Full wwPDB X-ray Structure Validation Report ⓘ

Jun 13, 2024 – 10:42 AM EDT

PDB ID	:	4JPP
Title	:	Bacteriophage phiX174 H protein residues 143-282
Authors	:	Sun, L.; Young, L.N.; Boudko, S.B.; Fokine, A.; Zhang, X.; Rossmann, M.G.; Fane, B.A.
Deposited on	:	2013-03-19
Resolution	:	2.40 Å(reported)

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

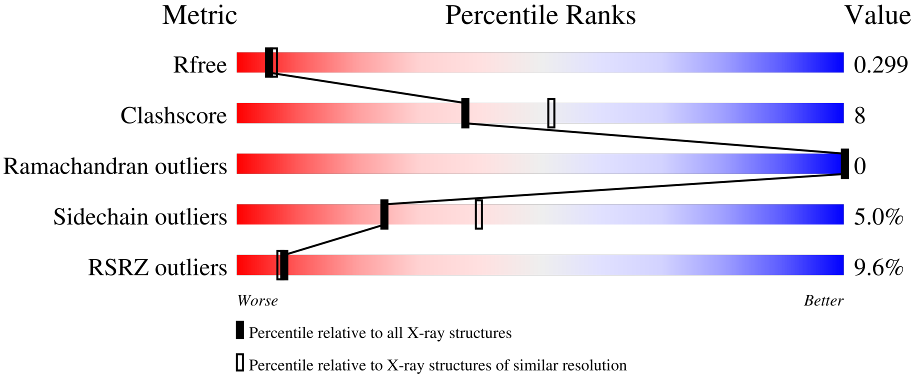
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance ⓘ

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

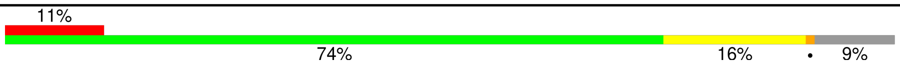
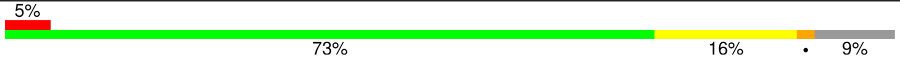
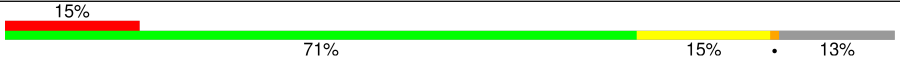
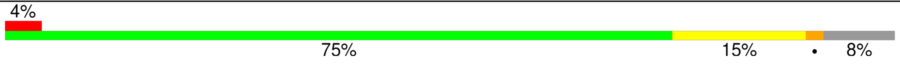
The reported resolution of this entry is 2.40 Å.

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	140	
1	B	140	
1	C	140	
1	D	140	
1	E	140	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5545 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 Minor spike protein H.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf	Trace
1	A	127	Total C N O S 1018 612 185 214 7	0	0	0
1	B	127	Total C N O S 1023 616 185 215 7	0	0	0
1	C	122	Total C N O S 978 589 176 206 7	0	0	0
1	D	129	Total C N O S 1034 623 187 217 7	0	0	0
1	E	121	Total C N O S 974 587 175 205 7	0	0	0

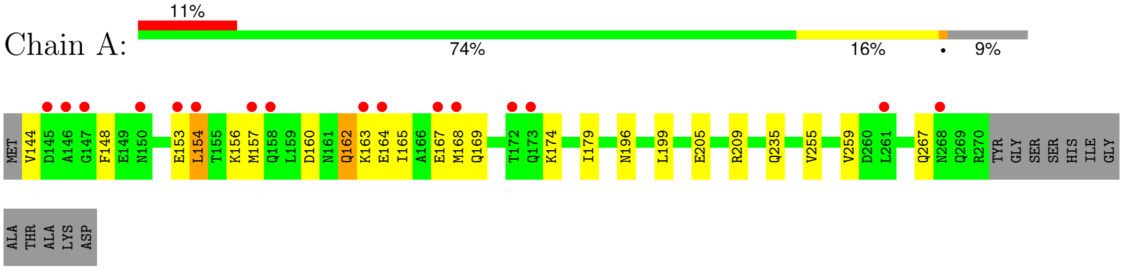
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	115	Total O 115 115	0	0
2	B	98	Total O 98 98	0	0
2	C	95	Total O 95 95	0	0
2	D	103	Total O 103 103	0	0
2	E	107	Total O 107 107	0	0

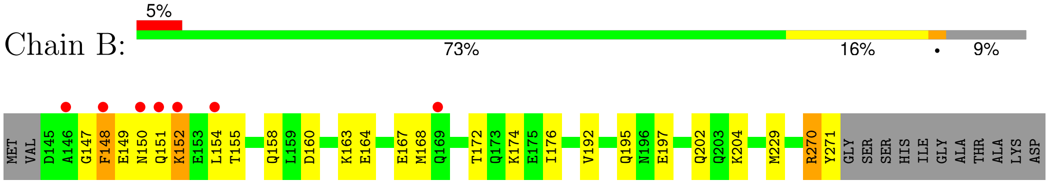
3 Residue-property plots

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.

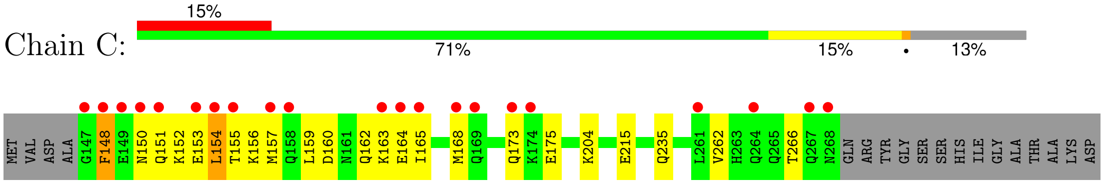
- Molecule 1: Minor spike protein H



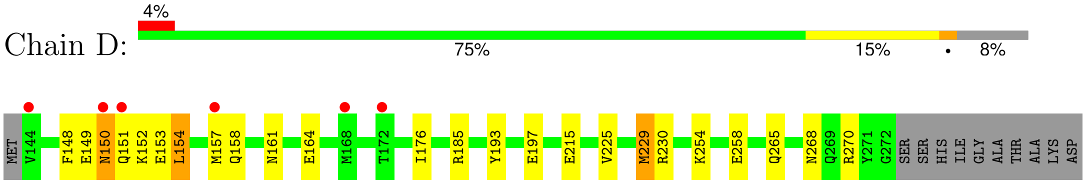
- Molecule 1: Minor spike protein H



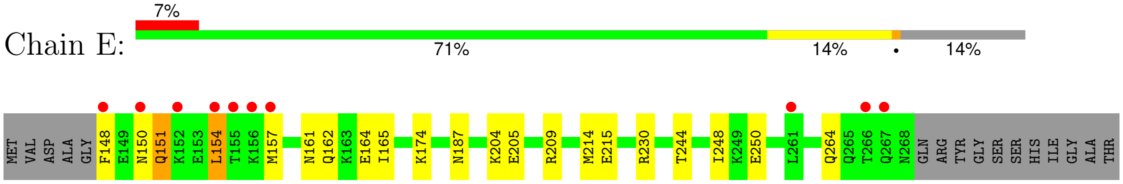
- Molecule 1: Minor spike protein H



- Molecule 1: Minor spike protein H



- Molecule 1: Minor spike protein H




ALA
LYS
ASP

4 Data and refinement statistics

Property	**Value**	**Source**
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	68.70Å 68.70Å 372.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.40 29.82 – 2.40	Depositor EDS
% Data completeness (in resolution range)	92.8 (8.00-2.40) 95.6 (29.82-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	4.16 (at 2.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.245 , 0.290 0.258 , 0.299	Depositor DCC
R_{free} test set	1737 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å^2)	16.3	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 65.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.54$, $< L^2 > = 0.38$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	5545	wwPDB-VP
Average B, all atoms (Å^2)	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.31% 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	0.61	0/1023	0.63	0/1371
1	B	0.52	0/1029	0.62	0/1379
1	C	0.58	0/983	0.62	0/1317
1	D	0.57	0/1040	0.60	0/1394
1	E	0.58	0/979	0.64	0/1312
All	All	0.57	0/5054	0.62	0/6773

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	1018	0	997	15	0
1	B	1023	0	997	23	0
1	C	978	0	958	19	0
1	D	1034	0	1009	21	0
1	E	974	0	955	17	0
2	A	115	0	0	5	0
2	B	98	0	0	3	0
2	C	95	0	0	1	0
2	D	103	0	0	4	0
2	E	107	0	0	5	0
All	All	5545	0	4916	77	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 8.

All (77) 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:150:ASN:OD1	1:B:151:GLN:N	2.08	0.86
2:A:395:HOH:O	1:B:229:MET:SD	2.37	0.82
1:A:154:LEU:HA	1:A:157:MET:HE2	1.66	0.78
1:B:155:THR:HG21	1:C:154:LEU:HD23	1.70	0.71
1:D:157:MET:HG3	2:D:383:HOH:O	1.95	0.65
1:D:270:ARG:HE	1:E:264:GLN:NE2	1.94	0.65
1:D:230:ARG:NH2	2:D:364:HOH:O	2.30	0.64
1:B:148:PHE:CE1	1:C:150:ASN:HB3	2.33	0.64
1:C:163:LYS:NZ	1:D:164:GLU:OE1	2.28	0.62
1:E:154:LEU:O	1:E:157:MET:HG2	2.00	0.61
1:C:159:LEU:HD13	1:D:161:ASN:HB2	1.81	0.61
1:E:174:LYS:NZ	2:E:380:HOH:O	2.32	0.61
1:C:159:LEU:O	1:D:161:ASN:ND2	2.34	0.60
1:A:165:ILE:HA	1:A:168:MET:HE2	1.83	0.60
1:A:196:ASN:O	1:A:199:LEU:HB3	2.02	0.60
1:C:148:PHE:O	1:C:151:GLN:HB2	2.04	0.58
2:B:357:HOH:O	1:C:164:GLU:HB2	2.03	0.58
1:E:148:PHE:N	1:E:151:GLN:HE21	2.03	0.56
1:C:173:GLN:HB3	1:D:176:ILE:HD11	1.88	0.55
1:C:151:GLN:HB3	1:D:154:LEU:HD11	1.88	0.55
2:A:367:HOH:O	1:B:204:LYS:HD2	2.07	0.54
1:A:235:GLN:HB2	2:A:395:HOH:O	2.07	0.54
1:A:205:GLU:HG2	1:A:209:ARG:NH1	2.23	0.53
1:B:154:LEU:O	1:B:158:GLN:HG3	2.08	0.53
1:B:155:THR:HG21	1:C:154:LEU:CD2	2.39	0.53
1:B:160:ASP:O	1:B:164:GLU:HG2	2.10	0.52
1:D:265:GLN:HA	2:D:390:HOH:O	2.10	0.52
1:A:162:GLN:NE2	2:A:386:HOH:O	2.42	0.52
1:D:150:ASN:OD1	1:D:151:GLN:HG2	2.09	0.52
1:E:161:ASN:O	1:E:165:ILE:HD12	2.10	0.51
1:E:214:MET:SD	2:E:321:HOH:O	2.59	0.51
1:E:161:ASN:O	1:E:164:GLU:HG2	2.10	0.51
1:A:174:LYS:HA	1:B:176:ILE:HD11	1.93	0.50
1:B:270:ARG:HG2	1:B:271:TYR:CZ	2.46	0.50
1:E:161:ASN:HA	1:E:164:GLU:HG2	1.94	0.50
1:B:152:LYS:O	1:B:155:THR:HB	2.13	0.49
1:C:159:LEU:HD11	1:D:158:GLN:HA	1.93	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:185:ARG:HD2	1:E:187:ASN:OD1	2.12	0.49
1:D:268:ASN:HB2	2:D:390:HOH:O	2.11	0.49
1:D:149:GLU:OE2	1:D:152:LYS:NZ	2.31	0.47
1:B:270:ARG:HG2	1:B:271:TYR:CE1	2.50	0.47
1:D:148:PHE:CD1	1:E:150:ASN:HB3	2.49	0.47
1:C:148:PHE:HZ	1:D:153:GLU:OE2	1.98	0.47
1:B:147:GLY:O	1:B:150:ASN:ND2	2.49	0.47
1:E:230:ARG:HD3	2:E:401:HOH:O	2.15	0.46
1:A:163:LYS:NZ	1:A:167:GLU:OE2	2.41	0.46
1:D:225:VAL:O	1:D:229:MET:HG2	2.16	0.45
1:B:152:LYS:HB3	1:B:152:LYS:HE2	1.32	0.45
1:E:250:GLU:OE2	2:E:310:HOH:O	2.21	0.45
1:B:147:GLY:O	1:B:150:ASN:CG	2.55	0.45
1:B:172:THR:O	1:B:176:ILE:HG12	2.17	0.45
1:D:254:LYS:HE2	1:D:258:GLU:OE2	2.17	0.44
1:A:163:LYS:NZ	1:B:168:MET:SD	2.87	0.44
1:A:174:LYS:HG2	1:B:176:ILE:HD11	2.00	0.43
1:C:235:GLN:HG2	2:C:377:HOH:O	2.18	0.43
1:B:174:LYS:NZ	1:C:175:GLU:OE1	2.50	0.43
1:E:204:LYS:HE2	1:E:204:LYS:HB3	1.86	0.43
1:E:205:GLU:HG2	1:E:209:ARG:NH1	2.34	0.43
1:E:244:THR:O	1:E:248:ILE:HG13	2.18	0.43
1:C:152:LYS:HE2	1:C:152:LYS:HB3	1.77	0.42
1:B:202:GLN:HA	2:B:358:HOH:O	2.18	0.42
1:A:255:VAL:O	1:A:259:VAL:HG23	2.19	0.42
1:B:197:GLU:OE1	2:B:390:HOH:O	2.22	0.42
1:B:270:ARG:HE	1:B:270:ARG:HB2	1.52	0.42
1:A:169:GLN:HB2	2:A:378:HOH:O	2.20	0.42
1:C:262:VAL:O	1:C:266:THR:HG23	2.20	0.41
1:C:204:LYS:HE2	1:C:204:LYS:HB3	1.83	0.41
1:A:156:LYS:NZ	1:A:160:ASP:OD2	2.50	0.41
1:A:153:GLU:O	1:A:157:MET:HG3	2.20	0.41
1:D:193:TYR:O	1:D:197:GLU:HG2	2.21	0.41
1:E:162:GLN:NE2	2:E:400:HOH:O	2.52	0.41
1:D:270:ARG:HE	1:E:264:GLN:HE22	1.64	0.41
1:D:229:MET:HG2	1:D:229:MET:H	1.72	0.41
1:C:160:ASP:O	1:C:164:GLU:HG2	2.21	0.40
1:C:165:ILE:HA	1:C:168:MET:HE2	2.03	0.40
1:A:168:MET:HE2	1:A:168:MET:HB2	1.88	0.40
1:B:192:VAL:O	1:B:195:GLN:HG2	2.22	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	125/140 (89%)	124 (99%)	1 (1%)	0	100	100
1	B	125/140 (89%)	123 (98%)	2 (2%)	0	100	100
1	C	120/140 (86%)	120 (100%)	0	0	100	100
1	D	127/140 (91%)	125 (98%)	2 (2%)	0	100	100
1	E	119/140 (85%)	116 (98%)	3 (2%)	0	100	100
All	All	616/700 (88%)	608 (99%)	8 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	113/122 (93%)	106 (94%)	7 (6%)	18	29
1	B	113/122 (93%)	107 (95%)	6 (5%)	22	37
1	C	109/122 (89%)	101 (93%)	8 (7%)	14	22
1	D	114/122 (93%)	110 (96%)	4 (4%)	36	55
1	E	109/122 (89%)	106 (97%)	3 (3%)	43	63
All	All	558/610 (92%)	530 (95%)	28 (5%)	24	40

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	144	VAL
1	A	148	PHE
1	A	154	LEU
1	A	162	GLN
1	A	164	GLU
1	A	179	ILE
1	A	267	GLN
1	B	148	PHE
1	B	149	GLU
1	B	152	LYS
1	B	163	LYS
1	B	167	GLU
1	B	270	ARG
1	C	148	PHE
1	C	153	GLU
1	C	154	LEU
1	C	155	THR
1	C	156	LYS
1	C	157	MET
1	C	162	GLN
1	C	215	GLU
1	D	150	ASN
1	D	154	LEU
1	D	215	GLU
1	D	229	MET
1	E	151	GLN
1	E	154	LEU
1	E	215	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	269	GLN
1	B	268	ASN
1	E	264	GLN

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	127/140 (90%)	0.52	16 (12%)	3	3	26, 42, 78, 90	0
1	B	127/140 (90%)	0.26	7 (5%)	25	24	27, 40, 81, 124	0
1	C	122/140 (87%)	0.74	21 (17%)	1	1	27, 42, 96, 139	0
1	D	129/140 (92%)	0.34	6 (4%)	31	30	27, 42, 82, 109	0
1	E	121/140 (86%)	0.42	10 (8%)	11	10	27, 42, 79, 103	0
All	All	626/700 (89%)	0.45	60 (9%)	8	7	26, 42, 82, 139	0

All (60) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	148	PHE	7.6
1	E	155	THR	7.1
1	A	168	MET	5.6
1	C	155	THR	5.0
1	C	157	MET	4.9
1	C	268	ASN	4.8
1	C	147	GLY	4.7
1	B	151	GLN	4.7
1	E	266	THR	4.5
1	C	154	LEU	4.4
1	C	264	GLN	4.3
1	A	154	LEU	4.2
1	C	150	ASN	4.1
1	A	146	ALA	4.0
1	C	148	PHE	3.9
1	C	158	GLN	3.9
1	A	150	ASN	3.7
1	C	267	GLN	3.7
1	C	151	GLN	3.6
1	B	148	PHE	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	157	MET	3.3
1	C	173	GLN	3.3
1	B	150	ASN	3.3
1	A	164	GLU	3.3
1	E	157	MET	3.3
1	D	168	MET	3.2
1	C	153	GLU	3.2
1	C	168	MET	3.1
1	A	172	THR	3.1
1	C	261	LEU	3.0
1	A	145	ASP	3.0
1	B	154	LEU	3.0
1	D	151	GLN	2.9
1	C	149	GLU	2.9
1	E	156	LYS	2.9
1	B	146	ALA	2.9
1	A	261	LEU	2.8
1	D	144	VAL	2.7
1	D	150	ASN	2.7
1	E	150	ASN	2.7
1	E	267	GLN	2.6
1	B	152	LYS	2.5
1	A	167	GLU	2.5
1	A	153	GLU	2.4
1	C	163	LYS	2.4
1	C	174	LYS	2.3
1	E	154	LEU	2.3
1	A	268	ASN	2.3
1	D	172	THR	2.3
1	E	152	LYS	2.2
1	D	157	MET	2.2
1	A	147	GLY	2.2
1	A	163	LYS	2.2
1	C	169	GLN	2.2
1	A	173	GLN	2.1
1	B	169	GLN	2.1
1	A	158	GLN	2.1
1	E	261	LEU	2.0
1	C	165	ILE	2.0
1	C	164	GLU	2.0

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.