wwPDB X-ray Structure Validation Summary Report (i)

Nov 22, 2023 – 10:44 PM JST

PDB ID	:	7XSC
Title	:	Crystal structure of SARS-CoV-2 spike receptor binding domain bound with P5S-2B10
Authors	:	Wang, X.; Wang, Z.; Lin, Z.
Deposited on	:	2022-05-13
Resolution	:	2.88 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 (i) 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 (1)) were used in the production of this report:

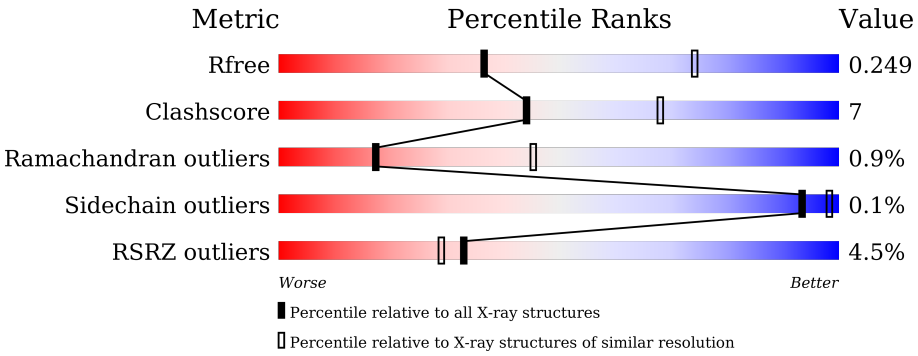
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.88 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	2691 (2.90-2.86)
Clashscore	141614	2947 (2.90-2.86)
Ramachandran outliers	138981	2868 (2.90-2.86)
Sidechain outliers	138945	2871 (2.90-2.86)
RSRZ outliers	127900	2629 (2.90-2.86)

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	220	83% 10% 6%
1	D	220	2% 75% 19% 6%
2	E	258	8% 64% 9% 25%
2	F	258	10% 62% 12% 25%
3	B	214	87% 10%
3	H	214	% 85% 14%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 9423 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 P5S-2B10 Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	206	Total	C	N	O	S	0	0	0
			1537	972	259	300	6			
1	D	207	Total	C	N	O	S	0	0	0
			1543	975	260	302	6			

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	193	Total	C	N	O	S	0	0	0
			1529	980	255	286	8			
2	F	193	Total	C	N	O	S	0	0	0
			1529	980	255	286	8			

There are 94 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	278	MET	-	initiating methionine	UNP P0DTC2
E	279	LEU	-	expression tag	UNP P0DTC2
E	280	LEU	-	expression tag	UNP P0DTC2
E	281	VAL	-	expression tag	UNP P0DTC2
E	282	ASN	-	expression tag	UNP P0DTC2
E	283	GLN	-	expression tag	UNP P0DTC2
E	284	SER	-	expression tag	UNP P0DTC2
E	285	HIS	-	expression tag	UNP P0DTC2
E	286	GLN	-	expression tag	UNP P0DTC2
E	287	GLY	-	expression tag	UNP P0DTC2
E	288	PHE	-	expression tag	UNP P0DTC2
E	289	ASN	-	expression tag	UNP P0DTC2
E	290	LYS	-	expression tag	UNP P0DTC2
E	291	GLU	-	expression tag	UNP P0DTC2
E	292	HIS	-	expression tag	UNP P0DTC2
E	293	THR	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	294	SER	-	expression tag	UNP P0DTC2
E	295	LYS	-	expression tag	UNP P0DTC2
E	296	MET	-	expression tag	UNP P0DTC2
E	297	VAL	-	expression tag	UNP P0DTC2
E	298	SER	-	expression tag	UNP P0DTC2
E	299	ALA	-	expression tag	UNP P0DTC2
E	300	ILE	-	expression tag	UNP P0DTC2
E	301	VAL	-	expression tag	UNP P0DTC2
E	302	LEU	-	expression tag	UNP P0DTC2
E	303	TYR	-	expression tag	UNP P0DTC2
E	304	VAL	-	expression tag	UNP P0DTC2
E	305	LEU	-	expression tag	UNP P0DTC2
E	306	LEU	-	expression tag	UNP P0DTC2
E	307	ALA	-	expression tag	UNP P0DTC2
E	308	ALA	-	expression tag	UNP P0DTC2
E	309	ALA	-	expression tag	UNP P0DTC2
E	310	ALA	-	expression tag	UNP P0DTC2
E	311	HIS	-	expression tag	UNP P0DTC2
E	312	SER	-	expression tag	UNP P0DTC2
E	313	ALA	-	expression tag	UNP P0DTC2
E	314	PHE	-	expression tag	UNP P0DTC2
E	315	ALA	-	expression tag	UNP P0DTC2
E	316	ALA	-	expression tag	UNP P0DTC2
E	317	ASP	-	expression tag	UNP P0DTC2
E	318	PRO	-	expression tag	UNP P0DTC2
E	530	HIS	-	expression tag	UNP P0DTC2
E	531	HIS	-	expression tag	UNP P0DTC2
E	532	HIS	-	expression tag	UNP P0DTC2
E	533	HIS	-	expression tag	UNP P0DTC2
E	534	HIS	-	expression tag	UNP P0DTC2
E	535	HIS	-	expression tag	UNP P0DTC2
F	278	MET	-	initiating methionine	UNP P0DTC2
F	279	LEU	-	expression tag	UNP P0DTC2
F	280	LEU	-	expression tag	UNP P0DTC2
F	281	VAL	-	expression tag	UNP P0DTC2
F	282	ASN	-	expression tag	UNP P0DTC2
F	283	GLN	-	expression tag	UNP P0DTC2
F	284	SER	-	expression tag	UNP P0DTC2
F	285	HIS	-	expression tag	UNP P0DTC2
F	286	GLN	-	expression tag	UNP P0DTC2
F	287	GLY	-	expression tag	UNP P0DTC2
F	288	PHE	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	289	ASN	-	expression tag	UNP P0DTC2
F	290	LYS	-	expression tag	UNP P0DTC2
F	291	GLU	-	expression tag	UNP P0DTC2
F	292	HIS	-	expression tag	UNP P0DTC2
F	293	THR	-	expression tag	UNP P0DTC2
F	294	SER	-	expression tag	UNP P0DTC2
F	295	LYS	-	expression tag	UNP P0DTC2
F	296	MET	-	expression tag	UNP P0DTC2
F	297	VAL	-	expression tag	UNP P0DTC2
F	298	SER	-	expression tag	UNP P0DTC2
F	299	ALA	-	expression tag	UNP P0DTC2
F	300	ILE	-	expression tag	UNP P0DTC2
F	301	VAL	-	expression tag	UNP P0DTC2
F	302	LEU	-	expression tag	UNP P0DTC2
F	303	TYR	-	expression tag	UNP P0DTC2
F	304	VAL	-	expression tag	UNP P0DTC2
F	305	LEU	-	expression tag	UNP P0DTC2
F	306	LEU	-	expression tag	UNP P0DTC2
F	307	ALA	-	expression tag	UNP P0DTC2
F	308	ALA	-	expression tag	UNP P0DTC2
F	309	ALA	-	expression tag	UNP P0DTC2
F	310	ALA	-	expression tag	UNP P0DTC2
F	311	HIS	-	expression tag	UNP P0DTC2
F	312	SER	-	expression tag	UNP P0DTC2
F	313	ALA	-	expression tag	UNP P0DTC2
F	314	PHE	-	expression tag	UNP P0DTC2
F	315	ALA	-	expression tag	UNP P0DTC2
F	316	ALA	-	expression tag	UNP P0DTC2
F	317	ASP	-	expression tag	UNP P0DTC2
F	318	PRO	-	expression tag	UNP P0DTC2
F	530	HIS	-	expression tag	UNP P0DTC2
F	531	HIS	-	expression tag	UNP P0DTC2
F	532	HIS	-	expression tag	UNP P0DTC2
F	533	HIS	-	expression tag	UNP P0DTC2
F	534	HIS	-	expression tag	UNP P0DTC2
F	535	HIS	-	expression tag	UNP P0DTC2

- Molecule 3 is a protein called P5S-2B10 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	213	Total	C	N	O	S	0	0	0
			1640	1026	276	333	5			

Continued on next page...

Continued from previous page...

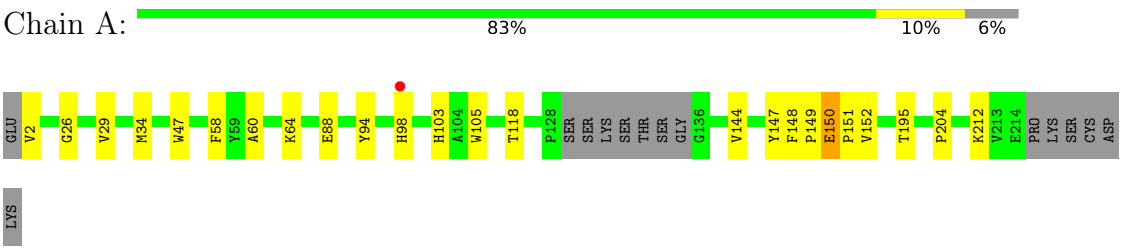
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	214	Total	C	N	O	S	0	0	0
			1645	1029	277	334	5			

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.

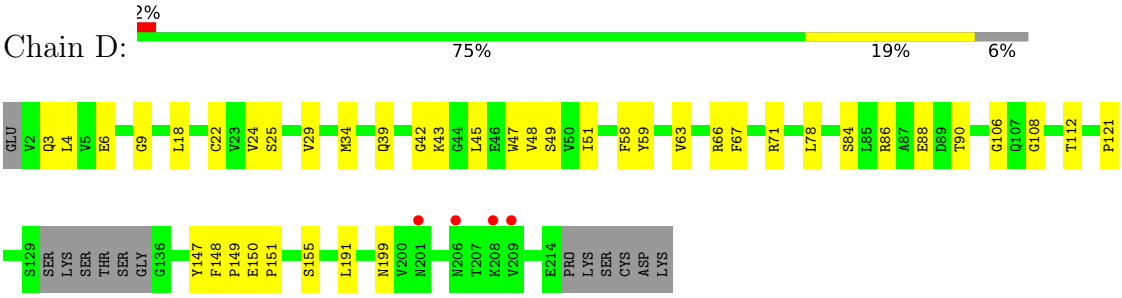
• Molecule 1: P5S-2B10 Heavy chain

Chain A:



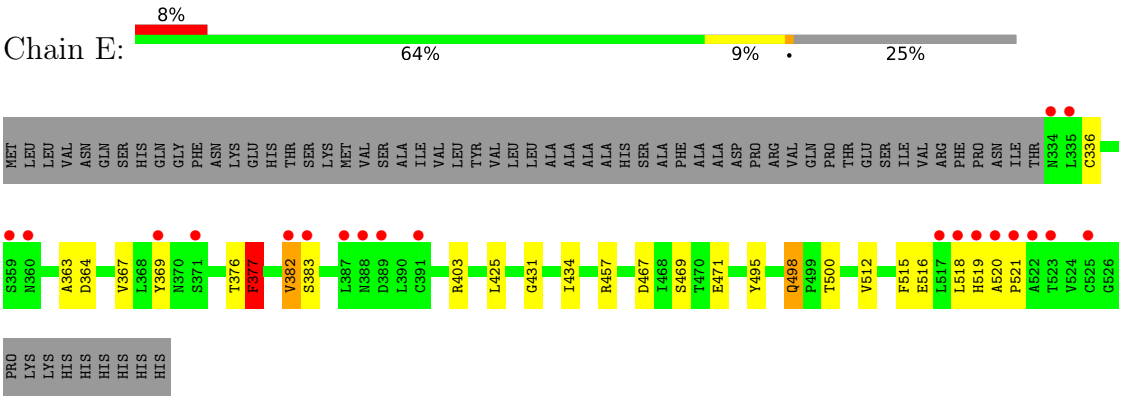
• Molecule 1: P5S-2B10 Heavy chain

Chain D:



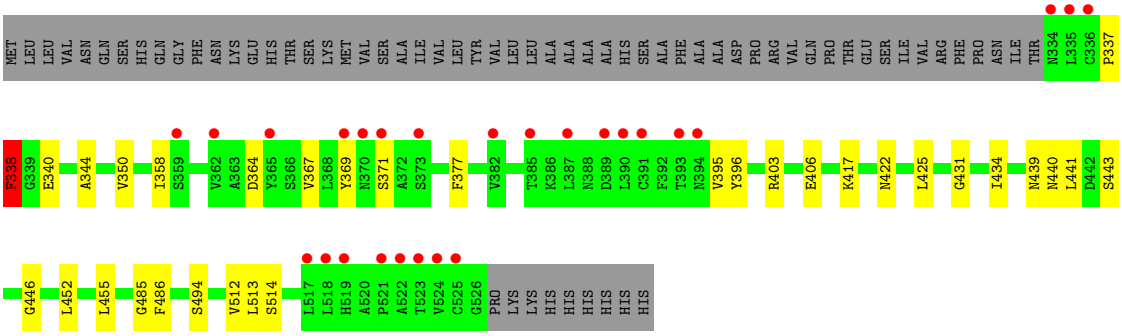
• Molecule 2: Spike protein S1

Chain E:

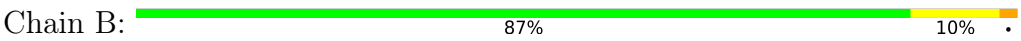


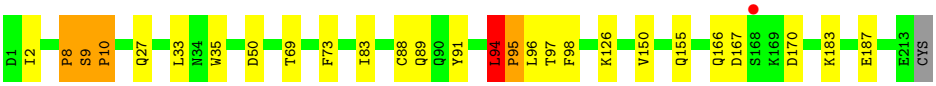
• Molecule 2: Spike protein S1

Chain F:

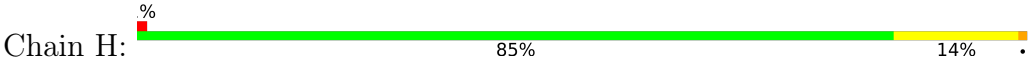


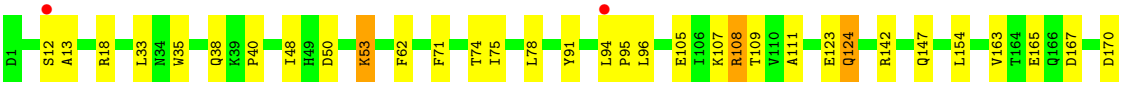
• Molecule 3: P5S-2B10 Light chain





• Molecule 3: P5S-2B10 Light chain





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	87.42Å 109.19Å 174.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.31 – 2.88 46.31 – 2.88	Depositor EDS
% Data completeness (in resolution range)	98.8 (46.31-2.88) 98.8 (46.31-2.88)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.08 (at 2.86Å)	Xtriage
Refinement program	PHENIX (1.18.2_3874: ???)	Depositor
R, R_{free}	0.200 , 0.250 0.200 , 0.249	Depositor DCC
R_{free} test set	1827 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	36.1	Xtriage
Anisotropy	0.422	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 35.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	9423	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.04% 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	0.29	0/1574	0.55	0/2146
1	D	0.28	0/1580	0.56	0/2154
2	E	0.33	0/1572	0.62	3/2139 (0.1%)
2	F	0.40	1/1572 (0.1%)	0.63	3/2139 (0.1%)
3	B	0.29	0/1675	0.55	0/2274
3	H	0.35	1/1680 (0.1%)	0.61	1/2281 (0.0%)
All	All	0.33	2/9653 (0.0%)	0.59	7/13133 (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
2	E	0	1
2	F	0	1
3	B	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	53	LYS	CE-NZ	-5.71	1.34	1.49
2	F	338	PHE	CE1-CZ	5.61	1.48	1.37

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	H	108	ARG	NE-CZ-NH1	10.79	125.69	120.30
2	F	338	PHE	CB-CG-CD2	-10.23	113.64	120.80
2	E	377	PHE	CB-CG-CD2	-8.12	115.12	120.80
2	F	338	PHE	N-CA-CB	6.01	121.41	110.60
2	E	498	GLN	CA-CB-CG	5.55	125.61	113.40

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	94	LEU	Peptide
2	E	377	PHE	Sidechain
2	F	338	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	1537	0	1502	17	0
1	D	1543	0	1507	29	0
2	E	1529	0	1446	21	0
2	F	1529	0	1446	24	0
3	B	1640	0	1594	20	0
3	H	1645	0	1596	26	0
All	All	9423	0	9091	122	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 7.

The worst 5 of 122 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:E:364:ASP:OD1	2:E:367:VAL:HG23	1.79	0.82
2:E:498:GLN:NE2	2:F:485:GLY:HA3	1.95	0.81
1:D:39:GLN:HE22	3:H:38:GLN:HE22	1.28	0.81
2:E:369:TYR:HA	2:E:377:PHE:CE2	2.18	0.79
2:E:369:TYR:HA	2:E:377:PHE:HE2	1.46	0.78

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	202/220 (92%)	195 (96%)	4 (2%)	3 (2%)	10 32
1	D	203/220 (92%)	197 (97%)	6 (3%)	0	100 100
2	E	191/258 (74%)	180 (94%)	9 (5%)	2 (1%)	15 42
2	F	191/258 (74%)	181 (95%)	10 (5%)	0	100 100
3	B	211/214 (99%)	199 (94%)	7 (3%)	5 (2%)	6 20
3	H	212/214 (99%)	202 (95%)	9 (4%)	1 (0%)	29 59
All	All	1210/1384 (87%)	1154 (95%)	45 (4%)	11 (1%)	17 45

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	E	519	HIS
3	B	8	PRO
3	B	95	PRO
3	H	124	GLN
1	A	148	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	A	173/186 (93%)	173 (100%)	0	100 100
1	D	174/186 (94%)	174 (100%)	0	100 100
2	E	166/222 (75%)	166 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	F	166/222 (75%)	165 (99%)	1 (1%)	86 95
3	B	187/188 (100%)	187 (100%)	0	100 100
3	H	187/188 (100%)	187 (100%)	0	100 100
All	All	1053/1192 (88%)	1052 (100%)	1 (0%)	93 98

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	F	338	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	27	GLN
3	B	152	ASN
3	H	38	GLN
3	H	3	GLN
2	F	481	ASN

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

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	206/220 (93%)	0.01	1 (0%) 91 90	17, 30, 44, 55	0
1	D	207/220 (94%)	0.13	4 (1%) 66 65	18, 37, 47, 56	0
2	E	193/258 (74%)	0.37	20 (10%) 6 4	19, 31, 71, 96	0
2	F	193/258 (74%)	0.38	26 (13%) 3 2	16, 29, 79, 88	0
3	B	213/214 (99%)	-0.12	1 (0%) 91 90	16, 27, 41, 62	0
3	H	214/214 (100%)	0.13	3 (1%) 75 75	22, 34, 51, 78	0
All	All	1226/1384 (88%)	0.14	55 (4%) 33 29	16, 31, 65, 96	0

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	523	THR	5.1
2	E	383	SER	4.7
2	F	387	LEU	4.6
2	E	519	HIS	4.5
2	F	370	ASN	4.3

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 (i)

There are no such residues in this entry.