HA	1.99	0.43
1:A:143:GLY:HA3	1:A:313:GLN:HG3	2.00	0.43
1:A:308:HIS:HD2	1:A:310:ASP:H	1.67	0.43
1:A:89:THR:HA	1:A:92:PHE:CE2	2.54	0.43
5:T:215:GLU:O	5:T:217:ARG:N	2.52	0.43
1:A:313:GLN:O	1:A:317:GLU:HG2	2.18	0.43
1:A:351:ASN:ND2	1:A:364:ARG:HH11	2.16	0.42
1:A:138:HIS:CB	1:A:157:GLN:HG2	2.50	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:287:LYS:HG2	1:A:308:HIS:HB2	2.02	0.41
2:C:60:LEU:O	2:C:64:ILE:HG12	2.21	0.41
1:A:253:GLN:NE2	1:A:361:TYR:OH	2.46	0.41
1:A:225:ILE:O	1:A:228:MET:HG2	2.20	0.41
1:A:408:ALA:CB	1:A:409:ASP:HB2	2.49	0.41
2:C:13:HIS:HE1	2:C:20:GLU:OE1	2.04	0.41
1:A:83:GLN:HB3	1:A:248:LEU:HD21	2.02	0.40

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	397/410 (97%)	388 (98%)	6 (2%)	3 (1%)	21	19
2	C	143/146 (98%)	142 (99%)	1 (1%)	0	100	100
3	D	88/89 (99%)	86 (98%)	1 (1%)	1 (1%)	16	12
5	T	40/150 (27%)	36 (90%)	3 (8%)	1 (2%)	6	3
All	All	668/795 (84%)	652 (98%)	11 (2%)	5 (1%)	21	23

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	408	ALA
1	A	409	ASP
3	D	139	LEU
1	A	23	MET
5	T	216	GLY

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	345/360 (96%)	324 (94%)	21 (6%)	20	22
2	C	130/134 (97%)	122 (94%)	8 (6%)	20	21
3	D	72/75 (96%)	67 (93%)	5 (7%)	17	17
5	T	35/133 (26%)	33 (94%)	2 (6%)	23	25
All	All	582/702 (83%)	546 (94%)	36 (6%)	21	21

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	ASP
1	A	103	GLN
1	A	123	GLN
1	A	148	GLU
1	A	149	ASP
1	A	152	LYS
1	A	153	LEU
1	A	157	GLN
1	A	166	ARG
1	A	175	SER
1	A	177	ARG
1	A	206	ARG
1	A	217	ILE
1	A	277	ILE
1	A	317	GLU
1	A	329	ARG
1	A	337	TRP
1	A	348[A]	LEU
1	A	348[B]	LEU
1	A	406	ASN
1	A	407	VAL
2	C	7	LEU
2	C	22	LEU
2	C	32	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	58	GLU
2	C	62	ARG
2	C	82	ARG
2	C	96[A]	HIS
2	C	96[B]	HIS
3	D	95	GLU
3	D	103	HIS
3	D	109	ARG
3	D	122	GLU
3	D	138	ASP
5	T	196	GLN
5	T	226	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	121	GLN
1	A	123	GLN
1	A	157	GLN
1	A	230	ASN
1	A	308	HIS
1	A	351	ASN
1	A	356	ASN
2	C	13	HIS
2	C	86	GLN
3	D	81	HIS
3	D	90	HIS
3	D	103	HIS
3	D	128	GLN
5	T	196	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	E	5/15 (33%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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	$\# Z > 2$
7	ANP	A	1413	6	29,33,33	2.31	9 (31%)	29,52,52	2.14	7 (24%)

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
7	ANP	A	1413	6	-	0/13/38/38	0/3/3/3

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1413	ANP	PG-O2G	-3.33	1.47	1.56
7	A	1413	ANP	PB-O2B	-2.12	1.50	1.56
7	A	1413	ANP	C2-N3	2.20	1.35	1.32
7	A	1413	ANP	C5-C4	2.77	1.46	1.40
7	A	1413	ANP	PG-N3B	2.91	1.71	1.63
7	A	1413	ANP	O4'-C1'	3.07	1.45	1.41
7	A	1413	ANP	PB-N3B	3.58	1.72	1.63
7	A	1413	ANP	PG-O1G	4.88	1.51	1.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	1413	ANP	PB-O1B	7.65	1.54	1.46

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1413	ANP	N3-C2-N1	-7.66	122.31	128.86
7	A	1413	ANP	O1G-PG-N3B	-3.06	107.21	111.79
7	A	1413	ANP	C4-C5-N7	-2.34	107.15	109.41
7	A	1413	ANP	C1'-N9-C4	-2.33	122.61	126.64
7	A	1413	ANP	O3A-PB-N3B	-2.07	100.85	106.59
7	A	1413	ANP	O3G-PG-O2G	3.32	116.62	107.61
7	A	1413	ANP	O2B-PB-O1B	4.75	119.70	109.88

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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers [i](#)

Unable to reproduce the depositors R factor - this section is therefore empty.