



Full wwPDB X-ray Structure Validation Report ⓘ

Oct 1, 2024 – 01:57 pm BST

PDB ID : 8QP7
Title : Crystal structure of Hepatitis C Virus E2 glycoprotein epitopeI 411-424 scaffold design 4CIL_04
Authors : Nagarathinam, K.; Cramer, J.T.; Krey, T.
Deposited on : 2023-09-30
Resolution : 2.05 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.003 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.39

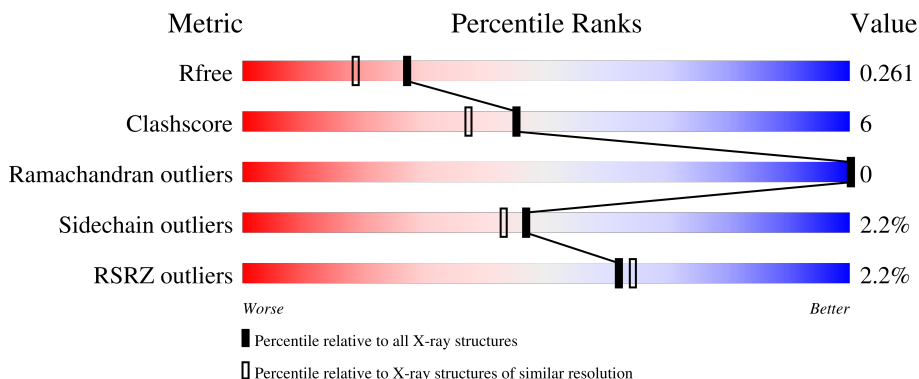
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.05 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}	164625	2096 (2.04-2.04)
Clashscore	180529	2229 (2.04-2.04)
Ramachandran outliers	177936	2217 (2.04-2.04)
Sidechain outliers	177891	2217 (2.04-2.04)
RSRZ outliers	164620	2096 (2.04-2.04)

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	296	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 2211 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 Yop effector YopM,Internalin B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	274	2084	1326	355	402	1	0	0	0

There are 34 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P74988
A	2	GLY	-	expression tag	UNP P74988
A	3	ALA	-	expression tag	UNP P74988
A	4	MET	-	expression tag	UNP P74988
A	5	GLY	-	expression tag	UNP P74988
A	24	GLY	PRO	engineered mutation	UNP P74988
A	40	ALA	CYS	engineered mutation	UNP P74988
A	45	SER	ALA	engineered mutation	UNP P74988
A	53	LEU	SER	engineered mutation	UNP P0DQD2
A	54	GLY	ASP	engineered mutation	UNP P0DQD2
A	55	LEU	ILE	engineered mutation	UNP P0DQD2
A	56	SER	LYS	engineered mutation	UNP P0DQD2
A	58	LEU	VAL	engineered mutation	UNP P0DQD2
A	59	PRO	GLN	engineered mutation	UNP P0DQD2
A	209	ILE	CYS	engineered mutation	UNP P0DQD2
A	228	ASP	ASN	engineered mutation	UNP P0DQD2
A	234	THR	VAL	engineered mutation	UNP P0DQD2
A	237	ALA	GLU	engineered mutation	UNP P0DQD2
A	266	ARG	TYR	engineered mutation	UNP P0DQD2
A	268	LEU	PRO	engineered mutation	UNP P0DQD2
A	270	ASN	THR	engineered mutation	UNP P0DQD2
A	273	GLY	LYS	engineered mutation	UNP P0DQD2
A	274	SER	ALA	engineered mutation	UNP P0DQD2
A	275	TRP	LYS	engineered mutation	UNP P0DQD2
A	277	ILE	ARG	engineered mutation	UNP P0DQD2
A	279	SER	HIS	engineered mutation	UNP P0DQD2
A	289	GLY	-	expression tag	UNP P0DQD2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	290	SER	-	expression tag	UNP P0DQD2
A	291	GLU	-	expression tag	UNP P0DQD2
A	292	ASN	-	expression tag	UNP P0DQD2
A	293	LEU	-	expression tag	UNP P0DQD2
A	294	TYR	-	expression tag	UNP P0DQD2
A	295	PHE	-	expression tag	UNP P0DQD2
A	296	GLN	-	expression tag	UNP P0DQD2

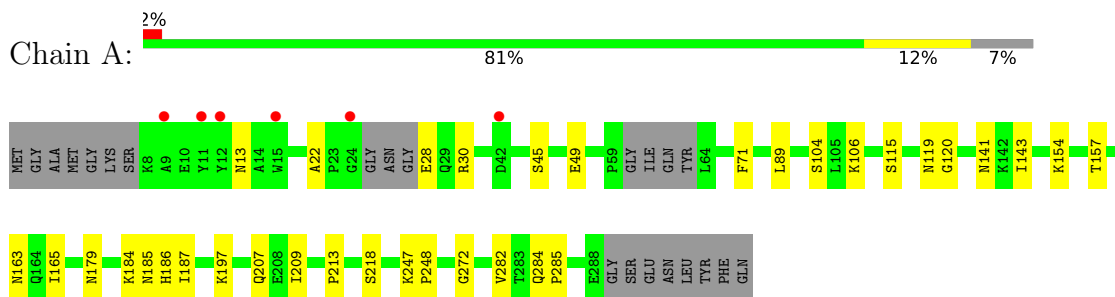
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	127	Total O 127 127	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: Yop effector YopM,Internalin B



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	60.29Å 31.67Å 135.75Å 90.00° 94.31° 90.00°	Depositor
Resolution (Å)	30.06 – 2.05 30.06 – 2.05	Depositor EDS
% Data completeness (in resolution range)	97.2 (30.06-2.05) 98.2 (30.06-2.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.20 (at 2.05Å)	Xtrriage
Refinement program	PHENIX 1.18_3861	Depositor
R, R_{free}	0.214 , 0.262 0.215 , 0.261	Depositor DCC
R_{free} test set	810 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	32.4	Xtrriage
Anisotropy	0.460	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 50.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	2211	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.87% 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	0.23	0/2119	0.44	0/2887

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	A	2084	0	2050	24	0
2	A	127	0	0	4	0
All	All	2211	0	2050	24	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.

All (24) 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:A:143:ILE:H	1:A:163:ASN:HD22	1.33	0.76
1:A:141:ASN:HB2	1:A:163:ASN:HD21	1.58	0.68
1:A:165:ILE:H	1:A:185:ASN:HD22	1.41	0.67
1:A:184:LYS:NZ	2:A:301:HOH:O	2.31	0.63

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:163:ASN:HB2	1:A:185:ASN:HD21	1.65	0.62
1:A:106:LYS:NZ	2:A:303:HOH:O	2.34	0.60
1:A:284:GLN:NE2	2:A:302:HOH:O	2.34	0.60
1:A:28:GLU:HG2	1:A:30:ARG:HG3	1.87	0.56
1:A:154:LYS:NZ	2:A:307:HOH:O	2.38	0.55
1:A:119:ASN:H	1:A:141:ASN:HD21	1.54	0.54
1:A:49:GLU:HA	1:A:71:PHE:HB2	1.88	0.53
1:A:143:ILE:H	1:A:163:ASN:ND2	2.05	0.53
1:A:119:ASN:H	1:A:141:ASN:ND2	2.09	0.51
1:A:185:ASN:HB2	1:A:207:GLN:HE22	1.78	0.47
1:A:187:ILE:H	1:A:207:GLN:HE21	1.64	0.45
1:A:209:ILE:HB	1:A:282:VAL:HG22	1.98	0.45
1:A:157:THR:HG23	1:A:179:ASN:HB2	2.00	0.44
1:A:197:LYS:HD3	1:A:272:GLY:HA2	2.01	0.43
1:A:247:LYS:HA	1:A:248:PRO:HA	1.88	0.43
1:A:213:PRO:HA	1:A:285:PRO:HG2	2.02	0.41
1:A:120:GLY:N	1:A:141:ASN:HD22	2.19	0.41
1:A:22:ALA:HB2	1:A:30:ARG:HD2	2.03	0.41
1:A:185:ASN:HB3	1:A:186:HIS:H	1.73	0.41
1:A:89:LEU:HD23	1:A:89:LEU:HA	1.87	0.40

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	268/296 (90%)	259 (97%)	9 (3%)	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	A	227/263 (86%)	222 (98%)	5 (2%)	47 43

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	13	ASN
1	A	45	SER
1	A	104	SER
1	A	115	SER
1	A	218	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	13	ASN
1	A	140	ASN
1	A	141	ASN
1	A	163	ASN
1	A	164	GLN
1	A	185	ASN
1	A	198	ASN
1	A	207	GLN
1	A	284	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 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	A	274/296 (92%)	0.28	6 (2%) 62 64	24, 36, 65, 86	0

All (6) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	TYR	2.7
1	A	15	TRP	2.7
1	A	24	GLY	2.6
1	A	12	TYR	2.2
1	A	42	ASP	2.1
1	A	9	ALA	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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.