wwPDB X-ray Structure Validation Summary Report (i)

Jan 31, 2016 – 06:42 PM GMT

PDB ID	:	1C2P
Title	:	HEPATITIS C VIRUS NS5B RNA-DEPENDENT RNA POLYMERASE
Authors	:	Lesburg, C.A.; Cable, M.B.; Ferrari, E.; Hong, Z.; Mannarino, A.F.; Weber, P.C.
Deposited on	:	1999-07-26
Resolution	:	1.90 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 (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

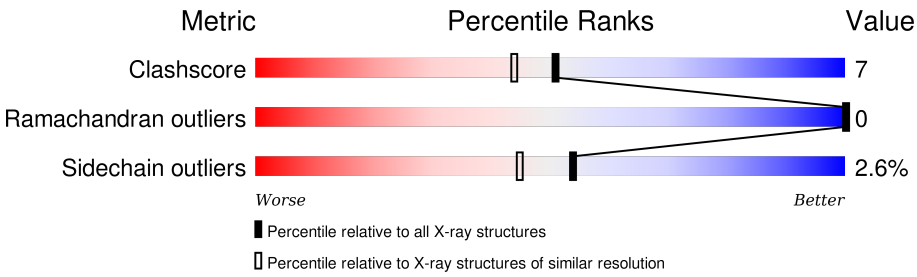
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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))
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)

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 <=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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	576	83% 13% • •
1	B	576	83% 13% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9438 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-DEPENDENT RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	558	4344	2741	769	804	19	11	0	0	0
1	B	561	4369	2756	775	808	19	11	0	0	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	ALA	-	EXPRESSION TAG	UNP P26663
A	-4	SER	-	EXPRESSION TAG	UNP P26663
A	-3	HIS	-	EXPRESSION TAG	UNP P26663
A	-2	HIS	-	EXPRESSION TAG	UNP P26663
A	-1	HIS	-	EXPRESSION TAG	UNP P26663
A	0	HIS	-	EXPRESSION TAG	UNP P26663
A	1	HIS	-	EXPRESSION TAG	UNP P26663
A	2	HIS	-	EXPRESSION TAG	UNP P26663
A	36	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	71	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	139	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	173	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	187	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	215	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	313	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	329	VAL	THR	CONFLICT	UNP P26663
A	338	ALA	VAL	CONFLICT	UNP P26663
A	343	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	414	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	423	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	426	MSE	MET	MODIFIED RESIDUE	UNP P26663
A	544	GLN	ARG	CONFLICT	UNP P26663
B	-5	ALA	-	EXPRESSION TAG	UNP P26663
B	-4	SER	-	EXPRESSION TAG	UNP P26663
B	-3	HIS	-	EXPRESSION TAG	UNP P26663

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	EXPRESSION TAG	UNP P26663
B	-1	HIS	-	EXPRESSION TAG	UNP P26663
B	0	HIS	-	EXPRESSION TAG	UNP P26663
B	1	HIS	-	EXPRESSION TAG	UNP P26663
B	2	HIS	-	EXPRESSION TAG	UNP P26663
B	36	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	71	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	139	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	173	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	187	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	215	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	313	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	329	VAL	THR	CONFLICT	UNP P26663
B	338	ALA	VAL	CONFLICT	UNP P26663
B	343	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	414	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	423	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	426	MSE	MET	MODIFIED RESIDUE	UNP P26663
B	544	GLN	ARG	CONFLICT	UNP P26663

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	365	365	365	0	0
			Total	O		
2	B	360	360	360	0	0

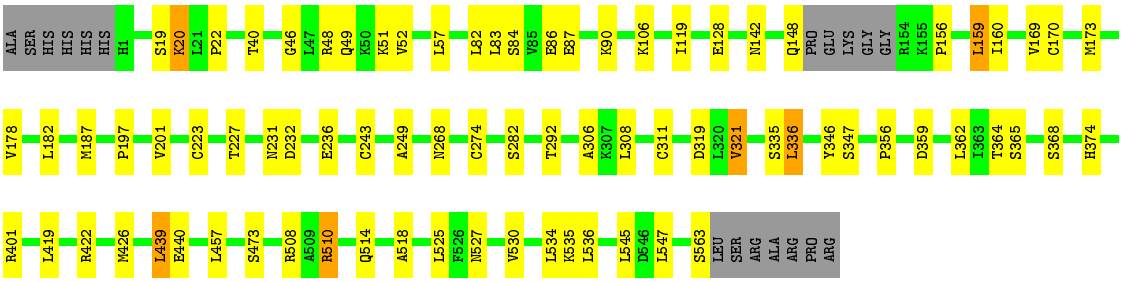
3 Residue-property plots (i)

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.

Note EDS was not executed.

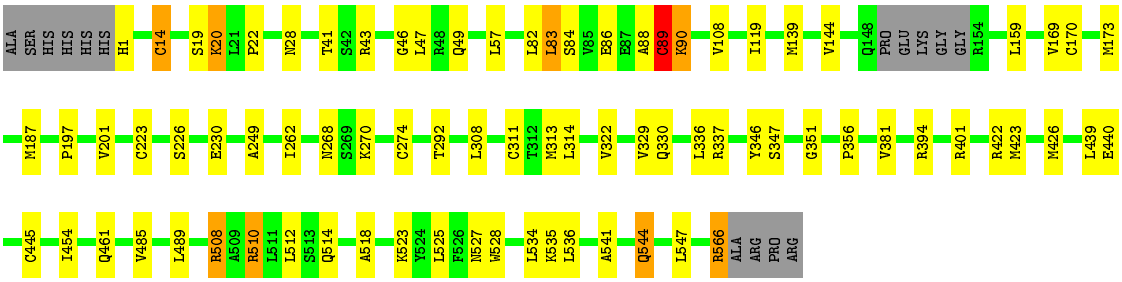
- Molecule 1: RNA-DEPENDENT RNA POLYMERASE

Chain A:



- Molecule 1: RNA-DEPENDENT RNA POLYMERASE

Chain B:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	86.84Å 105.20Å 127.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.90	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-1.90)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	X-PLOR 3.1	Depositor
R, R_{free}	0.211 , 0.255	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9438	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

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	0.49	0/4429	0.67	1/5993 (0.0%)
1	B	0.52	1/4454 (0.0%)	0.75	7/6026 (0.1%)
All	All	0.50	1/8883 (0.0%)	0.71	8/12019 (0.1%)

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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	88	ALA	C-N	9.72	1.56	1.34

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	89	CYS	O-C-N	-17.73	94.33	122.70
1	B	88	ALA	C-N-CA	-7.65	102.57	121.70
1	B	89	CYS	CB-CA-C	-7.16	96.08	110.40
1	B	90	LYS	N-CA-C	-6.53	93.37	111.00
1	B	89	CYS	CA-C-N	6.43	131.35	117.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	89	CYS	Mainchain

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	4344	0	4357	62	0
1	B	4369	0	4386	59	0
2	A	365	0	0	11	0
2	B	360	0	0	6	0
All	All	9438	0	8743	121	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 7.

The worst 5 of 121 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:223:CYS:SG	2:A:750:HOH:O	2.06	1.13
1:A:169:VAL:HG12	1:A:173:MSE:HE2	1.35	1.08
1:A:90:LYS:HD2	2:A:884:HOH:O	1.52	1.07
1:B:169:VAL:HG12	1:B:173:MSE:HE2	1.39	1.02
1:B:89:CYS:HB3	1:B:108:VAL:CG1	1.96	0.95

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	554/576 (96%)	539 (97%)	15 (3%)	0	100	100
1	B	557/576 (97%)	545 (98%)	12 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	1111/1152 (96%)	1084 (98%)	27 (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	A	475/478 (99%)	464 (98%)	11 (2%)	58	51
1	B	478/478 (100%)	464 (97%)	14 (3%)	50	40
All	All	953/956 (100%)	928 (97%)	25 (3%)	54	45

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	14	CYS
1	B	47	LEU
1	B	547	LEU
1	B	20	LYS
1	B	57	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	49	GLN
1	B	544	GLN
1	B	461	GLN
1	A	527	ASN
1	B	527	ASN

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)

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)

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates (i)

EDS was not executed - this section will therefore be empty.

6.4 Ligands (i)

EDS was not executed - this section will therefore be empty.

6.5 Other polymers (i)

EDS was not executed - this section will therefore be empty.