Full wwPDB X-ray Structure Validation Report ⓘ

Oct 11, 2023 – 06:21 PM EDT

PDB ID	:	7RTR
Title	:	YLQ-SG3 TCR in complex with SARS-CoV-2 Spike-derived peptide S269-277 (YLQPRTFLL) presented by HLA-A*02:01
Authors	:	Szeto, C.; Gras, S.
Deposited on	:	2021-08-14
Resolution	:	2.60 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:

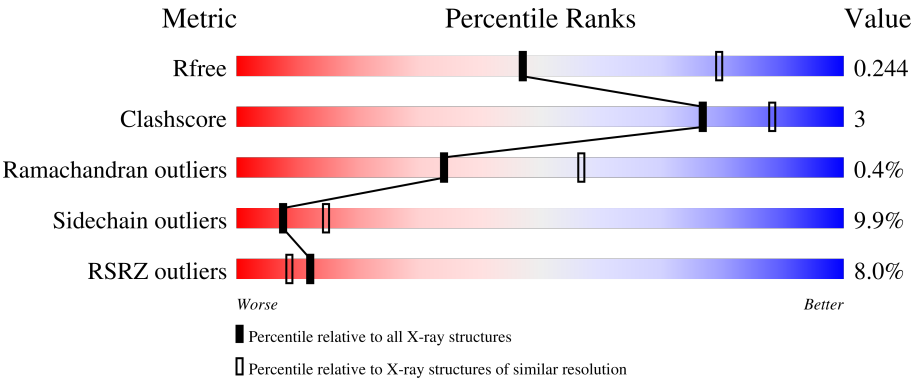
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	365	7% 63% 8% 27%
2	B	100	2% 88% 12%
3	C	9	56% 44%
4	D	203	10% 83% 12%
5	E	241	7% 83% 15%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6644 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 HLA class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	266	Total	C	N	O	S	0	9	0
			2160	1363	392	396	9			

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	100	Total	C	N	O	S	0	6	0
			864	551	145	164	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	9	Total	C	N	O	0	0	0
			82	56	13	13			

- Molecule 4 is a protein called YLQ-SG3 TCR alpha chain (TRAV12-2).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	197	Total	C	N	O	S	0	0	0
			1513	947	251	307	8			

- Molecule 5 is a protein called YLQ-SG3 TCR beta chain (TRBV7-9).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	239	Total	C	N	O	S	0	10	0
			1974	1239	353	375	7			

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	2	Total Na 2 2	0	0
6	D	2	Total Na 2 2	0	0

- Molecule 7 is water.

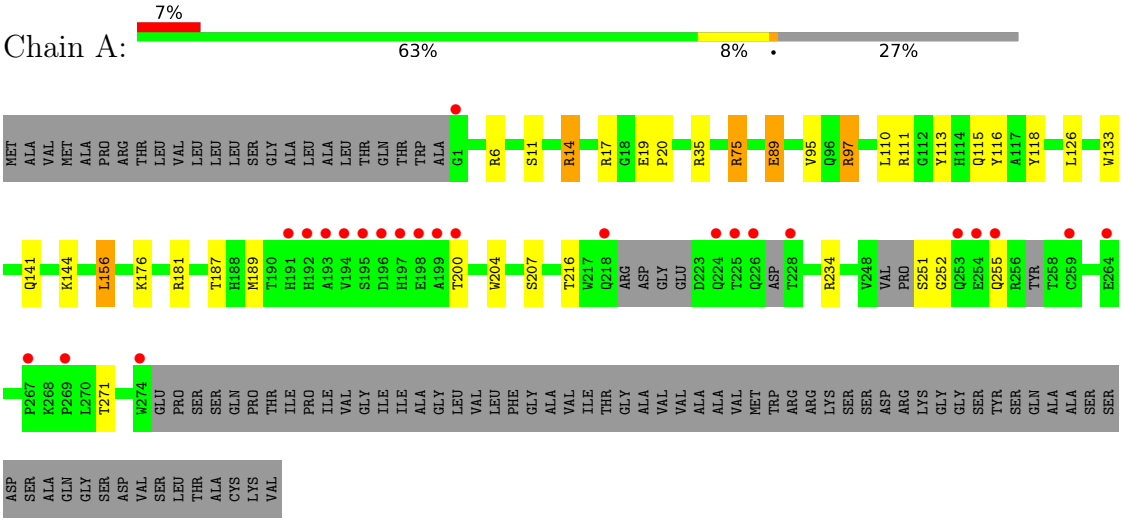
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	22	Total O 22 22	0	0
7	B	12	Total O 12 12	0	0
7	C	1	Total O 1 1	0	0
7	D	5	Total O 5 5	0	0
7	E	7	Total O 7 7	0	0

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.

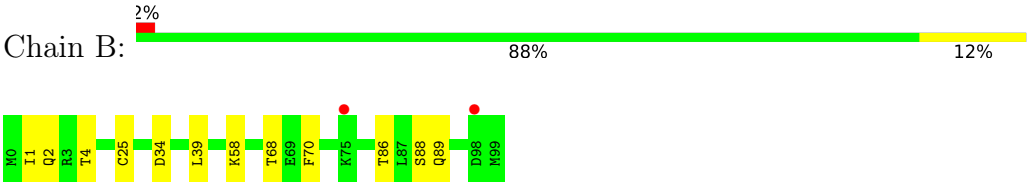
• Molecule 1: HLA class I antigen

Chain A:



• Molecule 2: Beta-2-microglobulin

Chain B:

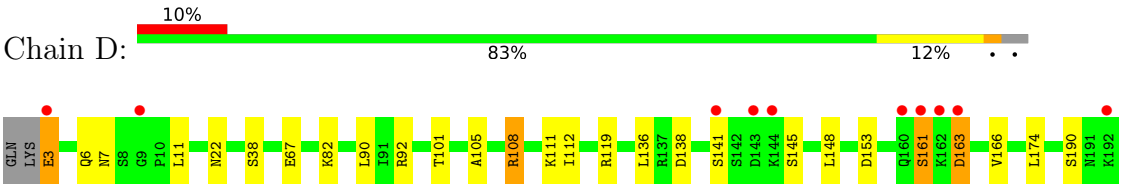


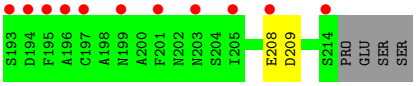
• Molecule 3: Spike protein S1

Chain C:

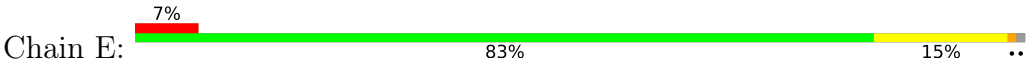
• Molecule 4: YLQ-SG3 TCR alpha chain (TRAV12-2)

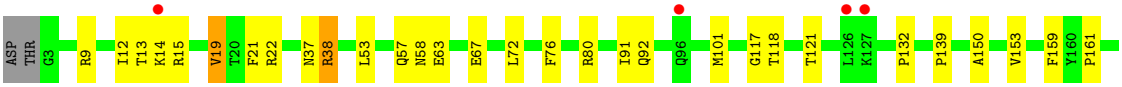
Chain D:

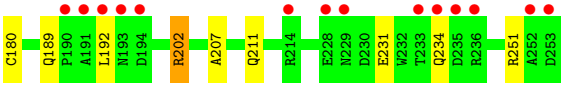




• Molecule 5: YLQ-SG3 TCR beta chain (TRBV7-9)







4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	225.36Å 49.62Å 91.72Å 90.00° 91.83° 90.00°	Depositor
Resolution (Å)	33.42 – 2.60 33.42 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (33.42-2.60) 100.0 (33.42-2.60)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.31 (at 2.61Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.191 , 0.235 0.199 , 0.244	Depositor DCC
R_{free} test set	1507 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å^2)	41.2	Xtriage
Anisotropy	0.929	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 67.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.010 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6644	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% 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

Bond lengths and bond angles in the following residue types are not validated in this section: NA

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	0.50	0/2244	0.72	0/3043
2	B	0.55	0/905	0.74	0/1224
3	C	0.63	0/84	0.81	0/112
4	D	0.57	0/1546	0.80	0/2100
5	E	0.49	0/2027	0.71	0/2758
All	All	0.52	0/6806	0.74	0/9237

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	2160	0	2003	10	0
2	B	864	0	846	4	0
3	C	82	0	88	0	0
4	D	1513	0	1398	9	0
5	E	1974	0	1845	16	0
6	A	2	0	0	0	0
6	D	2	0	0	0	0
7	A	22	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	12	0	0	0	0
7	C	1	0	0	0	0
7	D	5	0	0	0	0
7	E	7	0	0	0	0
All	All	6644	0	6180	37	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 (37) 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)
2:B:4[A]:THR:HG22	2:B:86:THR:HB	1.72	0.72
5:E:37:ASN:HD22	5:E:58:ASN:HD21	1.38	0.72
5:E:38:ARG:HH11	5:E:57:GLN:HE21	1.36	0.72
4:D:163:ASP:HB2	4:D:166:VAL:HG22	1.75	0.68
4:D:105:ALA:HB1	4:D:112:ILE:CG2	2.24	0.66
1:A:89:GLU:H	1:A:89:GLU:CD	2.02	0.61
4:D:3:GLU:HB2	4:D:108:ARG:HD2	1.81	0.61
4:D:138:ASP:HB3	4:D:141:SER:O	2.05	0.57
5:E:139:PRO:HG3	5:E:150:ALA:HB1	1.85	0.57
1:A:97:ARG:HG3	1:A:116:TYR:CE2	2.41	0.56
4:D:105:ALA:HB1	4:D:112:ILE:HG21	1.90	0.53
1:A:126:LEU:HD22	1:A:156:LEU:HD13	1.93	0.51
5:E:53:LEU:HD22	5:E:72:LEU:HD13	1.94	0.49
1:A:14:ARG:HB3	1:A:17:ARG:HB2	1.96	0.48
1:A:97:ARG:HG3	1:A:116:TYR:CZ	2.48	0.48
1:A:20:PRO:HD2	1:A:75:ARG:HH11	1.78	0.48
5:E:21:PHE:CD1	5:E:118:THR:HG21	2.49	0.47
4:D:108:ARG:O	4:D:111:LYS:HG2	2.14	0.47
5:E:132:PRO:HB3	5:E:159:PHE:CD2	2.50	0.47
2:B:25:CYS:HB2	2:B:39:LEU:HD21	1.98	0.45
4:D:136:LEU:HD22	5:E:139:PRO:HA	1.99	0.45
5:E:12:ILE:HD13	5:E:161:PRO:HG3	1.99	0.45
5:E:76:PHE:CE2	5:E:91:ILE:HG12	2.52	0.45
2:B:2:GLN:HB3	2:B:86:THR:HG22	2.00	0.44
5:E:121:THR:HG21	5:E:161:PRO:HB3	2.00	0.43
5:E:207:ALA:O	5:E:211:GLN:HG3	2.18	0.43
2:B:2:GLN:HB3	2:B:86:THR:CG2	2.49	0.43
5:E:13[A]:THR:HG21	5:E:19:VAL:HG13	1.99	0.43
5:E:153:VAL:HG22	5:E:202:ARG:HG3	2.00	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:95:VAL:HG12	1:A:118:TYR:HD1	1.84	0.42
4:D:166:VAL:HG12	4:D:190:SER:HB2	2.01	0.42
5:E:101[A]:MET:HE2	5:E:117:GLY:HA3	2.01	0.42
1:A:133:TRP:HB2	1:A:144:LYS:HD2	2.01	0.41
1:A:6[A]:ARG:HG2	1:A:113[A]:TYR:HE1	1.85	0.41
1:A:187:THR:HA	1:A:204:TRP:O	2.21	0.41
4:D:174:LEU:HB3	5:E:180:CYS:HB3	2.03	0.41
5:E:58:ASN:O	5:E:80:ARG:HD3	2.21	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	265/365 (73%)	254 (96%)	10 (4%)	1 (0%)	34	57
2	B	104/100 (104%)	100 (96%)	4 (4%)	0	100	100
3	C	7/9 (78%)	7 (100%)	0	0	100	100
4	D	195/203 (96%)	185 (95%)	9 (5%)	1 (0%)	29	52
5	E	247/241 (102%)	233 (94%)	13 (5%)	1 (0%)	34	57
All	All	818/918 (89%)	779 (95%)	36 (4%)	3 (0%)	34	57

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	161	SER
5	E	234	GLN
1	A	252	GLY

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	214/299 (72%)	190 (89%)	24 (11%)	6 10
2	B	101/95 (106%)	93 (92%)	8 (8%)	12 24
3	C	9/9 (100%)	5 (56%)	4 (44%)	0 0
4	D	167/183 (91%)	147 (88%)	20 (12%)	5 9
5	E	211/213 (99%)	197 (93%)	14 (7%)	16 33
All	All	702/799 (88%)	632 (90%)	70 (10%)	8 14

All (70) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	11	SER
1	A	14	ARG
1	A	19	GLU
1	A	35	ARG
1	A	75	ARG
1	A	89	GLU
1	A	97	ARG
1	A	110	LEU
1	A	111	ARG
1	A	115	GLN
1	A	141[A]	GLN
1	A	141[B]	GLN
1	A	156	LEU
1	A	176	LYS
1	A	181	ARG
1	A	189	MET
1	A	200	THR
1	A	207[A]	SER
1	A	207[B]	SER
1	A	216	THR
1	A	234	ARG
1	A	251	SER
1	A	255	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	271	THR
2	B	1	ILE
2	B	34	ASP
2	B	58	LYS
2	B	68	THR
2	B	70	PHE
2	B	88	SER
2	B	89[A]	GLN
2	B	89[B]	GLN
3	C	2	LEU
3	C	5	ARG
3	C	8	LEU
3	C	9	LEU
4	D	3	GLU
4	D	6	GLN
4	D	7	ASN
4	D	11	LEU
4	D	22	ASN
4	D	38	SER
4	D	67	GLU
4	D	82	LYS
4	D	90	LEU
4	D	92	ARG
4	D	101	THR
4	D	108	ARG
4	D	119	ARG
4	D	145	SER
4	D	148	LEU
4	D	153	ASP
4	D	161	SER
4	D	163	ASP
4	D	208	GLU
4	D	209	ASP
5	E	9	ARG
5	E	14	LYS
5	E	15	ARG
5	E	19	VAL
5	E	22	ARG
5	E	38	ARG
5	E	63	GLU
5	E	67	GLU
5	E	92	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	E	189	GLN
5	E	192	LEU
5	E	202	ARG
5	E	231	GLU
5	E	251	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
5	E	10	HIS
5	E	48	GLN
5	E	57	GLN
5	E	58	ASN
5	E	65	GLN
5	E	92	GLN
5	E	146	HIS

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)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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

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	266/365 (72%)	0.28	24 (9%) 9 6	25, 45, 100, 115	2 (0%)
2	B	100/100 (100%)	-0.11	2 (2%) 65 60	24, 40, 71, 81	0
3	C	9/9 (100%)	0.01	0 100 100	31, 32, 35, 37	0
4	D	197/203 (97%)	0.32	21 (10%) 6 3	29, 48, 96, 113	4 (2%)
5	E	239/241 (99%)	0.16	18 (7%) 14 10	28, 54, 88, 101	0
All	All	811/918 (88%)	0.20	65 (8%) 12 9	24, 48, 94, 115	6 (0%)

All (65) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	161	SER	7.7
4	D	141	SER	6.4
5	E	253	ASP	5.3
4	D	197	CYS	4.9
5	E	235	ASP	4.6
5	E	190	PRO	4.5
1	A	196	ASP	4.5
1	A	225	THR	4.4
5	E	191	ALA	4.3
1	A	254	GLU	4.3
4	D	208	GLU	4.3
1	A	193	ALA	4.1
4	D	162	LYS	4.1
4	D	196	ALA	4.1
5	E	127	LYS	4.0
1	A	191	HIS	3.8
5	E	228	GLU	3.7
1	A	199	ALA	3.7
5	E	193	ASN	3.4
1	A	194	VAL	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	163	ASP	3.3
1	A	253	GLN	3.3
4	D	203	ASN	3.3
1	A	274	TRP	3.3
1	A	228	THR	3.2
2	B	75	LYS	3.2
5	E	233	THR	3.2
4	D	205	ILE	3.2
1	A	226	GLN	3.1
1	A	197	HIS	3.1
4	D	143	ASP	3.1
1	A	255	GLN	3.1
4	D	9	GLY	3.0
1	A	195	SER	3.0
1	A	198	GLU	2.9
5	E	229	ASN	2.8
1	A	224	GLN	2.8
4	D	193	SER	2.8
4	D	199	ASN	2.8
5	E	126	LEU	2.7
5	E	234	GLN	2.7
1	A	192	HIS	2.6
4	D	194	ASP	2.6
5	E	192	LEU	2.6
1	A	259	CYS	2.6
4	D	3	GLU	2.5
4	D	214	SER	2.5
5	E	214	ARG	2.5
1	A	267	PRO	2.5
4	D	192	LYS	2.4
1	A	269	PRO	2.4
4	D	160	GLN	2.3
5	E	236	ARG	2.3
4	D	195	PHE	2.3
4	D	144	LYS	2.3
5	E	194	ASP	2.3
5	E	96	GLN	2.3
1	A	218	GLN	2.2
1	A	264	GLU	2.2
1	A	200	THR	2.2
2	B	98	ASP	2.2
1	A	1	GLY	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
4	D	201	PHE	2.2
5	E	252	ALA	2.2
5	E	14	LYS	2.2

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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	NA	A	402	1/1	0.89	0.70	53,53,53,53	0
6	NA	D	302	1/1	0.92	0.18	57,57,57,57	0
6	NA	A	401	1/1	0.94	0.21	48,48,48,48	0
6	NA	D	301	1/1	0.98	0.18	46,46,46,46	0

6.5 Other polymers (i)

There are no such residues in this entry.