



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 17, 2026 – 08:43 PM UTC

PDB ID : 4E7A / pdb_00004e7a
Title : Crystal structure of a product state assembly of HCV NS5B genotype 2a JFH-1 isolate with beta hairpin deletion bound to primer-template RNA with a 2',3'-ddC
Authors : Edwards, T.E.; Mosley, R.T.
Deposited on : 2012-03-16
Resolution : 3.00 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

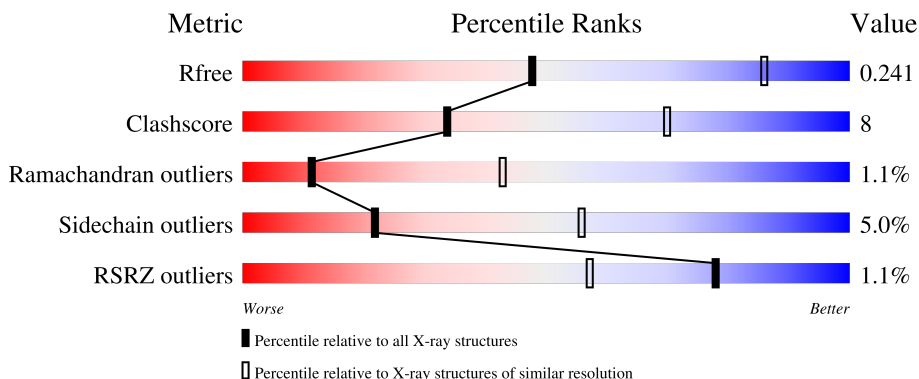
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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}	180053	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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	P	7	<div> <div>14%</div> <div>43%</div> <div>29%</div> <div>29%</div> </div>
1	T	7	<div> <div>43%</div> <div>14%</div> <div>14%</div> <div>29%</div> </div>
2	A	572	<div> <div>%</div> <div>76%</div> <div>16%</div> <div>6%</div> </div>

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 4253 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 DNA/RNA hybrid called 5'-R(*CP*AP*UP*GP*GP*C)-D(P*(DOC))-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	5	Total	C	N	O	P	0	0	0
			96	43	16	32	5			
1	P	5	Total	C	N	O	P	0	0	0
			101	47	18	32	4			

- Molecule 2 is a protein called RNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	536	Total	C	N	O	S	0	0	0
			4038	2549	710	752	27			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP Q99IB8
A	0	SER	-	expression tag	UNP Q99IB8
A	86	GLN	GLU	engineered mutation	UNP Q99IB8
A	87	GLN	GLU	engineered mutation	UNP Q99IB8
A	444	GLY	-	linker	UNP Q99IB8
A	445	GLY	-	linker	UNP Q99IB8
A	571	LEU	-	expression tag	UNP Q99IB8
A	572	GLU	-	expression tag	UNP Q99IB8
A	573	HIS	-	expression tag	UNP Q99IB8
A	574	HIS	-	expression tag	UNP Q99IB8
A	575	HIS	-	expression tag	UNP Q99IB8
A	576	HIS	-	expression tag	UNP Q99IB8
A	577	HIS	-	expression tag	UNP Q99IB8
A	578	HIS	-	expression tag	UNP Q99IB8

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	P	1	Total 1	O 1	0	0
3	A	17	Total 17	O 17	0	0

● Molecule 1: 5'-R(*CP*AP*UP*GP*GP*C)-D(P*(DOC))-3'



4 Data and refinement statistics

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α , β , γ	142.73Å 142.73Å 91.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 3.00 50.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-3.00) 99.5 (50.00-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.41 (at 3.01Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.194 , 0.241 0.199 , 0.241	Depositor DCC
R_{free} test set	1093 reflections (5.13%)	wwPDB-VP
Wilson B-factor (Å ²)	65.1	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 40.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.041 for h,-h-k,-l	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4253	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% 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](#)

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

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	P	0.54	0/92	1.18	2/142 (1.4%)
1	T	0.46	0/86	1.22	1/132 (0.8%)
2	A	0.78	0/4132	0.94	4/5639 (0.1%)
All	All	0.77	0/4310	0.95	7/5913 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	T	5	G	O3'-P-O5'	-7.57	92.64	104.00
1	P	6	C	C3'-C2'-O2'	6.47	120.40	110.70
1	P	3	U	P-O3'-C3'	6.10	129.35	120.20
2	A	93	PRO	CA-C-N	5.56	125.66	119.32
2	A	93	PRO	C-N-CA	5.56	125.66	119.32

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	P	101	0	56	0	0
1	T	96	0	51	2	0
2	A	4038	0	3930	62	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	17	0	0	0	0
3	P	1	0	0	0	0
All	All	4253	0	4037	64	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 8.

The worst 5 of 64 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:A:85:LEU:C	2:A:85:LEU:HD23	2.04	0.82
2:A:71:ILE:HD12	2:A:297:VAL:HG22	1.74	0.70
2:A:2:MET:HE3	2:A:53:THR:HG22	1.72	0.70
2:A:508:ARG:CZ	2:A:530:VAL:HG11	2.22	0.70
2:A:532:THR:O	2:A:532:THR:HG23	1.96	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
2	A	534/572 (93%)	480 (90%)	48 (9%)	6 (1%)	11 43

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	16	PRO
2	A	85	LEU
2	A	130	PRO
2	A	270	LYS
2	A	469	LEU

5.3.2 Protein sidechains ⓘ

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
2	A	421/486 (87%)	400 (95%)	21 (5%)	22 56

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	336	LEU
2	A	367	SER
2	A	480	HIS
2	A	369	ASN
2	A	366	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
2	A	63	HIS
2	A	184	GLN
2	A	194	GLN
2	A	327	GLN
2	A	467	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	DOC	T	7	1	16,19,20	0.36	0	20,26,29	0.45	0
1	DOC	P	7	1	16,19,20	0.41	0	20,26,29	0.55	0

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
1	DOC	T	7	1	-	2/7/18/19	0/2/2/2
1	DOC	P	7	1	-	0/7/18/19	0/2/2/2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	T	7	DOC	C3'-C4'-C5'-O5'
1	T	7	DOC	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

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	P	4/7 (57%)	1.14	1 (25%) 2 1	66, 71, 72, 117	0
1	T	4/7 (57%)	-0.17	0 100 100	53, 56, 64, 86	0
2	A	536/572 (93%)	-0.31	5 (0%) 81 61	33, 53, 78, 118	0
All	All	544/586 (92%)	-0.30	6 (1%) 78 57	33, 53, 79, 118	0

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	3	U	4.3
2	A	221	THR	4.3
2	A	222	ARG	3.2
2	A	543	ARG	3.0
2	A	365	SER	2.8

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	DOC	T	7	18/19	0.83	0.16	64,87,94,98	0
1	DOC	P	7	18/19	0.88	0.14	61,67,85,86	0

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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.