



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 15, 2024 – 01:26 pm GMT

PDB ID : 6ZBP
Title : H11-H4 complex with SARS-CoV-2
Authors : Naismith, J.H.; Huo, J.; Mikolajek, H.; Ward, P.; Dumoux, M.; Owens, R.J.;
LeBas, A.
Deposited on : 2020-06-08
Resolution : 1.85 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
Mogul : 1.8.4, CSD as541be (2020)
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

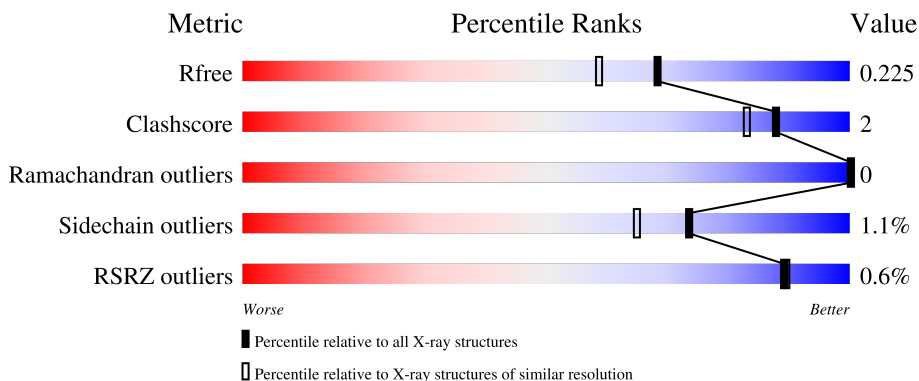
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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 $\leq 5\%$. 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	EEE	210	 89% 7%
2	FFF	134	 88% 7%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 5215 atoms, of which 2499 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 glycoprotein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	EEE	195	3050	1003	1488	258	293	8	90	4	0

There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
EEE	533	LYS	-	expression tag	UNP P0DTC2
EEE	534	HIS	-	expression tag	UNP P0DTC2
EEE	535	HIS	-	expression tag	UNP P0DTC2
EEE	536	HIS	-	expression tag	UNP P0DTC2
EEE	537	HIS	-	expression tag	UNP P0DTC2
EEE	538	HIS	-	expression tag	UNP P0DTC2
EEE	539	HIS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called H11-H4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	FFF	128	2029	654	997	176	194	8	58	5	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C₈H₁₅NO₆).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
3	EEE	1	28	8	14	1	5	3	0

- Molecule 4 is SULFATE ION (three-letter code: SO₄) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
4	FFF	1	5	4	1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	EEE	72	Total 72	O 72	0	0
5	FFF	31	Total 31	O 31	0	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.

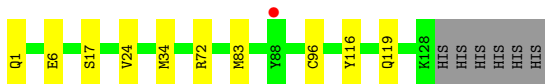
- Molecule 1: Spike glycoprotein

Chain EEE:  89% 7%



- Molecule 2: H11-H4

Chain FFF:  88% 7%



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	73.15Å 73.15Å 131.67Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	57.15 – 1.85 57.09 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.7 (57.15-1.85) 99.7 (57.09-1.85)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.74 (at 1.86Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.185 , 0.217 0.194 , 0.225	Depositor DCC
R_{free} test set	1808 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å ²)	38.3	Xtrriage
Anisotropy	0.192	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 51.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.030 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5215	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.71% 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

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, SO4

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	EEE	0.71	0/1618	0.90	0/2202
2	FFF	0.70	0/1066	0.88	0/1441
All	All	0.71	0/2684	0.89	0/3643

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	EEE	1562	1488	1490	4	2
2	FFF	1032	997	995	6	0
3	EEE	14	14	13	0	0
4	FFF	5	0	0	0	0
5	EEE	72	0	0	0	0
5	FFF	31	0	0	0	0
All	All	2716	2499	2498	10	2

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 2.

The worst 5 of 10 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:EEE:360:ASN:HA	1:EEE:523:THR:OG1	1.98	0.64
2:FFF:6:GLU:OE1	2:FFF:119:GLN:OE1	2.16	0.63
1:EEE:382:VAL:HG21	1:EEE:387:LEU:HD21	1.96	0.47
2:FFF:1:GLN:HA	2:FFF:116:TYR:CZ	2.52	0.45
2:FFF:6:GLU:HG3	2:FFF:96[A]:CYS:SG	2.58	0.44

All (2) 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:EEE:354:ASN:OD1	1:EEE:357:ARG:HH12[5_554]	1.54	0.06
1:EEE:354:ASN:OD1	1:EEE:357:ARG:NH1[5_554]	2.18	0.02

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	EEE	197/210 (94%)	191 (97%)	6 (3%)	0	100	100
2	FFF	131/134 (98%)	130 (99%)	1 (1%)	0	100	100
All	All	328/344 (95%)	321 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	EEE	172/183 (94%)	170 (99%)	2 (1%)	71	62
2	FFF	107/108 (99%)	106 (99%)	1 (1%)	78	72
All	All	279/291 (96%)	276 (99%)	3 (1%)	73	65

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	EEE	354	ASN
1	EEE	377	PHE
2	FFF	72	ARG

Sometimes 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](#)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	SO4	FFF	201	-	4,4,4	0.40	0	6,6,6	0.12	0
3	NAG	EEE	601	1	14,14,15	0.43	0	17,19,21	1.07	2 (11%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	EEE	601	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	EEE	601	NAG	O5-C1-C2	-2.12	107.94	111.29
3	EEE	601	NAG	C3-C4-C5	2.03	113.86	110.24

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	EEE	601	NAG	C4-C5-C6-O6
3	EEE	601	NAG	O5-C5-C6-O6

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 [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	EEE	195/210 (92%)	0.06	1 (0%) 91 91	33, 50, 78, 114	0
2	FFF	128/134 (95%)	0.02	1 (0%) 86 86	36, 49, 80, 112	0
All	All	323/344 (93%)	0.05	2 (0%) 89 89	33, 49, 80, 114	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	FFF	88	TYR	3.3
1	EEE	445	VAL	2.0

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, 95th 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
3	NAG	EEE	601	14/15	0.92	0.19	62,74,86,89	3
4	SO4	FFF	201	5/5	0.96	0.10	50,65,69,88	0

6.5 Other polymers [i](#)

There are no such residues in this entry.