



Full wwPDB X-ray Structure Validation Report ⓘ

May 18, 2017 – 01:35 PM EDT

PDB ID : 5NPH
Title : Structure of the Hepatitis C virus strain J4 glycoprotein E2 antigenic region
532-540 bound to the Fab fragment of the non-neutralizing antibody DAO5
Authors : Vasiliauskaite, I.; Rey, F.A.; Krey, T.
Deposited on : 2017-04-16
Resolution : 1.70 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

<http://wwpdb.org/validation/2016/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) : 1.9-1692
EDS : rb-20029077
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029077

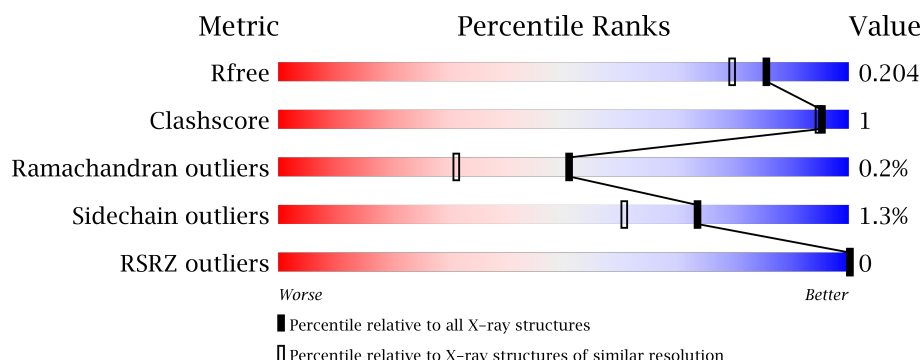
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION




The reported resolution of this entry is 1.70 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}	100719	3453 (1.70-1.70)
Clashscore	112137	3876 (1.70-1.70)
Ramachandran outliers	110173	3815 (1.70-1.70)
Sidechain outliers	110143	3815 (1.70-1.70)
RSRZ outliers	101464	3491 (1.70-1.70)

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. 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	12	 75% 25%
2	H	260	 80% 18% •
3	L	216	 94% 5% •

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3690 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 Genome polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	9	Total	C	N	O	S	0	0	0
			72	43	11	17	1			

- Molecule 2 is a protein called Heavy chain of Fab fragment derived from non-neutralizing antibody DAO5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	214	Total	C	N	O	S	6	0	0
			1622	1037	263	316	6			

- Molecule 3 is a protein called Light chain of Fab fragment derived from non-neutralizing antibody DAO5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	214	Total	C	N	O	S	2	0	0
			1667	1040	283	338	6			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	6	Total	O	0	0
			6	6		
4	H	162	Total	O	0	0
			162	162		
4	L	161	Total	O	0	0
			161	161		

i

- Molecule 1: Genome polyprotein



- [illegible]

-

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	48.68Å 80.72Å 54.56Å 90.00° 95.64° 90.00°	Depositor
Resolution (Å)	34.00 – 1.70 48.44 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.9 (34.00-1.70) 96.9 (48.44-1.70)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.96 (at 1.70Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.179 , 0.203 0.176 , 0.204	Depositor DCC
R_{free} test set	2254 reflections (5.30%)	DCC
Wilson B-factor (Å ²)	16.0	Xtriage
Anisotropy	0.650	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 43.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3690	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.44% 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	0.59	0/71	0.76	0/94
2	H	0.52	0/1669	0.68	0/2280
3	L	0.50	0/1706	0.64	0/2315
All	All	0.51	0/3446	0.66	0/4689

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	72	0	68	0	0
2	H	1622	0	1587	3	0
3	L	1667	0	1596	6	0
4	A	6	0	0	0	0
4	H	162	0	0	1	0
4	L	161	0	0	0	0
All	All	3690	0	3251	9	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 1.

All (9) 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)
3:L:150:ILE:HD12	3:L:155:ARG:HD3	1.64	0.78
3:L:13:VAL:HG11	3:L:19:VAL:CG1	2.18	0.74
3:L:13:VAL:HG11	3:L:19:VAL:HG11	1.72	0.72
3:L:150:ILE:HD11	3:L:179:LEU:HD21	1.74	0.69
2:H:80:HIS:HD2	4:H:313:HOH:O	1.92	0.51
3:L:91:THR:HA	3:L:96:LEU:HD22	1.93	0.51
3:L:13:VAL:HG11	3:L:19:VAL:HG12	1.91	0.50
2:H:149:VAL:HG11	2:H:157:VAL:HG11	2.00	0.43
2:H:206:HIS:HB3	2:H:211:THR:HB	2.03	0.41

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	7/12 (58%)	7 (100%)	0	0	100	100
2	H	210/260 (81%)	208 (99%)	2 (1%)	0	100	100
3	L	212/216 (98%)	206 (97%)	5 (2%)	1 (0%)	32	15
All	All	429/488 (88%)	421 (98%)	7 (2%)	1 (0%)	51	31

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	77	SER

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	9/11 (82%)	9 (100%)	0	100	100
2	H	186/217 (86%)	185 (100%)	1 (0%)	91	87
3	L	192/194 (99%)	188 (98%)	4 (2%)	59	40
All	All	387/422 (92%)	382 (99%)	5 (1%)	73	60

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

Mol	Chain	Res	Type
2	H	203	ASN
3	L	0	SER
3	L	56	SER
3	L	155	ARG
3	L	199	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	540	ASN
2	H	80	HIS
3	L	92	ASN

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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	9/12 (75%)	-0.18	0 100 100	16, 19, 32, 54	0
2	H	214/260 (82%)	-0.44	0 100 100	12, 17, 30, 44	1 (0%)
3	L	214/216 (99%)	-0.41	0 100 100	12, 17, 32, 44	1 (0%)
All	All	437/488 (89%)	-0.42	0 100 100	12, 17, 32, 54	2 (0%)

There are no RSRZ outliers to report.

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 carbohydrates 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.