WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 12, 2023 – 10:58 PM EDT

|		
PDB ID	:	4OOW
Title	:	HCV NS5B polymerase with a fragment of quercetagetin
Authors	:	Guichou, J.F.; Ahmed-Belkacem, A.; Rozenn, B.; Nazim, N.; Hernandez, E.; Pallier, C.; Pawlotsky, J.M.
Deposited on	:	2014-02-04
Resolution	:	2.57 Å(reported)

1   Overall quality at a glance ⓘ

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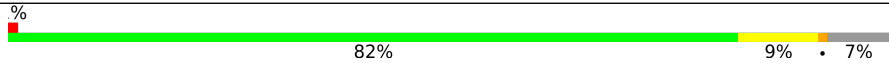*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.57 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.
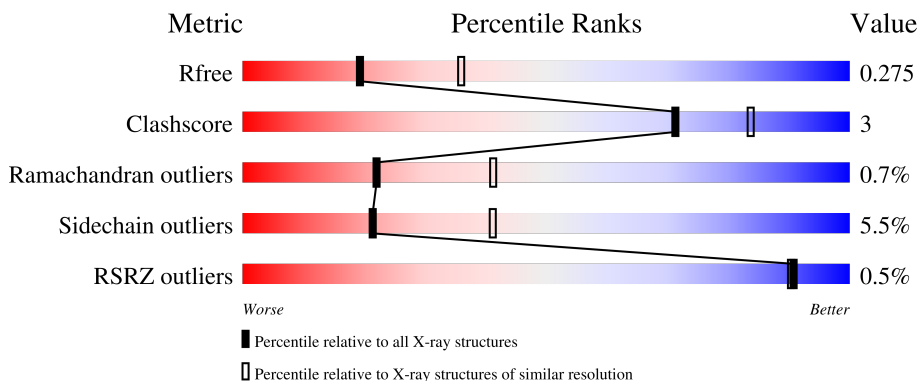
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3676 (2.60-2.56)
Clashscore	141614	4049 (2.60-2.56)
Ramachandran outliers	138981	3979 (2.60-2.56)
Sidechain outliers	138945	3979 (2.60-2.56)
RSRZ outliers	127900	3614 (2.60-2.56)

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 <=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	578	84% ... 10% ··
1	B	578	82% ... 9% · 7%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8841 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	555	Total	C	N	O	S	0	5	0
			4357	2748	771	806	32			
1	B	538	Total	C	N	O	S	0	8	0
			4250	2686	749	783	32			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	571	LEU	-	expression tag	UNP O92972
A	572	GLU	-	expression tag	UNP O92972
A	573	HIS	-	expression tag	UNP O92972
A	574	HIS	-	expression tag	UNP O92972
A	575	HIS	-	expression tag	UNP O92972
A	576	HIS	-	expression tag	UNP O92972
A	577	HIS	-	expression tag	UNP O92972
A	578	HIS	-	expression tag	UNP O92972
B	571	LEU	-	expression tag	UNP O92972
B	572	GLU	-	expression tag	UNP O92972
B	573	HIS	-	expression tag	UNP O92972
B	574	HIS	-	expression tag	UNP O92972
B	575	HIS	-	expression tag	UNP O92972
B	576	HIS	-	expression tag	UNP O92972
B	577	HIS	-	expression tag	UNP O92972
B	578	HIS	-	expression tag	UNP O92972

- Molecule 2 is CATECHOL (three-letter code: CAQ) (formula: $C_6H_6O_2$).

CAQ
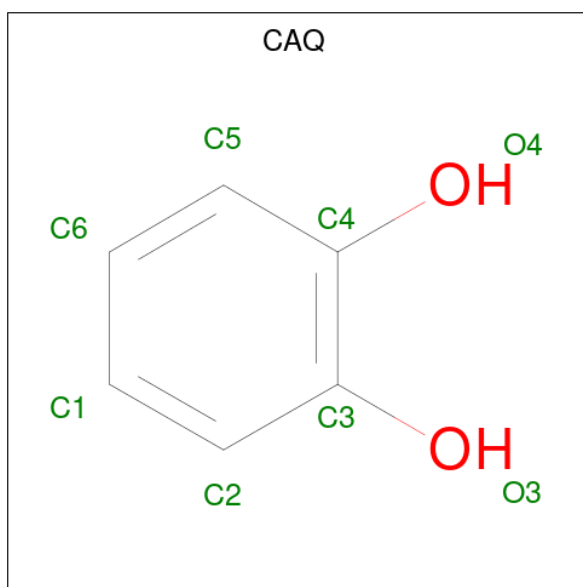
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total 8	C 6	O 2	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	135	Total 135	O 135	0	0
3	B	91	Total 91	O 91	0	0

3 Residue-property plots ⓘ

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: RNA-directed RNA polymerase

Chain A:

• Molecule 1: RNA-directed RNA polymerase

Chain B:

4   Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	106.60Å 108.87Å 134.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.90 – 2.57 84.74 – 2.57	Depositor EDS
% Data completeness (in resolution range)	94.7 (48.90-2.57) 94.7 (84.74-2.57)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.88 (at 2.58Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.216 , 0.277 0.215 , 0.275	Depositor DCC
R_{free} test set	2433 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å^2)	32.7	Xtriage
Anisotropy	0.557	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 33.9	EDS
L-test for twinning [2]	$<|L|> = 0.49, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	0.018 for k,h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	8841	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $<|L|>$, $<L^2>$ 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: CAQ

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	0.47	0/4466	0.69	4/6056 (0.1%)
1	B	0.44	0/4363	0.66	1/5911 (0.0%)
All	All	0.45	0/8829	0.68	5/11967 (0.0%)

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	A	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	57	LEU	CA-CB-CG	5.73	128.48	115.30
1	A	337	ARG	NE-CZ-NH1	-5.53	117.53	120.30
1	A	31	LEU	CA-CB-CG	5.15	127.15	115.30
1	B	419	LEU	CA-CB-CG	5.08	126.97	115.30
1	A	501	ARG	NE-CZ-NH1	5.03	122.81	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	532	THR	Peptide

5.2 Too-close contacts ⓘ

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	4357	0	4391	30	0
1	B	4250	0	4281	30	0
2	A	8	0	6	0	0
3	A	135	0	0	1	0
3	B	91	0	0	1	0
All	All	8841	0	8678	60	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 3.

The worst 5 of 60 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:539:ILE:HG22	1:B:540:PRO:HD3	1.30	1.06
1:B:524:TYR:OH	1:B:537:THR:O	1.87	0.93
1:B:539:ILE:CG2	1:B:540:PRO:HD3	2.10	0.78
1:B:531:ARG:O	1:B:532:THR:HG22	1.98	0.63
1:B:470:SER:O	1:B:474:LEU:HG	2.03	0.59

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	554/578 (96%)	533 (96%)	18 (3%)	3 (0%)	29	50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	534/578 (92%)	508 (95%)	22 (4%)	4 (1%)	22	41
All	All	1088/1156 (94%)	1041 (96%)	40 (4%)	7 (1%)	22	45

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	517	ARG
1	A	533	LYS
1	A	535	LYS
1	B	17	GLU
1	B	518	ALA

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	478/493 (97%)	455 (95%)	23 (5%)	25	47
1	B	466/493 (94%)	436 (94%)	30 (6%)	17	34
All	All	944/986 (96%)	891 (94%)	53 (6%)	21	40

5 of 53 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	148	GLN
1	B	270	LYS
1	B	535	LYS
1	B	160[A]	ILE
1	B	220	ASP

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

Mol	Chain	Res	Type
1	A	110	ASN
1	A	446	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	475	HIS
1	B	273	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	$\#|Z| > 2$
2	CAQ	A	601	-	8,8,8	0.62	0	10,10,10	1.00	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
2	CAQ	A	601	-	-	-	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

6.1   Protein, DNA and RNA chains ⓘ

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(Å²)	Q<0.9
1	A	555/578 (96%)	-0.26	1 (0%)	95	95	9, 21, 43, 74	20 (3%)
1	B	538/578 (93%)	-0.12	5 (0%)	84	83	11, 31, 61, 82	22 (4%)
All	All	1093/1156 (94%)	-0.19	6 (0%)	91	90	9, 25, 57, 82	42 (3%)

The worst 5 of 6 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	511	LEU	2.9
1	A	546	ASP	2.2
1	B	526	PHE	2.2
1	B	517	ARG	2.2
1	B	540	PRO	2.1

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3   Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4   Ligands ⓘ

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
2	CAQ	A	601	8/8	0.77	0.20	49,53,54,56	0

6.5 Other polymers ⓘ

There are no such residues in this entry.