



wwPDB X-ray Structure Validation Summary Report

Nov 21, 2023 – 05:01 PM JST

PDB ID : 7VYR
Title : Crystal structure of SARS-CoV-2 Spike RBD in complex with the D27 neutralizing antibody Fab fragment
Authors : Jeong, B.-S.; Oh, B.-H.
Deposited on : 2021-11-15
Resolution : 2.20 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

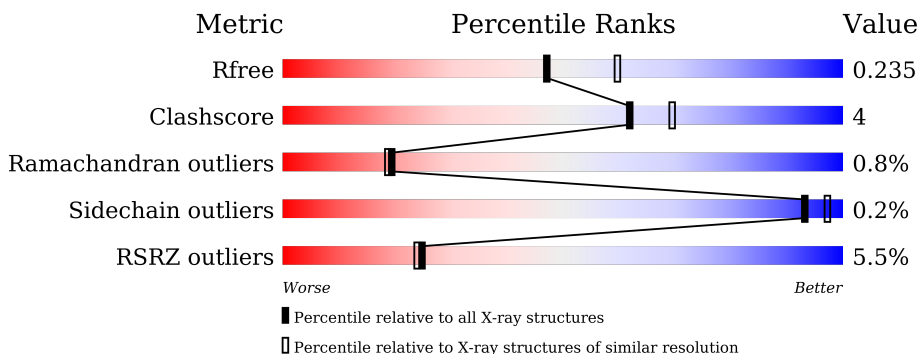
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	252	
1	H	252	
2	B	238	
2	L	238	
3	C	262	
3	R	262	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 9940 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 D27 heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	H	219	Total 1612	C 1003	N 272	O 330	S 7	0	0	0
1	A	219	Total 1612	C 1003	N 272	O 330	S 7	0	0	0

- Molecule 2 is a protein called D27 light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	L	214	Total 1644	C 1024	N 279	O 336	S 5	0	0	0
2	B	214	Total 1644	C 1024	N 279	O 336	S 5	0	0	0

- Molecule 3 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	R	186	Total 1484	C 953	N 247	O 277	S 7	0	0	0
3	C	186	Total 1484	C 953	N 247	O 277	S 7	0	0	0

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
R	280	MET	-	initiating methionine	UNP P0DTC2
R	281	VAL	-	expression tag	UNP P0DTC2
R	282	LEU	-	expression tag	UNP P0DTC2
R	283	VAL	-	expression tag	UNP P0DTC2
R	284	ASN	-	expression tag	UNP P0DTC2
R	285	GLN	-	expression tag	UNP P0DTC2
R	286	SER	-	expression tag	UNP P0DTC2
R	287	HIS	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
R	288	GLN	-	expression tag	UNP P0DTC2
R	289	GLY	-	expression tag	UNP P0DTC2
R	290	PHE	-	expression tag	UNP P0DTC2
R	291	ASN	-	expression tag	UNP P0DTC2
R	292	LYS	-	expression tag	UNP P0DTC2
R	293	GLU	-	expression tag	UNP P0DTC2
R	294	HIS	-	expression tag	UNP P0DTC2
R	295	THR	-	expression tag	UNP P0DTC2
R	296	SER	-	expression tag	UNP P0DTC2
R	297	LYS	-	expression tag	UNP P0DTC2
R	298	MET	-	expression tag	UNP P0DTC2
R	299	VAL	-	expression tag	UNP P0DTC2
R	300	SER	-	expression tag	UNP P0DTC2
R	301	ALA	-	expression tag	UNP P0DTC2
R	302	ILE	-	expression tag	UNP P0DTC2
R	303	VAL	-	expression tag	UNP P0DTC2
R	304	LEU	-	expression tag	UNP P0DTC2
R	305	TYR	-	expression tag	UNP P0DTC2
R	306	VAL	-	expression tag	UNP P0DTC2
R	307	LEU	-	expression tag	UNP P0DTC2
R	308	LEU	-	expression tag	UNP P0DTC2
R	309	ALA	-	expression tag	UNP P0DTC2
R	310	ALA	-	expression tag	UNP P0DTC2
R	311	ALA	-	expression tag	UNP P0DTC2
R	312	ALA	-	expression tag	UNP P0DTC2
R	313	HIS	-	expression tag	UNP P0DTC2
R	314	SER	-	expression tag	UNP P0DTC2
R	315	ALA	-	expression tag	UNP P0DTC2
R	316	PHE	-	expression tag	UNP P0DTC2
R	317	ALA	-	expression tag	UNP P0DTC2
R	318	ALA	-	expression tag	UNP P0DTC2
C	280	MET	-	initiating methionine	UNP P0DTC2
C	281	VAL	-	expression tag	UNP P0DTC2
C	282	LEU	-	expression tag	UNP P0DTC2
C	283	VAL	-	expression tag	UNP P0DTC2
C	284	ASN	-	expression tag	UNP P0DTC2
C	285	GLN	-	expression tag	UNP P0DTC2
C	286	SER	-	expression tag	UNP P0DTC2
C	287	HIS	-	expression tag	UNP P0DTC2
C	288	GLN	-	expression tag	UNP P0DTC2
C	289	GLY	-	expression tag	UNP P0DTC2
C	290	PHE	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	291	ASN	-	expression tag	UNP P0DTC2
C	292	LYS	-	expression tag	UNP P0DTC2
C	293	GLU	-	expression tag	UNP P0DTC2
C	294	HIS	-	expression tag	UNP P0DTC2
C	295	THR	-	expression tag	UNP P0DTC2
C	296	SER	-	expression tag	UNP P0DTC2
C	297	LYS	-	expression tag	UNP P0DTC2
C	298	MET	-	expression tag	UNP P0DTC2
C	299	VAL	-	expression tag	UNP P0DTC2
C	300	SER	-	expression tag	UNP P0DTC2
C	301	ALA	-	expression tag	UNP P0DTC2
C	302	ILE	-	expression tag	UNP P0DTC2
C	303	VAL	-	expression tag	UNP P0DTC2
C	304	LEU	-	expression tag	UNP P0DTC2
C	305	TYR	-	expression tag	UNP P0DTC2
C	306	VAL	-	expression tag	UNP P0DTC2
C	307	LEU	-	expression tag	UNP P0DTC2
C	308	LEU	-	expression tag	UNP P0DTC2
C	309	ALA	-	expression tag	UNP P0DTC2
C	310	ALA	-	expression tag	UNP P0DTC2
C	311	ALA	-	expression tag	UNP P0DTC2
C	312	ALA	-	expression tag	UNP P0DTC2
C	313	HIS	-	expression tag	UNP P0DTC2
C	314	SER	-	expression tag	UNP P0DTC2
C	315	ALA	-	expression tag	UNP P0DTC2
C	316	PHE	-	expression tag	UNP P0DTC2
C	317	ALA	-	expression tag	UNP P0DTC2
C	318	ALA	-	expression tag	UNP P0DTC2

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	H	50	Total O 50 50	0	0
4	L	107	Total O 107 107	0	0
4	R	50	Total O 50 50	0	0
4	A	112	Total O 112 112	0	0
4	B	95	Total O 95 95	0	0

Continued on next page...

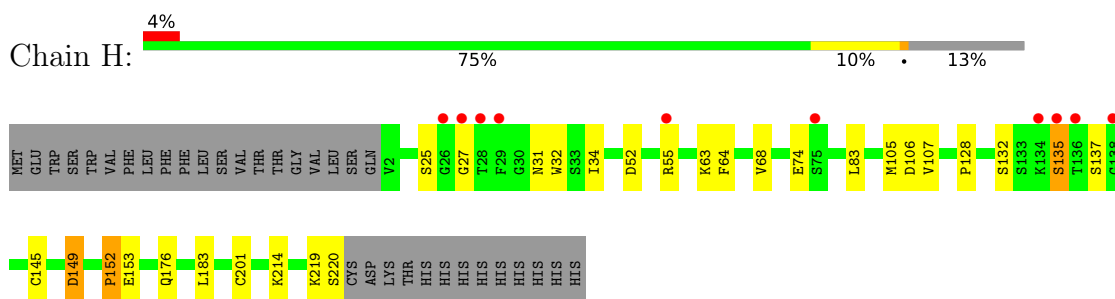
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	C	46	Total	O	0	0
			46	46		

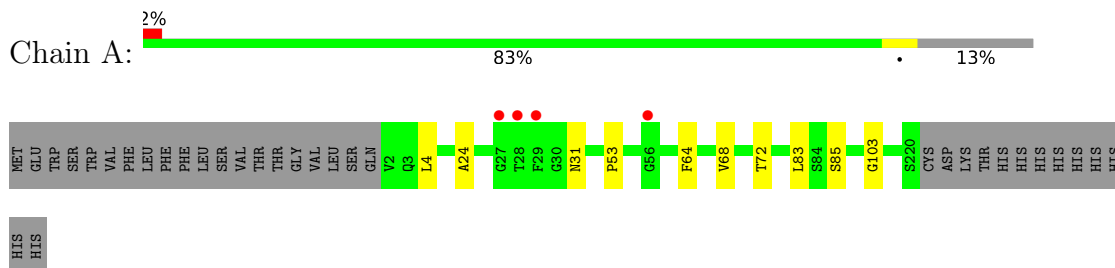
3 Residue-property plots [i](#)

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.

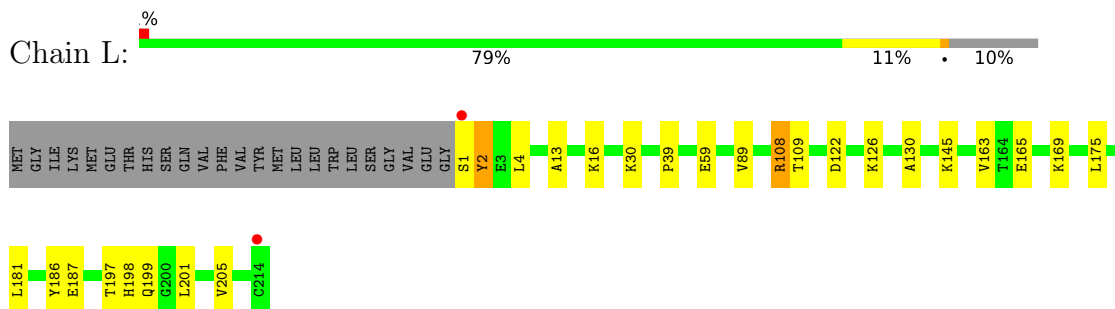
- Molecule 1: D27 heavy chain



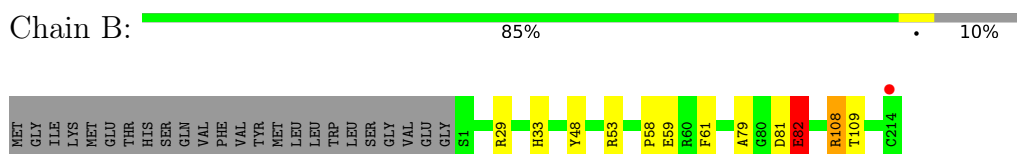
- Molecule 1: D27 heavy chain



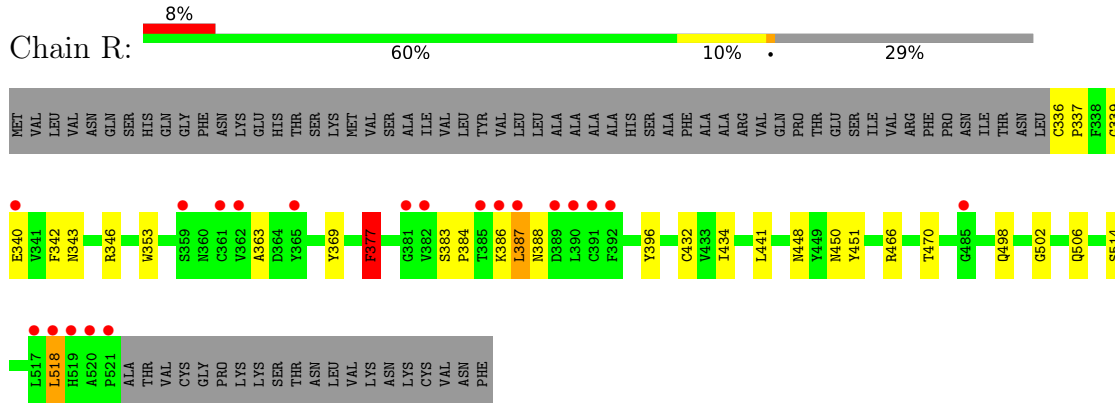
- Molecule 2: D27 light chain



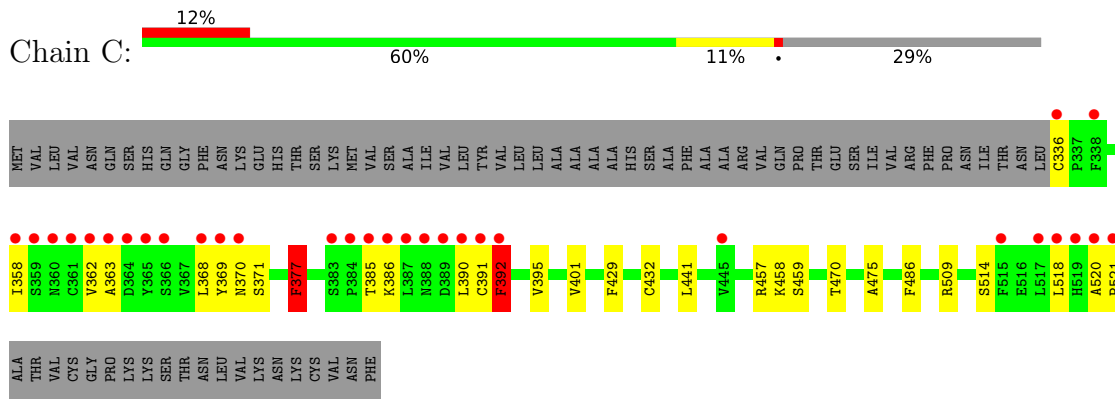
- Molecule 2: D27 light chain



• Molecule 3: Spike protein S1



• Molecule 3: Spike protein S1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.80Å 191.68Å 65.90Å 90.00° 93.99° 90.00°	Depositor
Resolution (Å)	47.92 – 2.20 47.92 – 1.89	Depositor EDS
% Data completeness (in resolution range)	95.4 (47.92-2.20) 90.8 (47.92-1.89)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 1.90Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.192 , 0.235 0.192 , 0.235	Depositor DCC
R_{free} test set	2000 reflections (1.77%)	wwPDB-VP
Wilson B-factor (Å ²)	19.8	Xtrriage
Anisotropy	0.312	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 39.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	9940	wwPDB-VP
Average B, all atoms (Å ²)	29.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.43% 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	0.44	0/1647	0.62	0/2241
1	H	0.51	0/1647	0.73	1/2241 (0.0%)
2	B	0.52	2/1680 (0.1%)	0.66	1/2284 (0.0%)
2	L	0.53	0/1680	0.75	4/2284 (0.2%)
3	C	0.41	0/1527	0.70	2/2077 (0.1%)
3	R	0.41	0/1527	0.71	3/2077 (0.1%)
All	All	0.47	2/9708 (0.0%)	0.70	11/13204 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	H	0	4
2	B	0	3
3	C	0	2
3	R	0	2
All	All	0	12

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	82	GLU	CD-OE1	7.33	1.33	1.25
2	B	82	GLU	CD-OE2	5.26	1.31	1.25

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	169	LYS	CD-CE-NZ	-9.14	90.69	111.70
2	L	108	ARG	CG-CD-NE	-8.39	94.18	111.80
3	R	387	LEU	CA-CB-CG	7.35	132.20	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	L	187	GLU	CA-CB-CG	7.34	129.55	113.40
2	B	108	ARG	CG-CD-NE	-6.88	97.35	111.80

There are no chirality outliers.

5 of 12 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	H	105	MET	Peptide
1	H	135	SER	Peptide
1	H	137	SER	Peptide
1	H	152	PRO	Peptide
3	R	377	PHE	Sidechain

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	1612	0	1571	8	0
1	H	1612	0	1573	15	0
2	B	1644	0	1580	7	0
2	L	1644	0	1580	15	0
3	C	1484	0	1402	24	0
3	R	1484	0	1402	20	0
4	A	112	0	0	0	0
4	B	95	0	0	1	0
4	C	46	0	0	1	0
4	H	50	0	0	0	0
4	L	107	0	0	0	0
4	R	50	0	0	0	0
All	All	9940	0	9108	82	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 4.

The worst 5 of 82 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)
3:R:340:GLU:HA	3:R:343:ASN:HB2	1.56	0.87
3:R:346:ARG:NH1	3:R:451:TYR:OH	2.13	0.80
3:R:369:TYR:HA	3:R:377:PHE:CE2	2.26	0.71
2:L:198:HIS:ND1	2:L:199:GLN:O	2.26	0.68
3:C:386:LYS:O	3:C:390:LEU:HD22	1.94	0.67

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	217/252 (86%)	214 (99%)	3 (1%)	0	100	100
1	H	217/252 (86%)	209 (96%)	7 (3%)	1 (0%)	29	31
2	B	212/238 (89%)	204 (96%)	7 (3%)	1 (0%)	29	31
2	L	212/238 (89%)	204 (96%)	6 (3%)	2 (1%)	17	16
3	C	184/262 (70%)	170 (92%)	12 (6%)	2 (1%)	14	12
3	R	184/262 (70%)	161 (88%)	19 (10%)	4 (2%)	6	4
All	All	1226/1504 (82%)	1162 (95%)	54 (4%)	10 (1%)	19	19

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	R	337	PRO
3	R	518	LEU
2	B	59	GLU
3	R	339	GLY
3	C	385	THR

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	182/214 (85%)	182 (100%)	0	100	100
1	H	182/214 (85%)	182 (100%)	0	100	100
2	B	186/207 (90%)	185 (100%)	1 (0%)	88	94
2	L	186/207 (90%)	186 (100%)	0	100	100
3	C	161/226 (71%)	160 (99%)	1 (1%)	86	93
3	R	161/226 (71%)	161 (100%)	0	100	100
All	All	1058/1294 (82%)	1056 (100%)	2 (0%)	93	97

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	82	GLU
3	C	392	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
3	R	498	GLN
1	A	3	GLN
1	A	43	GLN
2	B	33	HIS
2	B	199	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](#)

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	219/252 (86%)	-0.51	4 (1%) 68 66	12, 20, 50, 87	0
1	H	219/252 (86%)	-0.35	10 (4%) 32 31	12, 22, 55, 96	0
2	B	214/238 (89%)	-0.58	1 (0%) 91 90	12, 24, 42, 55	0
2	L	214/238 (89%)	-0.65	2 (0%) 84 83	11, 18, 31, 61	0
3	C	186/262 (70%)	0.43	31 (16%) 1 1	18, 32, 92, 123	0
3	R	186/262 (70%)	0.37	20 (10%) 5 5	17, 36, 83, 119	0
All	All	1238/1504 (82%)	-0.24	68 (5%) 25 24	11, 24, 70, 123	0

The worst 5 of 68 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	385	THR	15.0
1	H	138	GLY	10.1
3	C	368	LEU	9.5
1	A	27	GLY	9.2
3	C	518	LEU	7.9

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.