



Full wwPDB X-ray Structure Validation Report i

Feb 1, 2016 – 03:09 PM GMT

PDB ID : 4BOC
Title : Structure of mitochondrial RNA polymerase elongation complex
Authors : Schwinghammer, K.; Cheung, A.; Morozov, Y.; Agaronyan, K.; Temiakov, D.; Cramer, P.
Deposited on : 2013-05-18
Resolution : 2.65 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 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

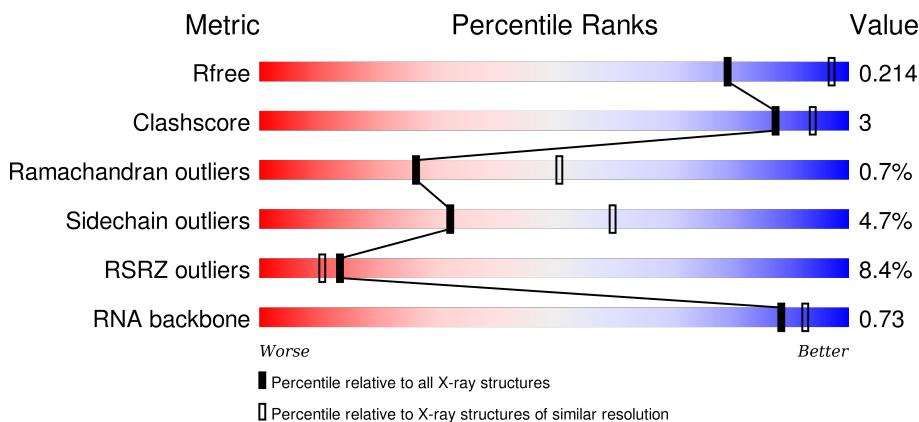
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

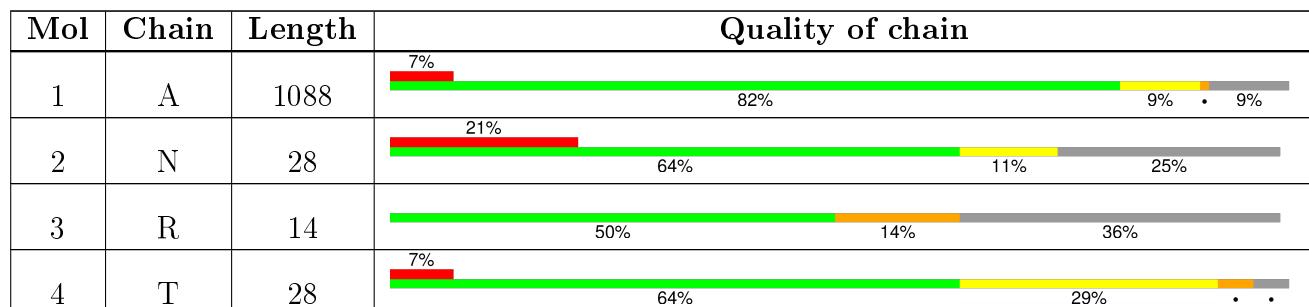
The reported resolution of this entry is 2.65 Å.

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	3152 (2.70-2.62)
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)
RSRZ outliers	91569	3161 (2.70-2.62)
RNA backbone	2183	1001 (3.08-2.24)

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.



2 Entry composition i

There are 5 unique types of molecules in this entry. The entry contains 9302 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 DNA-DIRECTED RNA POLYMERASE, MITOCHONDRIAL.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	989	7880	5017	1423	1392	48	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	143	MET	-	EXPRESSION TAG	UNP O00411
A	144	GLY	-	EXPRESSION TAG	UNP O00411
A	145	HIS	-	EXPRESSION TAG	UNP O00411
A	146	HIS	-	EXPRESSION TAG	UNP O00411
A	147	HIS	-	EXPRESSION TAG	UNP O00411
A	148	HIS	-	EXPRESSION TAG	UNP O00411
A	149	HIS	-	EXPRESSION TAG	UNP O00411
A	150	HIS	-	EXPRESSION TAG	UNP O00411

- Molecule 2 is a DNA chain called 5'-D(*CP*AP*TP*GP*GP*GP*GP*TP*AP*AP*TP*TP*T P*AP*TP *TP*TP*CP*GP*AP*CP*GP*CP*CP*AP*GP*AP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	N	21	437	205	89	122	21	0	0	0

- Molecule 3 is a RNA chain called 5'-R(*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*C P*GP*CP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	R	9	195	86	37	63	9	0	0	0

- Molecule 4 is a DNA chain called 5'-D(*CP*GP*TP*CP*TP*GP*GP*CP*GP*TP*GP*C P*GP*CP *GP*CP*CP*GP*CP*TP*AP*CP*CP*CP*CP*AP*TP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	T	27	Total	C	N	O	P	0	0	0
			546	258	96	165	27			

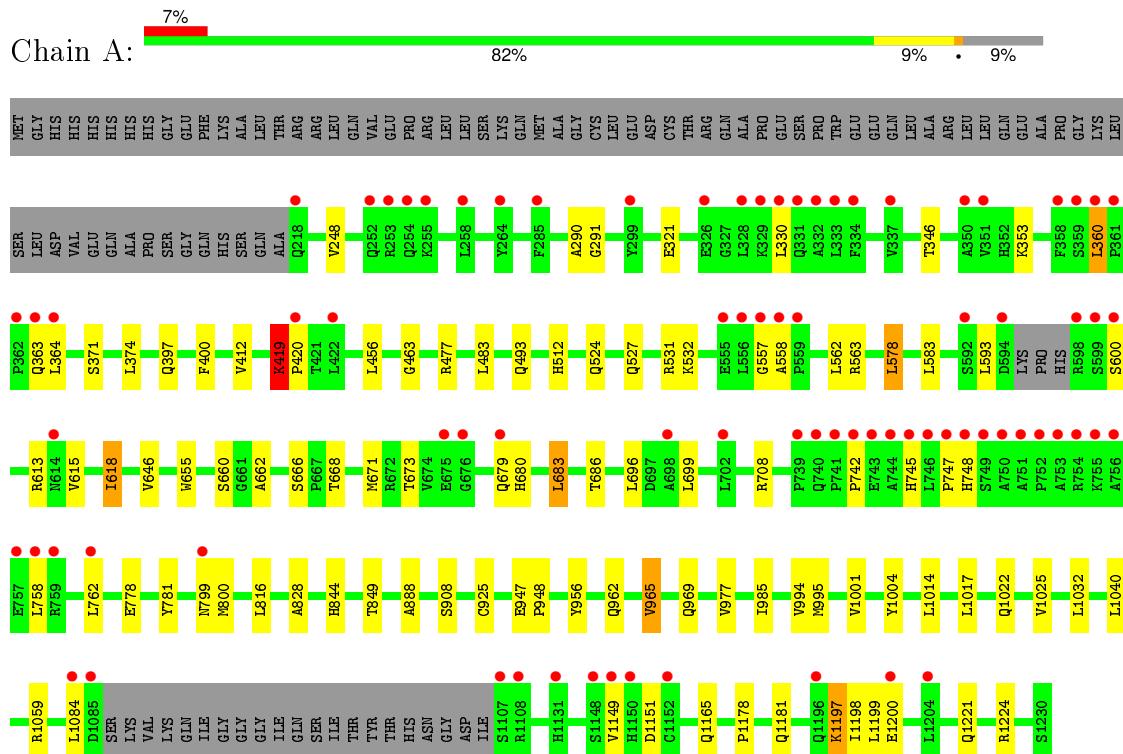
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	227	Total	O	0	0
			227	227		
5	N	7	Total	O	0	0
			7	7		
5	R	5	Total	O	0	0
			5	5		
5	T	5	Total	O	0	0
			5	5		

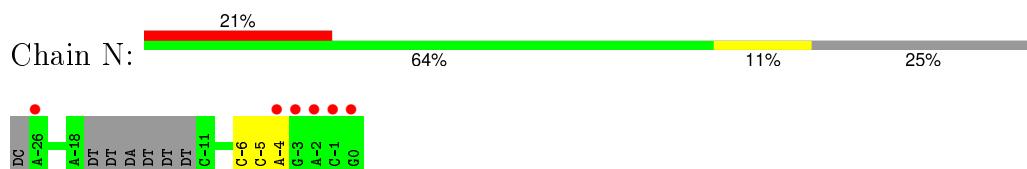
3 Residue-property plots

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 ($RSRZ > 2$). 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: DNA-DIRECTED RNA POLYMERASE, MITOCHONDRIAL



- Molecule 2: 5'-D(*CP*AP*TP*GP*GP*GP*GP*TP*AP*AP*TP*TP*AP*TP *TP*TP*CP*GP*AP*CP*GP*CP*CP*AP*GP*AP*CP*G)-3'



- Molecule 3: 5'-R(*AP*GP*UP*CP*UP*GP*CP*GP*GP*CP*GP*CP*GP*CP)-3'



- Molecule 4: 5'-D(*CP*GP*TP*CP*TP*GP*GP*CP*GP*TP*GP*CP*GP*CP *GP*CP*CP*GP*CP*TP*AP*CP*CP*CP*CP*AP*TP*G)-3'



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α , β , γ	225.19 Å 225.19 Å 225.19 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.81 – 2.65 39.81 – 2.65	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.81-2.65) 100.0 (39.81-2.65)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ ¹	1.83 (at 2.65 Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R , R_{free}	0.179 , 0.209 0.186 , 0.214	Depositor DCC
R_{free} test set	2741 reflections (5.25%)	DCC
Wilson B-factor (Å ²)	74.8	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 68.4	EDS
Estimated twinning fraction	0.023 for -l,-k,-h	Xtriage
L-test for twinning ²	$< L > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	1 of 55192 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	9302	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.23% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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 i

5.1 Standard geometry i

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	0.44	0/8071	0.65	0/10952
2	N	1.07	0/491	0.94	0/754
3	R	1.13	0/217	0.96	0/337
4	T	1.07	0/609	1.00	3/936 (0.3%)
All	All	0.58	0/9388	0.71	3/12979 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
4	T	0	DT	O4'-C4'-C3'	-6.48	101.91	104.50
4	T	4	DC	C4'-C3'-C2'	-5.12	98.49	103.10
4	T	0	DT	P-O3'-C3'	5.09	125.81	119.70

There are no chirality outliers.

There are no planarity outliers.

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	7880	0	7966	40	0
2	N	437	0	235	2	0
3	R	195	0	100	2	0
4	T	546	0	303	7	0
5	A	227	0	0	1	0
5	N	7	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	R	5	0	0	0	0
5	T	5	0	0	0	0
All	All	9302	0	8604	45	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 3.

All (45) 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:419:LYS:HB2	1:A:420:PRO:HA	1.39	1.04
1:A:363:GLN:HB3	1:A:364:LEU:HG	1.60	0.83
1:A:419:LYS:CB	1:A:420:PRO:HA	2.13	0.78
1:A:747:PRO:HB2	1:A:748:HIS:HA	1.71	0.71
1:A:419:LYS:HB2	1:A:420:PRO:CA	2.19	0.68
1:A:248:VAL:HG11	1:A:463:GLY:HA3	1.77	0.65
1:A:248:VAL:HG11	1:A:463:GLY:CA	2.28	0.63
1:A:985:ILE:HD12	1:A:1032:LEU:HD21	1.82	0.62
1:A:493:GLN:HG3	4:T:8:DG:C8	2.35	0.61
1:A:699:LEU:HD11	1:A:800:MET:HG3	1.82	0.61
1:A:412:VAL:HG21	1:A:646:VAL:HG21	1.83	0.60
1:A:1178:PRO:HB2	1:A:1181:GLN:HB2	1.83	0.59
1:A:708:ARG:HD3	1:A:828:ALA:HA	1.85	0.59
1:A:844:HIS:HD2	5:A:2149:HOH:O	1.86	0.58
1:A:747:PRO:CB	1:A:748:HIS:HA	2.36	0.56
4:T:-5:DT:HG2"	4:T:-4:DG:C8	2.42	0.55
1:A:1022:GLN:HA	1:A:1025:VAL:HG23	1.90	0.53
1:A:483:LEU:HB3	1:A:583:LEU:HD13	1.90	0.52
1:A:1004:TYR:CD2	4:T:0:DT:HG2"	2.46	0.50
4:T:-4:DG:HG2"	4:T:-3:DG:C8	2.47	0.49
1:A:1014:LEU:HA	1:A:1017:LEU:HD12	1.94	0.49
1:A:397:GLN:HG2	1:A:531:ARG:HG2	1.94	0.49
1:A:1001:VAL:HA	4:T:1:DG:HG2"	1.96	0.48
1:A:360:LEU:HD13	1:A:360:LEU:HG	1.79	0.47
2:N:-5:DC:HG2"	2:N:-4:DA:O5'	2.14	0.47
1:A:747:PRO:HB2	1:A:748:HIS:CA	2.41	0.47
1:A:816:LEU:HG	1:A:1149:VAL:HG13	1.97	0.47
1:A:655:TRP:HB3	1:A:696:LEU:HD22	1.97	0.46
1:A:1197:LYS:HD3	1:A:1200:GLU:HG2	1.97	0.45
1:A:781:TYR:HE2	4:T:6:DC:HG2"	1.81	0.45
1:A:618:ILE:HG21	3:R:1:G:HG1'	1.99	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:371:SER:HB3	1:A:374:LEU:HB2	1.97	0.45
1:A:948:PRO:HD3	1:A:1221:GLN:HB3	1.99	0.45
3:R:1:G:H2'	3:R:2:C:C6	2.53	0.44
1:A:400:PHE:HE2	1:A:527:GLN:HG2	1.83	0.43
1:A:557:GLY:HA3	1:A:558:ALA:HA	1.80	0.43
1:A:956:TYR:OH	1:A:995:MET:HG3	2.18	0.43
1:A:778:GLU:HG3	4:T:5:DG:H2"	2.00	0.43
1:A:290:ALA:N	1:A:291:GLY:HA2	2.34	0.41
1:A:965:VAL:O	1:A:969:GLN:HG3	2.20	0.41
1:A:683:LEU:HA	1:A:686:THR:HB	2.02	0.41
2:N:-6:DC:H2'	2:N:-5:DC:C6	2.55	0.41
1:A:849:THR:HA	1:A:888:ALA:HB1	2.02	0.41
1:A:994:VAL:HG13	1:A:1040:LEU:HD21	2.03	0.41
1:A:456:LEU:HD11	1:A:578:LEU:HD11	2.03	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	983/1088 (90%)	947 (96%)	29 (3%)	7 (1%)	26 51

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	419	LYS
1	A	600	SER
1	A	615	VAL
1	A	662	ALA
1	A	673	THR
1	A	745	HIS
1	A	742	PRO

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	857/943 (91%)	817 (95%)	40 (5%)	32 60

All (40) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	321	GLU
1	A	330	LEU
1	A	346	THR
1	A	353	LYS
1	A	360	LEU
1	A	419	LYS
1	A	477	ARG
1	A	512	HIS
1	A	524	GLN
1	A	532	LYS
1	A	562	LEU
1	A	563	ARG
1	A	578	LEU
1	A	593	LEU
1	A	613	ARG
1	A	618	ILE
1	A	660	SER
1	A	666	SER
1	A	668	THR
1	A	671	MET
1	A	679	GLN
1	A	680	HIS
1	A	683	LEU
1	A	758	LEU
1	A	762	LEU
1	A	799	ASN
1	A	908	SER
1	A	925	CYS
1	A	947	GLU
1	A	962	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	965	VAL
1	A	977	VAL
1	A	1059	ARG
1	A	1084	LEU
1	A	1151	ASP
1	A	1165	GLN
1	A	1197	LYS
1	A	1198	ILE
1	A	1199	LEU
1	A	1224	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	331	GLN

5.3.3 RNA [\(i\)](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	R	9/14 (64%)	1 (11%)	1 (11%)

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	R	2	C

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
3	R	1	G

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 [\(i\)](#)

There are no ligands in this entry.

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	989/1088 (90%)	0.37	80 (8%) 15 12	51, 84, 152, 198	0
2	N	21/28 (75%)	0.42	6 (28%) 1 0	87, 145, 199, 230	0
3	R	9/14 (64%)	-0.11	0 100 100	63, 72, 105, 116	0
4	T	27/28 (96%)	0.29	2 (7%) 17 14	73, 111, 233, 238	0
All	All	1046/1158 (90%)	0.36	88 (8%) 14 11	51, 85, 158, 238	0

All (88) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	218	GLN	7.9
1	A	745	HIS	7.3
1	A	749	SER	6.8
1	A	744	ALA	6.8
1	A	364	LEU	6.4
1	A	748	HIS	6.3
1	A	747	PRO	6.2
1	A	743	GLU	5.9
4	T	17	DT	5.8
1	A	742	PRO	5.6
1	A	750	ALA	5.6
1	A	746	LEU	5.2
1	A	253	ARG	5.2
1	A	679	GLN	5.1
1	A	254	GLN	4.9
1	A	362	PRO	4.9
1	A	752	PRO	4.9
1	A	758	LEU	4.8
1	A	755	LYS	4.6
1	A	332	ALA	4.6
1	A	598	ARG	4.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	600	SER	4.4
1	A	599	SER	4.3
1	A	333	LEU	4.1
1	A	258	LEU	4.0
4	T	-9	DC	4.0
1	A	594	ASP	3.9
1	A	1084	LEU	3.8
1	A	326	GLU	3.8
1	A	358	PHE	3.8
1	A	1107	SER	3.7
1	A	420	PRO	3.7
1	A	741	PRO	3.5
1	A	754	ARG	3.4
1	A	330	LEU	3.4
2	N	-1	DC	3.3
2	N	-3	DG	3.2
1	A	753	ALA	3.2
1	A	1108	ARG	3.2
1	A	363	GLN	3.2
1	A	759	ARG	3.1
1	A	614	ASN	3.1
2	N	0	DG	3.1
1	A	1196	GLN	2.9
1	A	331	GLN	2.9
1	A	762	LEU	2.8
1	A	255	LYS	2.8
1	A	1149	VAL	2.8
1	A	361	PRO	2.8
1	A	299	TYR	2.8
1	A	556	LEU	2.8
1	A	555	GLU	2.7
1	A	1148	SER	2.7
1	A	739	PRO	2.7
1	A	756	ALA	2.7
1	A	751	ALA	2.7
1	A	1085	ASP	2.7
1	A	559	PRO	2.7
1	A	757	GLU	2.6
1	A	351	VAL	2.6
1	A	557	GLY	2.5
1	A	799	ASN	2.5
1	A	285	PHE	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1204	LEU	2.5
1	A	360	LEU	2.5
1	A	359	SER	2.5
1	A	350	ALA	2.4
1	A	698	ALA	2.4
1	A	675	GLU	2.4
2	N	-26	DA	2.2
1	A	252	GLN	2.2
1	A	558	ALA	2.2
1	A	1152	CYS	2.2
1	A	329	LYS	2.2
1	A	702	LEU	2.2
2	N	-2	DA	2.2
2	N	-4	DA	2.2
1	A	422	LEU	2.1
1	A	328	LEU	2.1
1	A	337	VAL	2.1
1	A	740	GLN	2.1
1	A	1150	HIS	2.1
1	A	334	PHE	2.1
1	A	264	TYR	2.1
1	A	592	SER	2.0
1	A	676	GLY	2.0
1	A	1131	HIS	2.0
1	A	1200	GLU	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 [\(i\)](#)

There are no ligands in this entry.

6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.