



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2020 – 05:07 PM BST

PDB ID : 6ZH9
Title : Ternary complex CR3022 H11-H4 and RBD (SARS-CoV-2)
Authors : Naismith, J.H.; Mikolajek, H.; Le Bas, A.
Deposited on : 2020-06-21
Resolution : 3.31 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 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 : 2.14.4.dev1
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) : 2.14.4.dev1

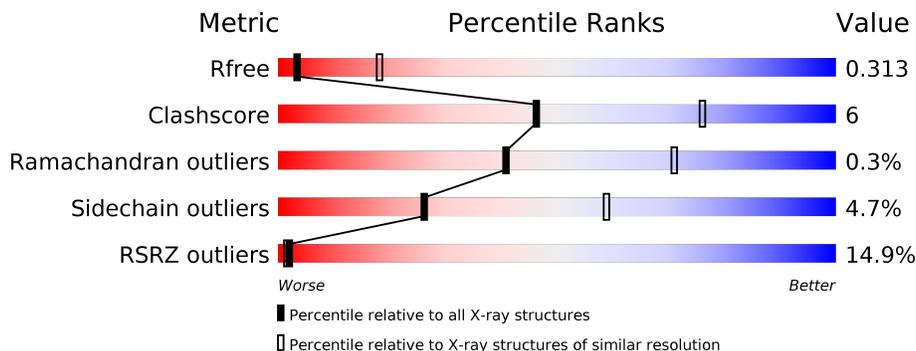
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.31 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	1089 (3.36-3.28)
Clashscore	141614	1137 (3.36-3.28)
Ramachandran outliers	138981	1115 (3.36-3.28)
Sidechain outliers	138945	1114 (3.36-3.28)
RSRZ outliers	127900	1059 (3.36-3.28)

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 on 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	HHH	216	
2	LLL	219	
3	EEE	197	
4	FFF	134	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 5906 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 CR3022 heavy.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	HHH	216	1609	1023	261	317	8	0	1	0

- Molecule 2 is a protein called CR3022 Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	LLL	219	1703	1070	282	347	4	0	0	0

- Molecule 3 is a protein called Spike glycoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	EEE	195	1562	1003	258	293	8	0	4	0

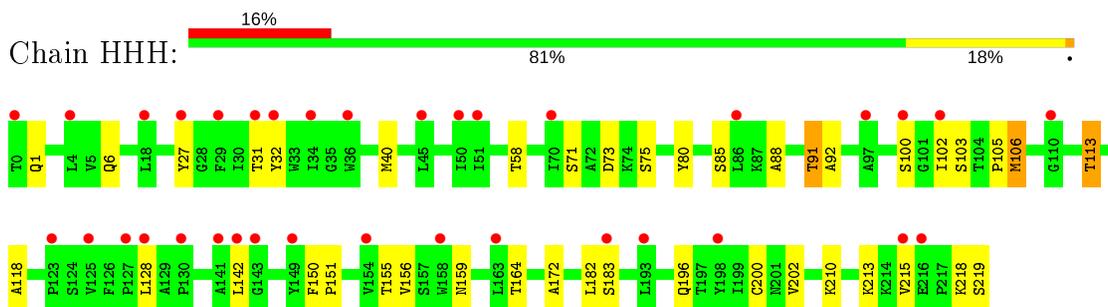
- Molecule 4 is a protein called Nanobody H11-H4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	FFF	128	1032	654	176	194	8	0	5	0

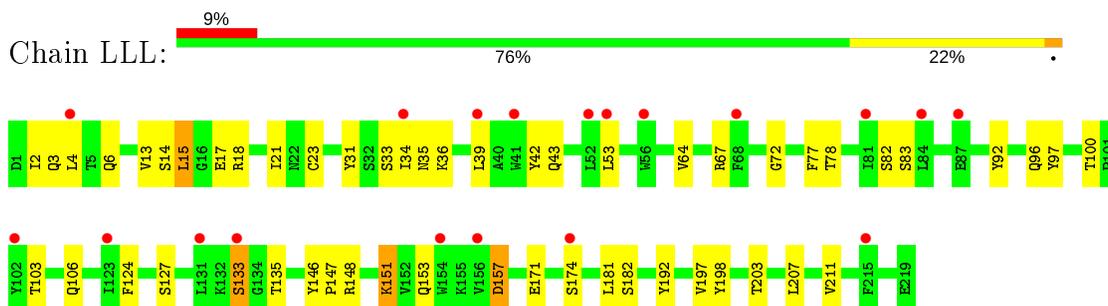
3 Residue-property plots [i](#)

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.

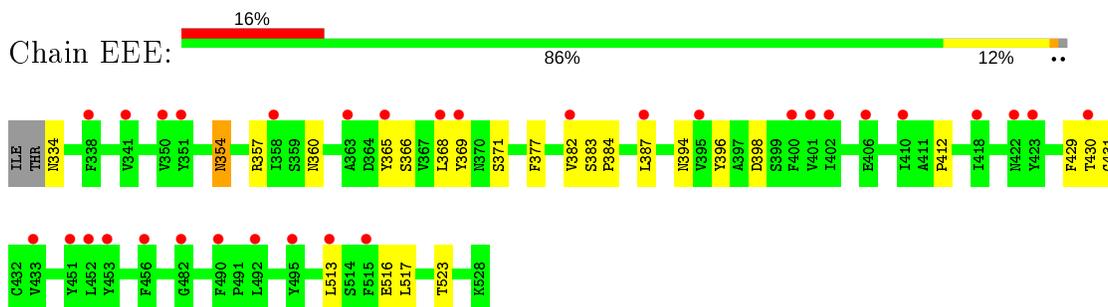
- Molecule 1: CR3022 heavy



- Molecule 2: CR3022 Light chain

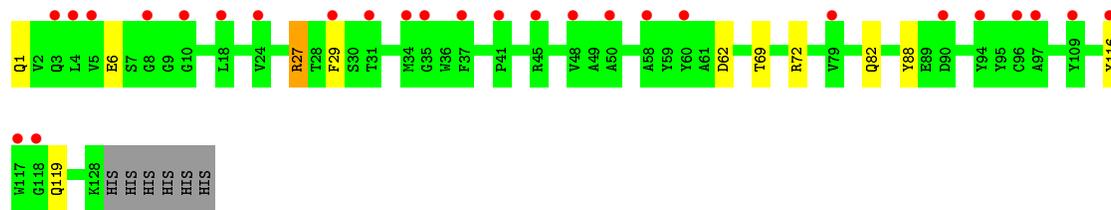


- Molecule 3: Spike glycoprotein



- Molecule 4: Nanobody H11-H4





4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α , β , γ	156.41Å 156.41Å 116.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.24 – 3.31 80.11 – 3.31	Depositor EDS
% Data completeness (in resolution range)	100.0 (80.24-3.31) 100.0 (80.11-3.31)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.08 (at 3.33Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.261 , 0.305 0.266 , 0.313	Depositor DCC
R_{free} test set	1088 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å ²)	159.7	Xtrriage
Anisotropy	0.259	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 178.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.86	EDS
Total number of atoms	5906	wwPDB-VP
Average B, all atoms (Å ²)	199.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.12% 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	HHH	0.71	0/1654	0.89	0/2253
2	LLL	0.72	0/1741	0.94	0/2367
3	EEE	0.70	0/1618	0.94	0/2202
4	FFF	0.72	0/1066	0.86	0/1441
All	All	0.71	0/6079	0.91	0/8263

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 [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	HHH	1609	0	1585	27	0
2	LLL	1703	0	1649	36	0
3	EEE	1562	0	1491	23	0
4	FFF	1032	0	995	4	0
All	All	5906	0	5720	73	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 6.

The worst 5 of 73 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:LLL:96:GLN:HE21	2:LLL:103:THR:HG22	1.44	0.83
2:LLL:34:ILE:HG22	3:EEE:517:LEU:HB2	1.74	0.69
2:LLL:157:ASP:HA	2:LLL:197:VAL:HG22	1.78	0.66
2:LLL:133:SER:HB2	2:LLL:135:THR:HG22	1.78	0.66
1:HHH:128:LEU:HB3	2:LLL:124:PHE:CD1	2.32	0.65

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	HHH	213/216 (99%)	199 (93%)	14 (7%)	0	100	100
2	LLL	217/219 (99%)	204 (94%)	11 (5%)	2 (1%)	17	49
3	EEE	197/197 (100%)	189 (96%)	8 (4%)	0	100	100
4	FFF	131/134 (98%)	130 (99%)	1 (1%)	0	100	100
All	All	758/766 (99%)	722 (95%)	34 (4%)	2 (0%)	41	71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	LLL	35	ASN
2	LLL	15	LEU

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	HHH	182/181 (101%)	170 (93%)	12 (7%)	16	47
2	LLL	194/194 (100%)	183 (94%)	11 (6%)	20	52
3	EEE	172/170 (101%)	169 (98%)	3 (2%)	60	79
4	FFF	107/108 (99%)	103 (96%)	4 (4%)	34	64
All	All	655/653 (100%)	625 (95%)	30 (5%)	26	60

5 of 30 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	LLL	14	SER
2	LLL	127	SER
4	FFF	62	ASP
2	LLL	18	ARG
2	LLL	133	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	HHH	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	HHH	132:SER	C	137:GLY	N	3.83

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	HHH	216/216 (100%)	0.99	35 (16%) 1 1	136, 194, 241, 279	0
2	LLL	219/219 (100%)	0.71	19 (8%) 10 10	120, 176, 223, 254	0
3	EEE	195/197 (98%)	0.81	32 (16%) 1 1	132, 180, 240, 281	0
4	FFF	128/134 (95%)	1.03	27 (21%) 1 0	203, 275, 331, 352	0
All	All	758/766 (98%)	0.87	113 (14%) 2 1	120, 191, 292, 352	0

The worst 5 of 113 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	FFF	109	TYR	8.5
1	HHH	142	LEU	7.4
4	FFF	8	GLY	6.7
4	FFF	35	GLY	6.1
4	FFF	118	GLY	5.9

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

There are no such residues in this entry.