



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 6, 2026 – 03:27 PM JST

PDB ID : 9LFO / pdb_00009lfo
Title : Complex structure of THC-2199 Fab bound to SARS-CoV-2 XBB.1.5 RBD
Authors : Wang, X.; Guo, F.
Deposited on : 2025-01-08
Resolution : 2.91 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	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

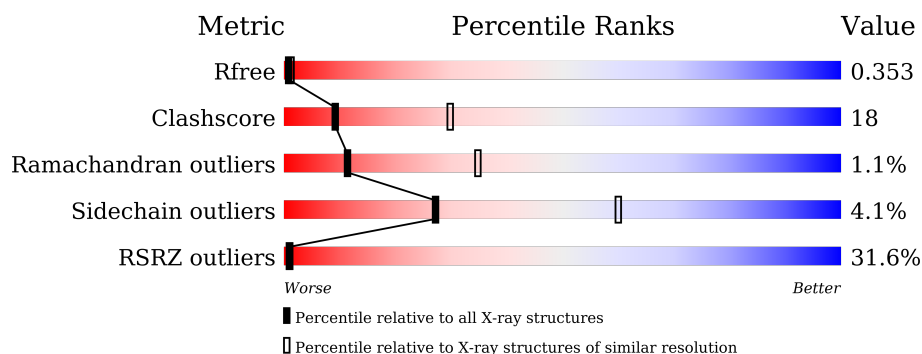
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.91 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}	180053	2995 (2.94-2.90)
Clashscore	190562	3213 (2.94-2.90)
Ramachandran outliers	187476	3128 (2.94-2.90)
Sidechain outliers	187428	3130 (2.94-2.90)
RSRZ outliers	180081	2995 (2.94-2.90)

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	C	219	26% 71% 15% • 12%
1	F	219	22% 58% 30% • 12%
2	A	232	27% 63% 34% •
2	D	232	35% 71% 25% • •
3	B	219	27% 71% 26% •
3	E	219	45% 63% 32% 5%

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 9932 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	192	Total	C	N	O	S	0	0	0
			1527	987	255	277	8			
1	F	193	Total	C	N	O	S	0	0	0
			1535	991	257	279	8			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	339	HIS	GLY	variant	UNP P0DTC2
C	346	THR	ARG	variant	UNP P0DTC2
C	368	ILE	LEU	variant	UNP P0DTC2
C	371	PHE	SER	variant	UNP P0DTC2
C	373	PRO	SER	variant	UNP P0DTC2
C	375	PHE	SER	variant	UNP P0DTC2
C	376	ALA	THR	variant	UNP P0DTC2
C	405	ASN	ASP	variant	UNP P0DTC2
C	408	SER	ARG	variant	UNP P0DTC2
C	417	ASN	LYS	variant	UNP P0DTC2
C	440	LYS	ASN	variant	UNP P0DTC2
C	445	PRO	VAL	variant	UNP P0DTC2
C	446	SER	GLY	variant	UNP P0DTC2
C	460	LYS	ASN	variant	UNP P0DTC2
C	477	ASN	SER	variant	UNP P0DTC2
C	478	LYS	THR	variant	UNP P0DTC2
C	484	ALA	GLU	variant	UNP P0DTC2
C	486	PRO	PHE	variant	UNP P0DTC2
C	490	SER	PHE	variant	UNP P0DTC2
C	498	ARG	GLN	variant	UNP P0DTC2
C	501	TYR	ASN	variant	UNP P0DTC2
C	505	HIS	TYR	variant	UNP P0DTC2
F	339	HIS	GLY	variant	UNP P0DTC2
F	346	THR	ARG	variant	UNP P0DTC2
F	368	ILE	LEU	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	371	PHE	SER	variant	UNP P0DTC2
F	373	PRO	SER	variant	UNP P0DTC2
F	375	PHE	SER	variant	UNP P0DTC2
F	376	ALA	THR	variant	UNP P0DTC2
F	405	ASN	ASP	variant	UNP P0DTC2
F	408	SER	ARG	variant	UNP P0DTC2
F	417	ASN	LYS	variant	UNP P0DTC2
F	440	LYS	ASN	variant	UNP P0DTC2
F	445	PRO	VAL	variant	UNP P0DTC2
F	446	SER	GLY	variant	UNP P0DTC2
F	460	LYS	ASN	variant	UNP P0DTC2
F	477	ASN	SER	variant	UNP P0DTC2
F	478	LYS	THR	variant	UNP P0DTC2
F	484	ALA	GLU	variant	UNP P0DTC2
F	486	PRO	PHE	variant	UNP P0DTC2
F	490	SER	PHE	variant	UNP P0DTC2
F	498	ARG	GLN	variant	UNP P0DTC2
F	501	TYR	ASN	variant	UNP P0DTC2
F	505	HIS	TYR	variant	UNP P0DTC2

- Molecule 2 is a protein called THC-2199 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	232	Total	C	N	O	S	0	0	0
			1760	1112	293	347	8			
2	D	228	Total	C	N	O	S	0	0	0
			1730	1096	288	339	7			

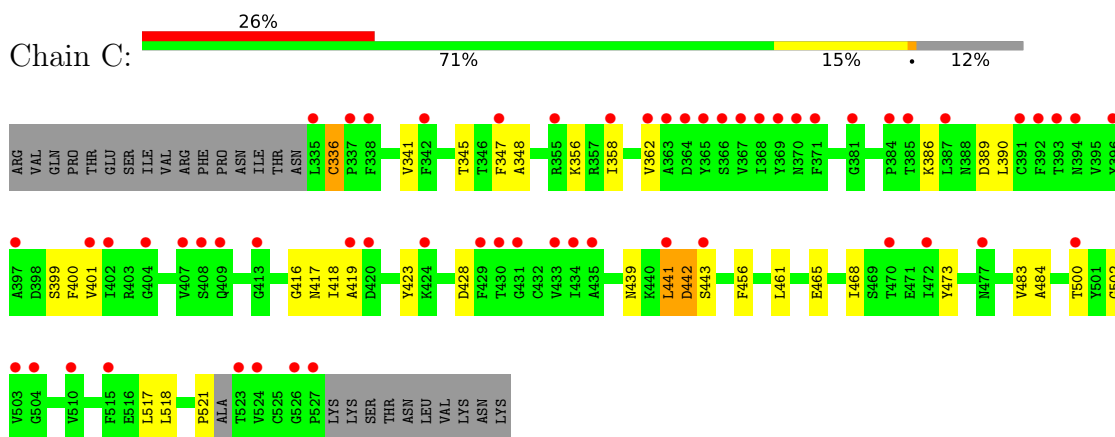
- Molecule 3 is a protein called THC-2199 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	219	Total	C	N	O	S	0	0	0
			1690	1054	290	339	7			
3	E	219	Total	C	N	O	S	0	0	0
			1690	1054	290	339	7			

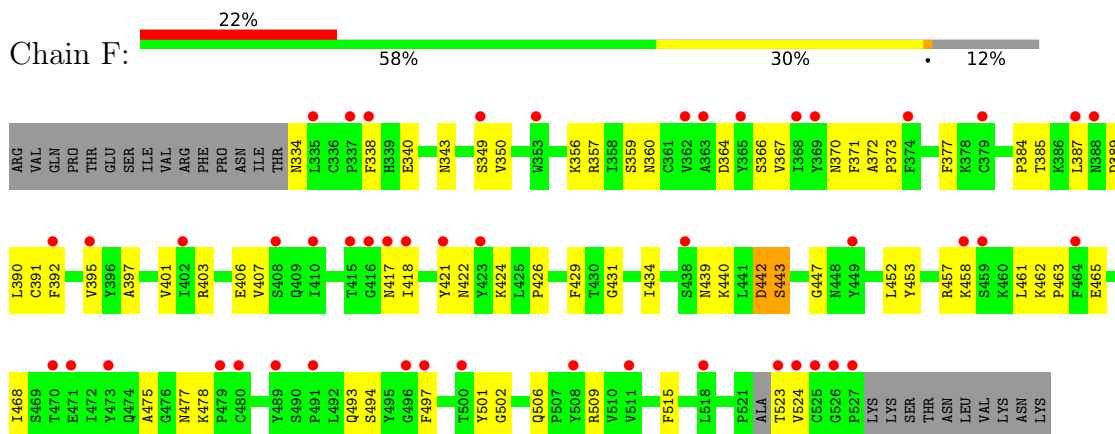
3 Residue-property plots

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.

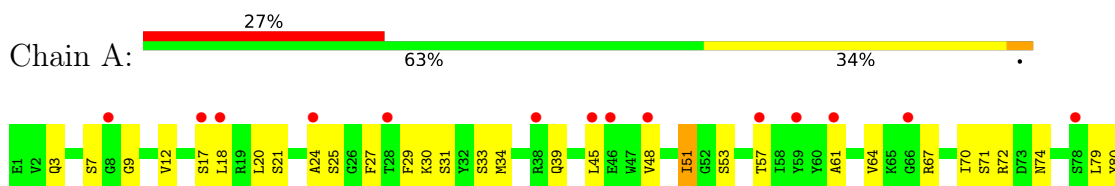
• Molecule 1: Spike protein S1

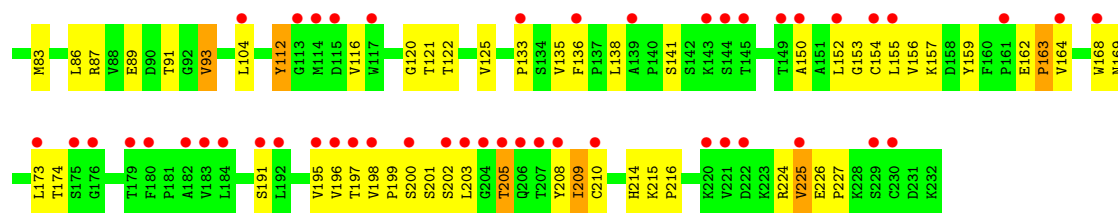


• Molecule 1: Spike protein S1

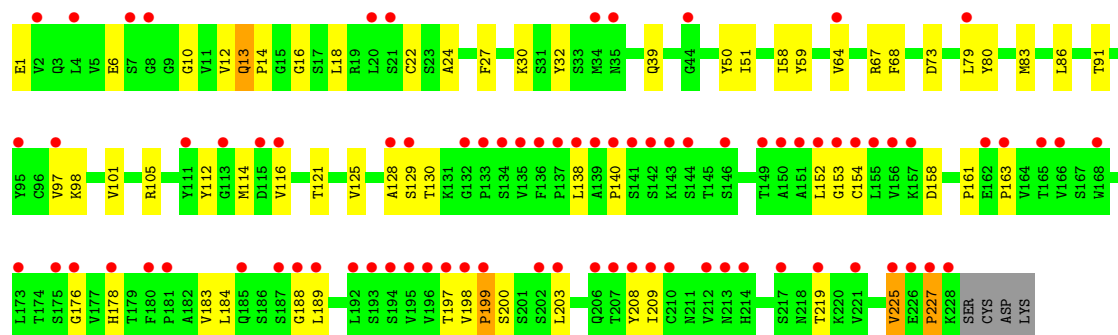


• Molecule 2: THC-2199 Heavy chain

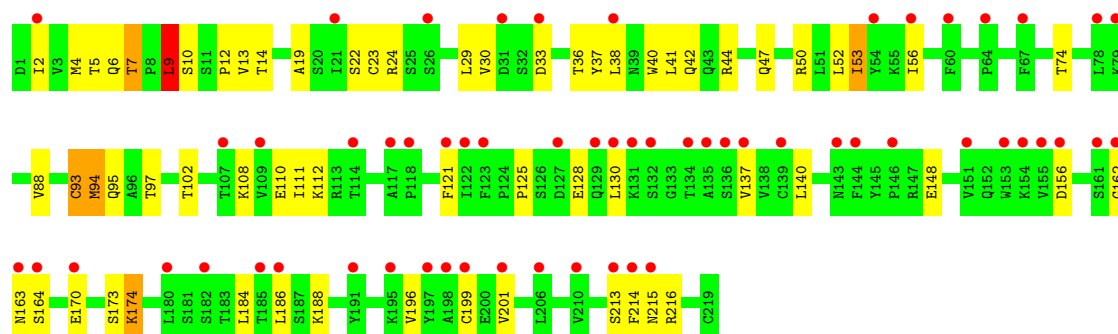




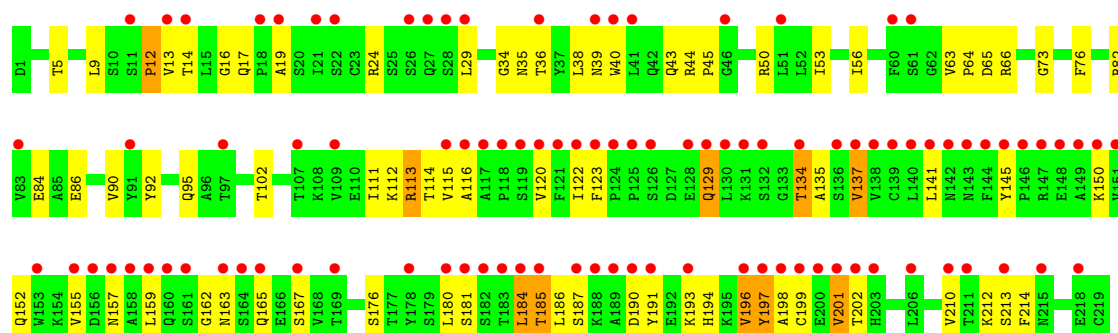
• Molecule 2: THC-2199 Heavy chain



• Molecule 3: THC-2199 Light chain



• Molecule 3: THC-2199 Light chain



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	97.79Å 120.30Å 317.11Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.89 – 2.91 48.89 – 2.91	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.89-2.91) 99.8 (48.89-2.91)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.36 (at 2.91Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.300 , 0.351 0.351 , 0.353	Depositor DCC
R_{free} test set	2000 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	62.3	Xtriage
Anisotropy	0.479	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 78.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.35$, $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.74	EDS
Total number of atoms	9932	wwPDB-VP
Average B, all atoms (Å ²)	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% 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

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	C	0.91	0/1575	1.23	3/2145 (0.1%)
1	F	0.93	0/1583	1.24	5/2156 (0.2%)
2	A	0.95	0/1803	1.24	2/2450 (0.1%)
2	D	0.97	1/1773 (0.1%)	1.28	5/2412 (0.2%)
3	B	0.95	1/1725 (0.1%)	1.27	3/2340 (0.1%)
3	E	0.97	0/1725	1.28	5/2340 (0.2%)
All	All	0.95	2/10184 (0.0%)	1.26	23/13843 (0.2%)

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
3	B	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	199	PRO	N-CD	-7.90	1.36	1.47
3	B	9	LEU	CA-C	5.49	1.60	1.52

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	439	ASN	N-CA-C	-11.50	98.68	113.17
1	F	338	PHE	N-CA-C	-9.85	100.39	111.82
1	F	340	GLU	N-CA-C	8.56	121.69	111.33
3	B	164	SER	N-CA-C	8.19	122.60	108.76
3	E	137	VAL	N-CA-C	8.11	119.84	107.99

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	7	THR	Peptide
3	B	9	LEU	Mainchain

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	C	1527	0	1454	58	4
1	F	1535	0	1460	56	24
2	A	1760	0	1719	75	59
2	D	1730	0	1692	58	0
3	B	1690	0	1653	52	0
3	E	1690	0	1653	80	83
All	All	9932	0	9631	356	87

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 18.

The worst 5 of 356 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:C:347:PHE:HZ	1:C:400:PHE:N	1.09	1.50
3:E:191:TYR:HA	3:E:194:HIS:CE1	1.51	1.46
1:C:347:PHE:CD2	1:C:401:VAL:HG23	1.53	1.42
3:E:44:ARG:NH1	3:E:45:PRO:HG2	1.36	1.36
3:E:129:GLN:NE2	3:E:134:THR:HG22	1.38	1.34

The worst 5 of 87 symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:370:ASN:ND2	3:E:84:GLU:CD[5_545]	0.40	1.80
2:A:205:THR:O	3:E:194:HIS:CD2[4_555]	0.50	1.70
2:A:226:GLU:CB	3:E:159:LEU:CD1[4_555]	0.60	1.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:199:PRO:C	3:E:186:LEU:CG[4_555]	0.66	1.54
2:A:226:GLU:CA	3:E:159:LEU:CB[4_555]	0.69	1.51

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	C	188/219 (86%)	171 (91%)	16 (8%)	1 (0%)	24	53
1	F	189/219 (86%)	172 (91%)	16 (8%)	1 (0%)	24	53
2	A	230/232 (99%)	216 (94%)	12 (5%)	2 (1%)	14	40
2	D	226/232 (97%)	212 (94%)	11 (5%)	3 (1%)	9	30
3	B	217/219 (99%)	206 (95%)	9 (4%)	2 (1%)	14	40
3	E	217/219 (99%)	199 (92%)	13 (6%)	5 (2%)	5	18
All	All	1267/1340 (95%)	1176 (93%)	77 (6%)	14 (1%)	11	35

5 of 14 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	442	ASP
3	B	213	SER
1	F	442	ASP
3	E	129	GLN
3	B	214	PHE

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	C	166/192 (86%)	163 (98%)	3 (2%)	51 79
1	F	167/192 (87%)	162 (97%)	5 (3%)	36 68
2	A	200/200 (100%)	188 (94%)	12 (6%)	17 45
2	D	196/200 (98%)	190 (97%)	6 (3%)	35 67
3	B	194/194 (100%)	184 (95%)	10 (5%)	21 50
3	E	194/194 (100%)	184 (95%)	10 (5%)	21 50
All	All	1117/1172 (95%)	1071 (96%)	46 (4%)	27 60

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	468	ILE
2	D	219	THR
1	F	501	TYR
2	D	101	VAL
3	E	122	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
3	E	39	ASN
3	E	157	ASN
3	E	194	HIS
3	E	163	ASN
3	B	171	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 [i](#)

There are no oligosaccharides 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	C	192/219 (87%)	1.56	57 (29%) 1 1	26, 58, 82, 92	0
1	F	193/219 (88%)	1.53	48 (24%) 2 1	30, 54, 82, 99	0
2	A	232/232 (100%)	1.55	62 (26%) 1 1	25, 50, 101, 106	0
2	D	228/232 (98%)	1.98	82 (35%) 1 1	28, 57, 111, 118	0
3	B	219/219 (100%)	1.51	59 (26%) 1 1	27, 59, 90, 100	0
3	E	219/219 (100%)	1.93	98 (44%) 0 0	27, 61, 104, 109	0
All	All	1283/1340 (95%)	1.68	406 (31%) 1 1	25, 57, 101, 118	0

The worst 5 of 406 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	195	VAL	8.5
2	D	208	TYR	8.5
3	E	124	PRO	6.8
2	D	137	PRO	6.6
2	D	226	GLU	6.3

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 oligosaccharides 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.