



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2023 – 01:39 pm BST

PDB ID : 8P0Y
Title : The crystal structure of the C-terminal domain of Mengla nucleoprotein
Authors : Ferrero, D.S.; Tomas Gilabert, O.; Verdaguer, N.
Deposited on : 2023-05-11
Resolution : 4.12 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

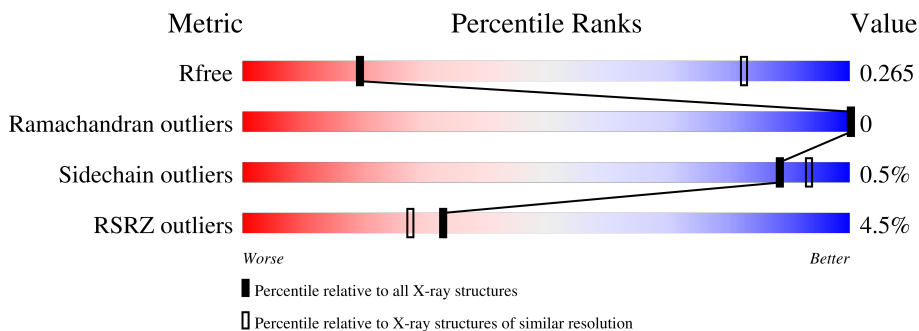
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.12 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}	130704	1024 (4.50-3.74)
Ramachandran outliers	138981	1043 (4.50-3.74)
Sidechain outliers	138945	1030 (4.50-3.74)
RSRZ outliers	127900	1041 (4.54-3.70)

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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. 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	130	
1	O	130	
1	P	130	
1	Q	130	
1	R	130	
1	S	130	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	T	130	<p>% 50% 49%</p>
1	U	130	<p>3% 52% 47%</p>
1	V	130	<p>4% 50% 49%</p>
1	W	130	<p>% 49% 49%</p>
1	X	130	<p>7% 49% 50%</p>
1	Y	130	<p>2% 50% 50%</p>
1	Z	130	<p>2% 48% 51%</p>
1	a	130	<p>4% 53% 46%</p>
2	C	5	<p>100%</p>
3	B	33	<p>100%</p>
4	D	34	<p>91% 6%</p>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8000 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 Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	V	66	552	350	94	106	2	0	0	0
1	Z	64	535	341	91	101	2	0	0	0
1	Y	65	543	345	92	104	2	0	0	0
1	T	66	552	350	94	106	2	0	0	0
1	P	65	543	345	92	104	2	0	0	0
1	S	64	535	341	91	101	2	0	0	0
1	X	65	543	345	92	104	2	0	0	0
1	W	66	552	350	94	106	2	0	0	0
1	Q	65	543	345	92	104	2	0	0	0
1	R	63	521	330	89	100	2	0	0	0
1	U	69	572	360	97	113	2	0	0	0
1	a	70	565	359	97	107	2	0	0	0
1	O	65	543	345	92	104	2	0	0	0
1	A	66	552	350	94	106	2	0	0	0

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
V	568	GLY	-	expression tag	UNP A0A1Q1NMU1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
V	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
V	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
V	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
V	572	SER	-	expression tag	UNP A0A1Q1NMMU1
Z	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
Z	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
Z	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
Z	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
Z	572	SER	-	expression tag	UNP A0A1Q1NMMU1
Y	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
Y	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
Y	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
Y	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
Y	572	SER	-	expression tag	UNP A0A1Q1NMMU1
T	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
T	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
T	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
T	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
T	572	SER	-	expression tag	UNP A0A1Q1NMMU1
P	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
P	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
P	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
P	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
P	572	SER	-	expression tag	UNP A0A1Q1NMMU1
S	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
S	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
S	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
S	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
S	572	SER	-	expression tag	UNP A0A1Q1NMMU1
X	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
X	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
X	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
X	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
X	572	SER	-	expression tag	UNP A0A1Q1NMMU1
W	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
W	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
W	570	LEU	-	expression tag	UNP A0A1Q1NMMU1
W	571	GLY	-	expression tag	UNP A0A1Q1NMMU1
W	572	SER	-	expression tag	UNP A0A1Q1NMMU1
Q	568	GLY	-	expression tag	UNP A0A1Q1NMMU1
Q	569	PRO	-	expression tag	UNP A0A1Q1NMMU1
Q	570	LEU	-	expression tag	UNP A0A1Q1NMMU1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
Q	571	GLY	-	expression tag	UNP A0A1Q1NMU1
Q	572	SER	-	expression tag	UNP A0A1Q1NMU1
R	568	GLY	-	expression tag	UNP A0A1Q1NMU1
R	569	PRO	-	expression tag	UNP A0A1Q1NMU1
R	570	LEU	-	expression tag	UNP A0A1Q1NMU1
R	571	GLY	-	expression tag	UNP A0A1Q1NMU1
R	572	SER	-	expression tag	UNP A0A1Q1NMU1
U	568	GLY	-	expression tag	UNP A0A1Q1NMU1
U	569	PRO	-	expression tag	UNP A0A1Q1NMU1
U	570	LEU	-	expression tag	UNP A0A1Q1NMU1
U	571	GLY	-	expression tag	UNP A0A1Q1NMU1
U	572	SER	-	expression tag	UNP A0A1Q1NMU1
a	568	GLY	-	expression tag	UNP A0A1Q1NMU1
a	569	PRO	-	expression tag	UNP A0A1Q1NMU1
a	570	LEU	-	expression tag	UNP A0A1Q1NMU1
a	571	GLY	-	expression tag	UNP A0A1Q1NMU1
a	572	SER	-	expression tag	UNP A0A1Q1NMU1
O	568	GLY	-	expression tag	UNP A0A1Q1NMU1
O	569	PRO	-	expression tag	UNP A0A1Q1NMU1
O	570	LEU	-	expression tag	UNP A0A1Q1NMU1
O	571	GLY	-	expression tag	UNP A0A1Q1NMU1
O	572	SER	-	expression tag	UNP A0A1Q1NMU1
A	568	GLY	-	expression tag	UNP A0A1Q1NMU1
A	569	PRO	-	expression tag	UNP A0A1Q1NMU1
A	570	LEU	-	expression tag	UNP A0A1Q1NMU1
A	571	GLY	-	expression tag	UNP A0A1Q1NMU1
A	572	SER	-	expression tag	UNP A0A1Q1NMU1

- Molecule 2 is a protein called C-terminal domain of Mengla nucleoprotein.

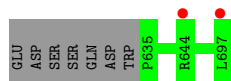
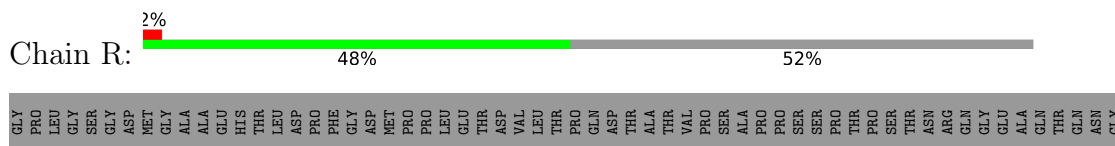
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	5	Total	C	N	O	0	0	0
			25	15	5	5			

- Molecule 3 is a protein called C-terminal domain of Mengla nucleoprotein.

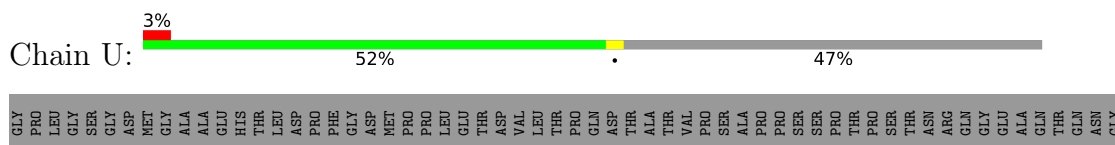
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	B	33	Total	C	N	O	0	0	0
			164	98	33	33			

- Molecule 4 is a protein called C-terminal domain of Mengla nucleoprotein.

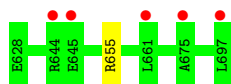
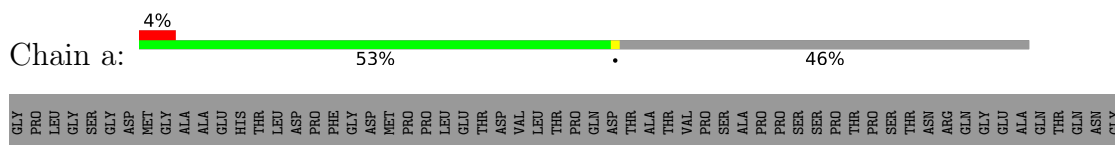
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	D	32	160	96	32	32	0	0	0



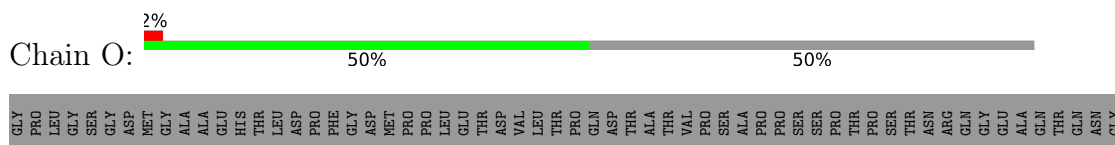
- Molecule 1: Nucleoprotein



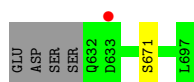
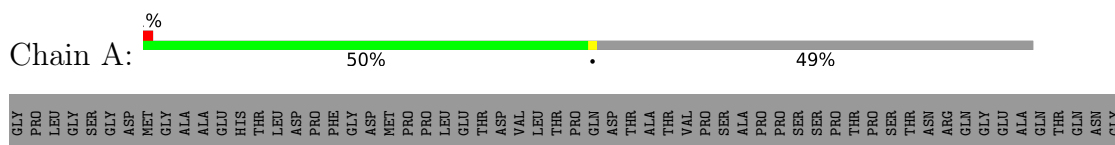
- Molecule 1: Nucleoprotein



- Molecule 1: Nucleoprotein



- Molecule 1: Nucleoprotein



- Molecule 2: C-terminal domain of Mengla nucleoprotein

Chain C:  100%

There are no outlier residues recorded for this chain.

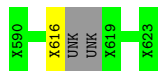
- Molecule 3: C-terminal domain of Mengla nucleoprotein

Chain B:  100%

There are no outlier residues recorded for this chain.

- Molecule 4: C-terminal domain of Mengla nucleoprotein

Chain D:  91% • 6%



4 Data and refinement statistics i

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	177.18Å 177.18Å 90.95Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	48.90 – 4.12 48.90 – 4.12	Depositor EDS
% Data completeness (in resolution range)	51.0 (48.90-4.12) 51.1 (48.90-4.12)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 4.14Å)	Xtrriage
Refinement program	REFMAC dev_4788, PHENIX dev_4788	Depositor
R, R_{free}	0.236 , 0.268 0.245 , 0.265	Depositor DCC
R_{free} test set	672 reflections (5.35%)	wwPDB-VP
Wilson B-factor (Å ²)	153.6	Xtrriage
Anisotropy	0.124	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 188.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	0.057 for -h,-k,l 0.328 for h,-h-k,-l 0.046 for -k,-h,-l	Xtrriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	8000	wwPDB-VP
Average B, all atoms (Å ²)	176.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.27% 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](#)

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	0.29	0/567	0.57	0/770
1	O	0.26	0/558	0.53	0/758
1	P	0.26	0/558	0.48	0/758
1	Q	0.28	0/558	0.52	0/758
1	R	0.27	0/534	0.52	0/723
1	S	0.36	0/550	0.54	0/747
1	T	0.39	1/567 (0.2%)	0.55	0/770
1	U	0.29	0/587	0.61	1/797 (0.1%)
1	V	0.27	0/567	0.59	1/770 (0.1%)
1	W	0.26	0/567	0.49	0/770
1	X	0.26	0/558	0.54	0/758
1	Y	0.27	0/558	0.53	0/758
1	Z	0.26	0/550	0.55	0/747
1	a	0.26	0/580	0.52	0/789
All	All	0.29	1/7859 (0.0%)	0.54	2/10673 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	T	635	PRO	N-CD	6.34	1.56	1.47

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	U	672	PRO	N-CA-C	5.59	126.63	112.10
1	V	633	ASP	CB-CG-OD2	5.12	122.91	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	64/130 (49%)	60 (94%)	4 (6%)	0	100	100
1	O	63/130 (48%)	60 (95%)	3 (5%)	0	100	100
1	P	63/130 (48%)	61 (97%)	2 (3%)	0	100	100
1	Q	63/130 (48%)	59 (94%)	4 (6%)	0	100	100
1	R	61/130 (47%)	58 (95%)	3 (5%)	0	100	100
1	S	62/130 (48%)	58 (94%)	4 (6%)	0	100	100
1	T	64/130 (49%)	58 (91%)	6 (9%)	0	100	100
1	U	67/130 (52%)	62 (92%)	5 (8%)	0	100	100
1	V	64/130 (49%)	60 (94%)	4 (6%)	0	100	100
1	W	64/130 (49%)	62 (97%)	2 (3%)	0	100	100
1	X	63/130 (48%)	59 (94%)	4 (6%)	0	100	100
1	Y	63/130 (48%)	58 (92%)	5 (8%)	0	100	100
1	Z	62/130 (48%)	59 (95%)	3 (5%)	0	100	100
1	a	68/130 (52%)	63 (93%)	5 (7%)	0	100	100
All	All	891/1820 (49%)	837 (94%)	54 (6%)	0	100	100

There are no Ramachandran outliers to report.

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	62/114 (54%)	62 (100%)	0	100	100
1	O	61/114 (54%)	61 (100%)	0	100	100
1	P	61/114 (54%)	61 (100%)	0	100	100
1	Q	61/114 (54%)	61 (100%)	0	100	100
1	R	59/114 (52%)	59 (100%)	0	100	100
1	S	60/114 (53%)	60 (100%)	0	100	100
1	T	62/114 (54%)	62 (100%)	0	100	100
1	U	65/114 (57%)	65 (100%)	0	100	100
1	V	62/114 (54%)	62 (100%)	0	100	100
1	W	62/114 (54%)	60 (97%)	2 (3%)	39	61
1	X	61/114 (54%)	60 (98%)	1 (2%)	62	78
1	Y	61/114 (54%)	61 (100%)	0	100	100
1	Z	60/114 (53%)	59 (98%)	1 (2%)	60	77
1	a	60/114 (53%)	60 (100%)	0	100	100
All	All	857/1596 (54%)	853 (100%)	4 (0%)	88	93

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	Z	636	ARG
1	X	634	TRP
1	W	639	LYS
1	W	647	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	V	696	ASN
1	T	659	GLN
1	P	696	ASN
1	U	659	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 monosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	D	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	620:UNK	C	621:UNK	N	1.60

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	66/130 (50%)	0.40	1 (1%) 73 63	96, 155, 230, 270	0
1	O	65/130 (50%)	0.29	3 (4%) 32 27	120, 187, 241, 288	0
1	P	65/130 (50%)	0.34	0 100 100	99, 166, 238, 268	0
1	Q	65/130 (50%)	0.48	4 (6%) 20 17	92, 184, 237, 244	0
1	R	63/130 (48%)	0.28	2 (3%) 47 37	108, 166, 211, 226	0
1	S	64/130 (49%)	0.36	1 (1%) 72 62	108, 167, 217, 261	0
1	T	66/130 (50%)	0.33	1 (1%) 73 63	84, 177, 230, 260	0
1	U	69/130 (53%)	0.54	4 (5%) 23 19	94, 151, 261, 342	0
1	V	66/130 (50%)	0.38	5 (7%) 13 12	128, 184, 237, 392	0
1	W	66/130 (50%)	0.26	1 (1%) 73 63	109, 173, 228, 275	0
1	X	65/130 (50%)	0.64	9 (13%) 2 3	116, 180, 245, 278	0
1	Y	65/130 (50%)	0.36	2 (3%) 49 38	121, 176, 220, 250	0
1	Z	64/130 (49%)	0.36	3 (4%) 31 26	107, 166, 232, 252	0
1	a	70/130 (53%)	0.46	5 (7%) 16 13	100, 170, 243, 270	0
2	C	0/5	-	-	-	-
3	B	0/33	-	-	-	-
4	D	0/34	-	-	-	-
All	All	919/1892 (48%)	0.39	41 (4%) 33 27	84, 173, 240, 392	0

The worst 5 of 41 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	U	631	SER	4.8
1	T	675	ALA	3.3
1	a	645	GLU	3.1
1	a	675	ALA	3.1
1	R	697	LEU	2.9

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 monosaccharides 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.