



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:15 AM GMT

PDB ID : 2WHO
Title : CRYSTAL STRUCTURE OF HEPATITIS C VIRUS NS5B POLYMERASE
FROM 1B GENOTYPE IN COMPLEX WITH A NON-NUCLEOSIDE IN-
HIBITOR
Authors : Di Marco, S.
Deposited on : 2009-05-05
Resolution : 2.00 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
<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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

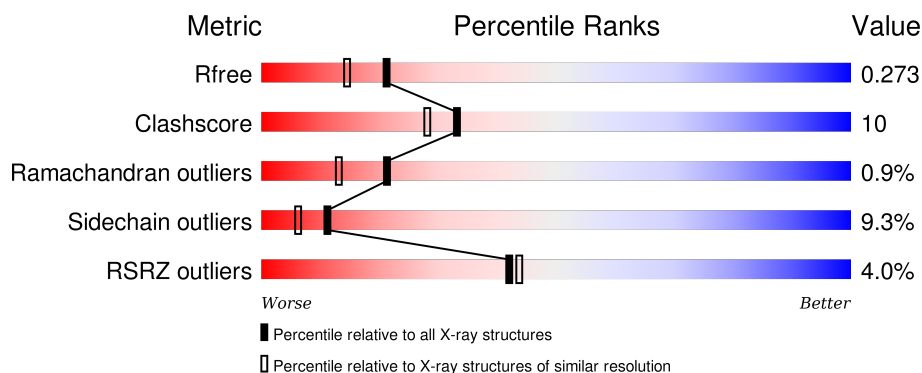
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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 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	536	<div> <div>4%</div> <div>75%</div> <div>19%</div> <div>5%</div> </div>
1	B	536	<div> <div>4%</div> <div>76%</div> <div>19%</div> <div>• •</div> </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 8817 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.

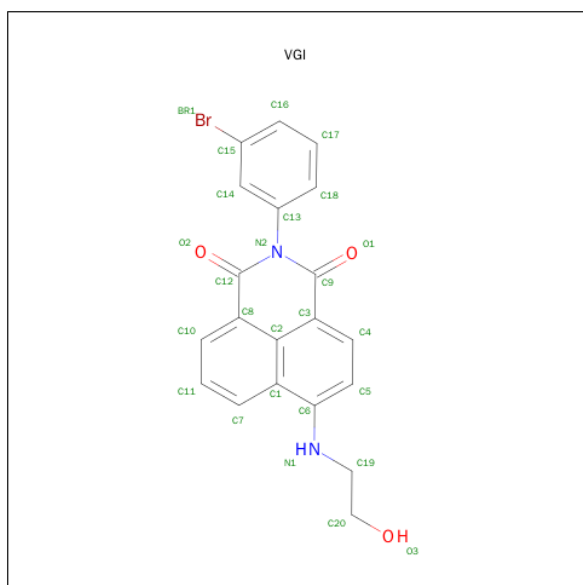
- Molecule 1 is a protein called RNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	531	Total	C	N	O	S	0	0	0
			4121	2595	727	768	31			
1	B	531	Total	C	N	O	S	0	0	0
			4121	2595	727	768	31			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	2	Total	Mn	0	0
			2	2		
2	A	2	Total	Mn	0	0
			2	2		

- Molecule 3 is 2-(3-BROMOPHENYL)-6-[(2-HYDROXYETHYL)AMINO]-1H-BENZO[DE]ISOQUINOLINE-1,3(2H)-DIONE (three-letter code: VGI) (formula: C₂₀H₁₅BrN₂O₃).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total 26	Br 1	C 20	N 2	O 3	0	0
3	B	1	Total 26	Br 1	C 20	N 2	O 3	0	0

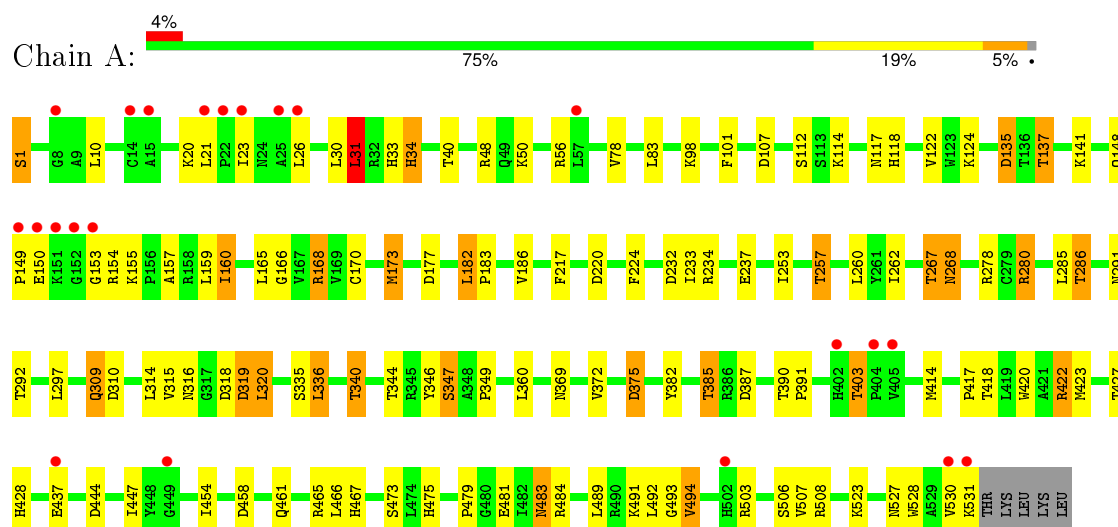
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	259	Total 259	O 259	0	0
4	B	260	Total 260	O 260	0	0

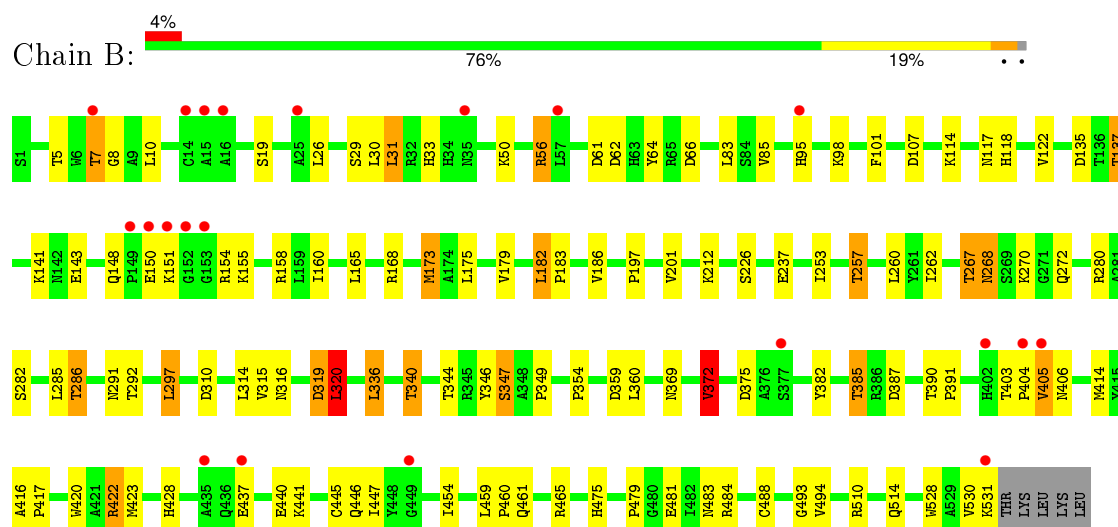
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: RNA-DIRECTED RNA POLYMERASE



• Molecule 1: RNA-DIRECTED RNA POLYMERASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	67.21Å 97.00Å 193.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	95.35 – 2.00 19.97 – 2.00	Depositor EDS
% Data completeness (in resolution range)	74.3 (95.35-2.00) 74.4 (19.97-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.211 , 0.266 0.221 , 0.273	Depositor DCC
R_{free} test set	3282 reflections (5.39%)	DCC
Wilson B-factor (Å ²)	25.5	Xtriage
Anisotropy	0.345	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 52.1	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	5 of 64166 reflections (0.008%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8817	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 41.96 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.1880e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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: MN, VGI

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	0.69	1/4209 (0.0%)	0.89	18/5712 (0.3%)
1	B	0.68	1/4209 (0.0%)	0.87	15/5712 (0.3%)
All	All	0.68	2/8418 (0.0%)	0.88	33/11424 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	173	MET	SD-CE	-8.57	1.29	1.77
1	B	173	MET	SD-CE	-5.58	1.46	1.77

The worst 5 of 33 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	422	ARG	NE-CZ-NH1	9.63	125.12	120.30
1	A	168	ARG	NE-CZ-NH2	-8.72	115.94	120.30
1	A	168	ARG	NE-CZ-NH1	8.42	124.51	120.30
1	A	444	ASP	CB-CG-OD2	8.17	125.65	118.30
1	B	310	ASP	CB-CG-OD2	7.45	125.01	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	7	THR	Peptide

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	4121	0	4134	90	0
1	B	4121	0	4134	83	0
2	A	2	0	0	0	0
2	B	2	0	0	0	0
3	A	26	0	15	3	0
3	B	26	0	15	3	0
4	A	259	0	0	18	0
4	B	260	0	0	18	0
All	All	8817	0	8298	171	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 10.

The worst 5 of 171 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:B:423:MET:HE2	3:B:1534:VGI:C14	1.87	1.04
1:A:423:MET:HE2	3:A:1534:VGI:H14	1.03	1.02
1:B:423:MET:CE	3:B:1534:VGI:H14	1.90	1.02
1:A:1:SER:HA	4:A:2003:HOH:O	1.62	0.98
1:A:423:MET:HE2	3:A:1534:VGI:C14	1.93	0.97

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	529/536 (99%)	510 (96%)	15 (3%)	4 (1%)	24	15
1	B	529/536 (99%)	506 (96%)	17 (3%)	6 (1%)	17	9
All	All	1058/1072 (99%)	1016 (96%)	32 (3%)	10 (1%)	21	13

5 of 10 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	268	ASN
1	B	268	ASN
1	A	149	PRO
1	A	347	SER
1	A	403	THR

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	
1	A	452/459 (98%)	409 (90%)	43 (10%)	11	6
1	B	452/459 (98%)	411 (91%)	41 (9%)	12	6
All	All	904/918 (98%)	820 (91%)	84 (9%)	11	6

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	B	29	SER
1	B	440	GLU
1	A	483	ASN
1	A	531	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN
1	B	58	GLN
1	B	461	GLN
1	A	475	HIS
1	B	33	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

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	$\# Z > 2$
3	VGI	A	1534	-	29,29,29	1.34	3 (10%)	31,42,42	1.48	5 (16%)
3	VGI	B	1534	-	29,29,29	1.47	3 (10%)	31,42,42	1.45	5 (16%)

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	VGI	A	1534	-	-	0/8/8/8	0/4/4/4
3	VGI	B	1534	-	-	2/8/8/8	0/4/4/4

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	1534	VGI	C13-N2	-4.92	1.40	1.46
3	A	1534	VGI	C13-N2	-4.08	1.41	1.46
3	A	1534	VGI	C6-N1	-2.41	1.32	1.36
3	B	1534	VGI	C6-N1	-2.07	1.32	1.36
3	B	1534	VGI	C12-N2	2.82	1.42	1.37

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1534	VGI	C5-C6-N1	-3.72	117.91	122.27
3	B	1534	VGI	C5-C6-N1	-2.92	118.85	122.27
3	B	1534	VGI	C14-C13-N2	-2.42	116.76	119.27
3	A	1534	VGI	C16-C15-C14	-2.19	118.12	121.44
3	A	1534	VGI	C14-C13-N2	-2.18	117.02	119.27

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1534	VGI	C5-C6-N1-C19
3	B	1534	VGI	C1-C6-N1-C19

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1534	VGI	3	0
3	B	1534	VGI	3	0

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	531/536 (99%)	0.16	22 (4%) 41 42	19, 31, 56, 71	0
1	B	531/536 (99%)	0.11	21 (3%) 42 44	20, 31, 51, 71	0
All	All	1062/1072 (99%)	0.13	43 (4%) 42 44	19, 31, 54, 71	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	152	GLY	5.9
1	A	150	GLU	5.8
1	B	153	GLY	5.6
1	A	14	CYS	5.4
1	B	150	GLU	5.3

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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
3	VGI	A	1534	26/26	0.95	0.14	0.21	43,48,51,52	0
3	VGI	B	1534	26/26	0.96	0.13	0.18	41,46,49,50	0
2	MN	B	1533	1/1	0.99	0.02	-2.94	28,28,28,28	0
2	MN	A	1533	1/1	1.00	0.04	-3.59	29,29,29,29	0
2	MN	B	1532	1/1	0.99	0.04	-	28,28,28,28	0
2	MN	A	1532	1/1	1.00	0.03	-	26,26,26,26	0

6.5 Other polymers [i](#)

There are no such residues in this entry.