Full wwPDB X-ray Structure Validation Report ⓘ

Jan 30, 2024 – 12:09 pm GMT

PDB ID	:	8Q8K
Title	:	KI Polyomavirus LTA NLS bound to importin alpha 2
Authors	:	Cross, E.M.; Forwood, J.K.; Alvisi, G.
Deposited on	:	2023-08-18
Resolution	:	2.70 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 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:

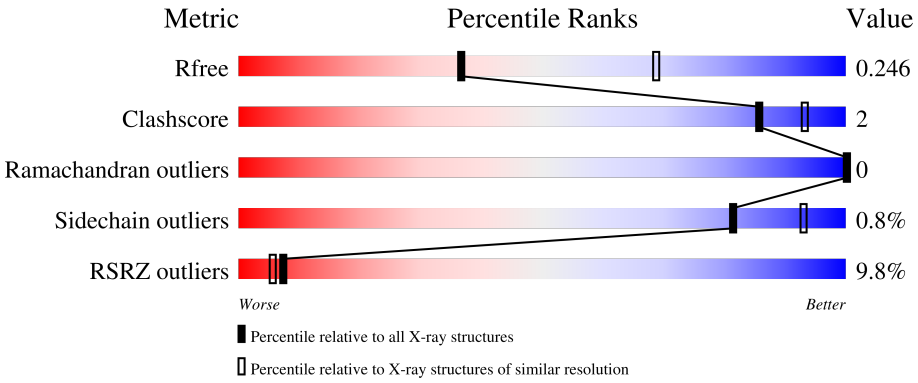
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.70 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	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 <=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	510	6% 78% 5% 17%
1	B	510	8% 78% 5% 17%
2	C	22	36% 50% 5% 45%
2	D	22	36% 55% 45%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12668 atoms, of which 6314 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 Importin subunit alpha-1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	423	Total	C	H	N	O	S	0	0	0
			6290	2001	3149	534	596	10			
1	B	421	Total	C	H	N	O	S	0	0	0
			5970	1921	2943	523	574	9			

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MET	-	initiating methionine	UNP P52293
A	21	HIS	-	expression tag	UNP P52293
A	22	HIS	-	expression tag	UNP P52293
A	23	HIS	-	expression tag	UNP P52293
A	24	HIS	-	expression tag	UNP P52293
A	25	HIS	-	expression tag	UNP P52293
A	26	HIS	-	expression tag	UNP P52293
A	27	SER	-	expression tag	UNP P52293
A	28	SER	-	expression tag	UNP P52293
A	29	GLY	-	expression tag	UNP P52293
A	30	LEU	-	expression tag	UNP P52293
A	31	VAL	-	expression tag	UNP P52293
A	32	PRO	-	expression tag	UNP P52293
A	33	ARG	-	expression tag	UNP P52293
A	34	GLY	-	expression tag	UNP P52293
A	35	SER	-	expression tag	UNP P52293
A	36	GLY	-	expression tag	UNP P52293
A	37	MET	-	expression tag	UNP P52293
A	38	LEU	-	expression tag	UNP P52293
A	39	GLU	-	expression tag	UNP P52293
A	40	THR	-	expression tag	UNP P52293
A	41	ALA	-	expression tag	UNP P52293
A	42	ALA	-	expression tag	UNP P52293
A	43	ALA	-	expression tag	UNP P52293
A	44	LEU	-	expression tag	UNP P52293

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	45	PHE	-	expression tag	UNP P52293
A	46	GLU	-	expression tag	UNP P52293
A	47	ARG	-	expression tag	UNP P52293
A	48	ASN	-	expression tag	UNP P52293
A	49	HIS	-	expression tag	UNP P52293
A	50	MET	-	expression tag	UNP P52293
A	51	ASP	-	expression tag	UNP P52293
A	52	SER	-	expression tag	UNP P52293
A	53	PRO	-	expression tag	UNP P52293
A	54	ASP	-	expression tag	UNP P52293
A	55	LEU	-	expression tag	UNP P52293
A	56	GLY	-	expression tag	UNP P52293
A	57	THR	-	expression tag	UNP P52293
A	58	ASP	-	expression tag	UNP P52293
A	59	ASP	-	expression tag	UNP P52293
A	60	ASP	-	expression tag	UNP P52293
A	61	ASP	-	expression tag	UNP P52293
A	62	LEU	-	expression tag	UNP P52293
A	63	ALA	-	expression tag	UNP P52293
A	64	MET	-	expression tag	UNP P52293
A	65	ALA	-	expression tag	UNP P52293
A	66	ASP	-	expression tag	UNP P52293
A	67	ILE	-	expression tag	UNP P52293
A	68	GLY	-	expression tag	UNP P52293
A	69	SER	-	expression tag	UNP P52293
B	20	MET	-	initiating methionine	UNP P52293
B	21	HIS	-	expression tag	UNP P52293
B	22	HIS	-	expression tag	UNP P52293
B	23	HIS	-	expression tag	UNP P52293
B	24	HIS	-	expression tag	UNP P52293
B	25	HIS	-	expression tag	UNP P52293
B	26	HIS	-	expression tag	UNP P52293
B	27	SER	-	expression tag	UNP P52293
B	28	SER	-	expression tag	UNP P52293
B	29	GLY	-	expression tag	UNP P52293
B	30	LEU	-	expression tag	UNP P52293
B	31	VAL	-	expression tag	UNP P52293
B	32	PRO	-	expression tag	UNP P52293
B	33	ARG	-	expression tag	UNP P52293
B	34	GLY	-	expression tag	UNP P52293
B	35	SER	-	expression tag	UNP P52293
B	36	GLY	-	expression tag	UNP P52293

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	37	MET	-	expression tag	UNP P52293
B	38	LEU	-	expression tag	UNP P52293
B	39	GLU	-	expression tag	UNP P52293
B	40	THR	-	expression tag	UNP P52293
B	41	ALA	-	expression tag	UNP P52293
B	42	ALA	-	expression tag	UNP P52293
B	43	ALA	-	expression tag	UNP P52293
B	44	LEU	-	expression tag	UNP P52293
B	45	PHE	-	expression tag	UNP P52293
B	46	GLU	-	expression tag	UNP P52293
B	47	ARG	-	expression tag	UNP P52293
B	48	ASN	-	expression tag	UNP P52293
B	49	HIS	-	expression tag	UNP P52293
B	50	MET	-	expression tag	UNP P52293
B	51	ASP	-	expression tag	UNP P52293
B	52	SER	-	expression tag	UNP P52293
B	53	PRO	-	expression tag	UNP P52293
B	54	ASP	-	expression tag	UNP P52293
B	55	LEU	-	expression tag	UNP P52293
B	56	GLY	-	expression tag	UNP P52293
B	57	THR	-	expression tag	UNP P52293
B	58	ASP	-	expression tag	UNP P52293
B	59	ASP	-	expression tag	UNP P52293
B	60	ASP	-	expression tag	UNP P52293
B	61	ASP	-	expression tag	UNP P52293
B	62	LEU	-	expression tag	UNP P52293
B	63	ALA	-	expression tag	UNP P52293
B	64	MET	-	expression tag	UNP P52293
B	65	ALA	-	expression tag	UNP P52293
B	66	ASP	-	expression tag	UNP P52293
B	67	ILE	-	expression tag	UNP P52293
B	68	GLY	-	expression tag	UNP P52293
B	69	SER	-	expression tag	UNP P52293

- Molecule 2 is a protein called Large T antigen.

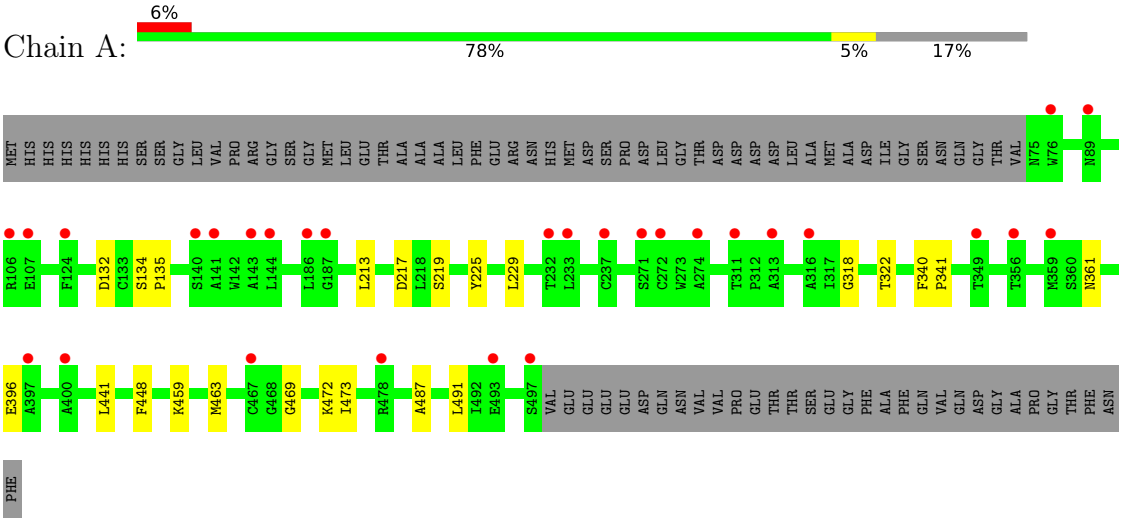
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	12	Total	C	H	N	O	0	0	0
			204	58	111	20	15			
2	D	12	Total	C	H	N	O	0	0	0
			204	58	111	20	15			

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

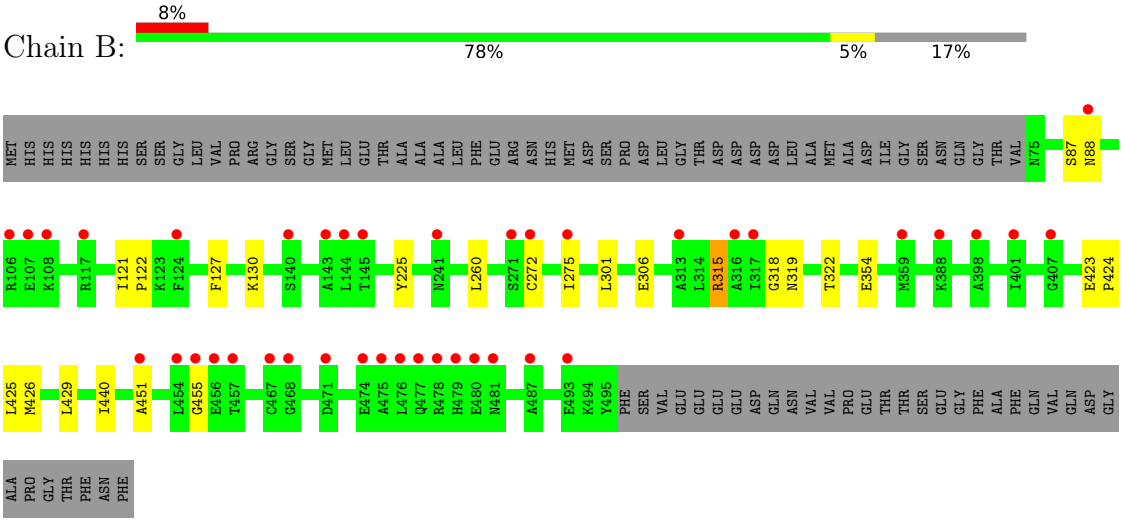
• Molecule 1: Importin subunit alpha-1

Chain A:



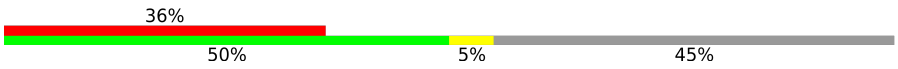
• Molecule 1: Importin subunit alpha-1

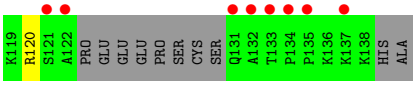
Chain B:



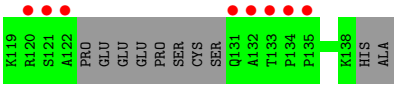
• Molecule 2: Large T antigen

Chain C:





- Molecule 2: Large T antigen



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	107.24Å 85.70Å 139.06Å 90.00° 95.96° 90.00°	Depositor
Resolution (Å)	19.88 – 2.70 19.88 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.2 (19.88-2.70) 99.2 (19.88-2.70)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.62 (at 2.71Å)	Xtriage
Refinement program	PHENIX 1.20.1-4487	Depositor
R, R_{free}	0.212 , 0.247 0.217 , 0.246	Depositor DCC
R_{free} test set	1703 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	67.9	Xtriage
Anisotropy	0.588	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 54.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12668	wwPDB-VP
Average B, all atoms (Å^2)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.24	0/3198	0.41	0/4370
1	B	0.26	0/3081	0.43	0/4225
2	C	0.24	0/93	0.48	0/120
2	D	0.23	0/93	0.51	0/120
All	All	0.25	0/6465	0.42	0/8835

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	3141	3149	3148	14	0
1	B	3027	2943	2942	14	0
2	C	93	111	107	1	0
2	D	93	111	107	0	0
All	All	6354	6314	6304	28	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 2.

All (28) 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:B:306:GLU:N	1:B:306:GLU:OE1	2.31	0.63
1:A:318:GLY:O	1:A:322:THR:HG23	1.99	0.62
1:A:396:GLU:OE1	2:C:120:ARG:NH2	2.35	0.60
1:A:213:LEU:HD13	1:A:229:LEU:HD13	1.85	0.59
1:B:260:LEU:HD13	1:B:275:ILE:HD12	1.87	0.56
1:A:459:LYS:O	1:A:463:MET:HG3	2.06	0.56
1:A:469:GLY:O	1:A:473:ILE:HD12	2.06	0.55
1:A:322:THR:HG22	1:A:361:ASN:ND2	2.24	0.52
1:B:426:MET:CE	1:B:429:LEU:HD12	2.42	0.50
1:A:469:GLY:O	1:A:472:LYS:N	2.45	0.49
1:B:318:GLY:O	1:B:322:THR:HG23	2.14	0.48
1:A:473:ILE:HD12	1:A:473:ILE:H	1.78	0.48
1:A:340:PHE:N	1:A:341:PRO:CD	2.79	0.46
1:B:426:MET:HE3	1:B:429:LEU:HD12	1.97	0.46
1:A:132:ASP:OD1	1:A:132:ASP:N	2.49	0.45
1:B:315:ARG:NH1	1:B:354:GLU:OE2	2.51	0.44
1:B:319:ASN:O	1:B:322:THR:OG1	2.31	0.44
1:B:423:GLU:HB3	1:B:424:PRO:HD3	2.00	0.43
1:B:272:CYS:SG	1:B:301:LEU:HD13	2.59	0.43
1:B:425:LEU:HG	1:B:440:ILE:HG23	2.00	0.42
1:B:87:SER:OG	1:B:88:ASN:N	2.53	0.42
1:A:134:SER:OG	1:A:135:PRO:HD3	2.20	0.42
1:A:217:ASP:OD1	1:A:219:SER:OG	2.25	0.41
1:A:134:SER:N	1:A:135:PRO:CD	2.83	0.41
1:A:487:ALA:O	1:A:491:LEU:HD13	2.19	0.41
1:B:127:PHE:HA	1:B:130:LYS:HG3	2.01	0.41
1:B:121:ILE:N	1:B:122:PRO:CD	2.84	0.41
1:B:451:ALA:O	1:B:455:GLY:N	2.53	0.41

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	421/510 (82%)	411 (98%)	10 (2%)	0	100	100
1	B	419/510 (82%)	412 (98%)	7 (2%)	0	100	100
2	C	8/22 (36%)	7 (88%)	1 (12%)	0	100	100
2	D	8/22 (36%)	8 (100%)	0	0	100	100
All	All	856/1064 (80%)	838 (98%)	18 (2%)	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	334/426 (78%)	331 (99%)	3 (1%)	78	92
1	B	305/426 (72%)	303 (99%)	2 (1%)	84	94
2	C	10/19 (53%)	10 (100%)	0	100	100
2	D	10/19 (53%)	10 (100%)	0	100	100
All	All	659/890 (74%)	654 (99%)	5 (1%)	81	93

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	225	TYR
1	A	441	LEU
1	A	448	PHE
1	B	225	TYR
1	B	315	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	352	GLN
1	A	375	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	427	ASN
1	B	137	GLN
1	B	297	GLN
1	B	427	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	423/510 (82%)	0.31	29 (6%)	16	15	54, 78, 110, 164	0
1	B	421/510 (82%)	0.41	40 (9%)	8	6	42, 77, 125, 138	0
2	C	12/22 (54%)	2.53	8 (66%)	0	0	82, 93, 105, 105	0
2	D	12/22 (54%)	2.54	8 (66%)	0	0	73, 91, 103, 113	0
All	All	868/1064 (81%)	0.42	85 (9%)	7	5	42, 78, 117, 164	0

All (85) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	133	THR	6.2
1	B	478	ARG	5.3
1	B	479	HIS	4.9
2	C	131	GLN	4.8
2	C	134	PRO	4.7
1	B	451	ALA	4.7
1	B	475	ALA	4.7
1	B	143	ALA	4.4
2	D	131	GLN	4.4
1	A	272	CYS	4.2
2	C	135	PRO	4.1
2	C	132	ALA	3.9
2	C	122	ALA	3.8
1	A	140	SER	3.8
1	B	476	LEU	3.7
1	A	497	SER	3.5
1	B	272	CYS	3.5
2	D	122	ALA	3.3
1	B	107	GLU	3.3
1	B	487	ALA	3.3
1	A	143	ALA	3.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	D	135	PRO	3.2
1	B	106	ARG	3.2
2	D	134	PRO	3.2
2	D	121	SER	3.2
1	A	232	THR	3.2
1	B	477	GLN	3.1
1	B	467	CYS	3.1
1	A	237	CYS	3.0
1	B	456	GLU	2.9
1	B	140	SER	2.9
1	A	349	THR	2.8
1	B	241	ASN	2.8
1	A	313	ALA	2.8
1	A	359	MET	2.8
1	A	107	GLU	2.8
1	B	144	LEU	2.8
1	A	397	ALA	2.7
1	B	455	GLY	2.7
1	A	316	ALA	2.6
1	B	480	GLU	2.6
1	A	144	LEU	2.6
1	B	388	LYS	2.6
1	B	359	MET	2.6
2	C	137	LYS	2.6
1	A	493	GLU	2.5
1	A	124	PHE	2.5
1	B	457	THR	2.5
1	B	407	GLY	2.5
1	B	468	GLY	2.5
1	B	108	LYS	2.5
1	A	106	ARG	2.4
1	A	467	CYS	2.4
1	B	317	ILE	2.4
1	B	481	ASN	2.4
1	B	401	ILE	2.4
1	A	141	ALA	2.4
1	A	271	SER	2.3
2	C	121	SER	2.3
1	B	471	ASP	2.3
1	A	233	LEU	2.3
1	B	117	ARG	2.2
1	B	313	ALA	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	478	ARG	2.2
1	B	88	ASN	2.2
1	A	274	ALA	2.2
2	D	132	ALA	2.2
1	A	311	THR	2.2
1	A	187	GLY	2.2
1	B	271	SER	2.2
1	A	76	TRP	2.2
1	B	493	GLU	2.2
1	B	316	ALA	2.2
1	B	275	ILE	2.1
1	A	400	ALA	2.1
1	B	398	ALA	2.1
1	B	124	PHE	2.1
2	D	120	ARG	2.1
1	B	145	THR	2.1
2	C	133	THR	2.1
1	A	89	ASN	2.0
1	A	186	LEU	2.0
1	B	454	LEU	2.0
1	B	474	GLU	2.0
1	A	356	THR	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 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.