



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 02:10 AM GMT

PDB ID : 3VQS
Title : Crystal structure of HCV NS5B RNA polymerase with a novel piperazine inhibitor
Authors : Adachi, T.; Doi, S.; Ando, I.; Sugimoto, K.; Orita, T.; Nomura, A.; Kamada, M.
Deposited on : 2012-03-30
Resolution : 1.90 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

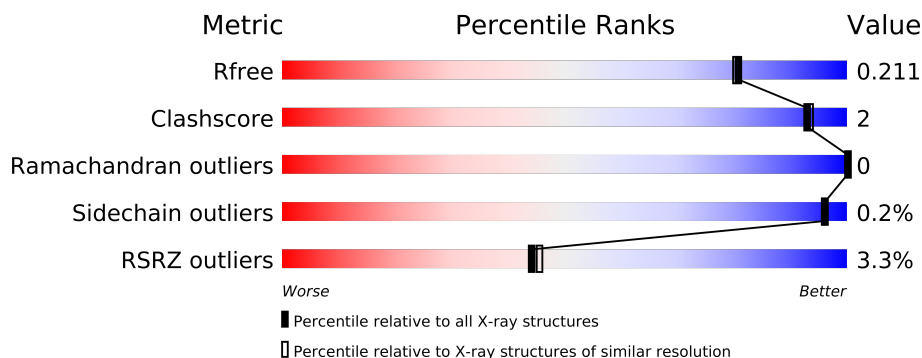
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

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))
R_{free}	66092	3684 (1.90-1.90)
Clashscore	79885	4465 (1.90-1.90)
Ramachandran outliers	78287	4413 (1.90-1.90)
Sidechain outliers	78261	4414 (1.90-1.90)
RSRZ outliers	66119	3686 (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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	578	
1	B	578	
1	C	578	
1	D	578	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 19445 atoms, of which 0 are hydrogen and 0 are deuterium.

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	553	Total	C	N	O	S	0	21	0
			4451	2798	784	835	34			
1	B	557	Total	C	N	O	S	0	14	0
			4425	2786	780	826	33			
1	C	556	Total	C	N	O	S	0	22	0
			4477	2814	789	840	34			
1	D	553	Total	C	N	O	S	0	19	0
			4437	2790	785	827	35			

There are 32 discrepancies between the modelled and reference sequences:

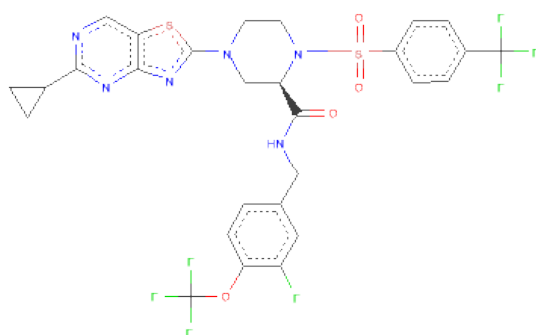
Chain	Residue	Modelled	Actual	Comment	Reference
A	571	GLY	-	EXPRESSION TAG	UNP D0PY27
A	572	SER	-	EXPRESSION TAG	UNP D0PY27
A	573	HIS	-	EXPRESSION TAG	UNP D0PY27
A	574	HIS	-	EXPRESSION TAG	UNP D0PY27
A	575	HIS	-	EXPRESSION TAG	UNP D0PY27
A	576	HIS	-	EXPRESSION TAG	UNP D0PY27
A	577	HIS	-	EXPRESSION TAG	UNP D0PY27
A	578	HIS	-	EXPRESSION TAG	UNP D0PY27
B	571	GLY	-	EXPRESSION TAG	UNP D0PY27
B	572	SER	-	EXPRESSION TAG	UNP D0PY27
B	573	HIS	-	EXPRESSION TAG	UNP D0PY27
B	574	HIS	-	EXPRESSION TAG	UNP D0PY27
B	575	HIS	-	EXPRESSION TAG	UNP D0PY27
B	576	HIS	-	EXPRESSION TAG	UNP D0PY27
B	577	HIS	-	EXPRESSION TAG	UNP D0PY27
B	578	HIS	-	EXPRESSION TAG	UNP D0PY27
C	571	GLY	-	EXPRESSION TAG	UNP D0PY27
C	572	SER	-	EXPRESSION TAG	UNP D0PY27
C	573	HIS	-	EXPRESSION TAG	UNP D0PY27
C	574	HIS	-	EXPRESSION TAG	UNP D0PY27
C	575	HIS	-	EXPRESSION TAG	UNP D0PY27

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	576	HIS	-	EXPRESSION TAG	UNP D0PY27
C	577	HIS	-	EXPRESSION TAG	UNP D0PY27
C	578	HIS	-	EXPRESSION TAG	UNP D0PY27
D	571	GLY	-	EXPRESSION TAG	UNP D0PY27
D	572	SER	-	EXPRESSION TAG	UNP D0PY27
D	573	HIS	-	EXPRESSION TAG	UNP D0PY27
D	574	HIS	-	EXPRESSION TAG	UNP D0PY27
D	575	HIS	-	EXPRESSION TAG	UNP D0PY27
D	576	HIS	-	EXPRESSION TAG	UNP D0PY27
D	577	HIS	-	EXPRESSION TAG	UNP D0PY27
D	578	HIS	-	EXPRESSION TAG	UNP D0PY27

- Molecule 2 is (2R)-4-(5-CYCLOPROPYL[1,3]THIAZOLO[4,5-D]PYRIMIDIN-2-YL)-N-[3-FLUORO-4-(TRIFLUOROMETHOXY)BENZYL]-1-[4-(TRIFLUOROMETHYL)PHE NYL]SULFONYL}PIPERAZINE-2-CARBOXAMIDE (three-letter code: JT1) (formula: C₂₈H₂₃F₇N₆O₄S₂).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	S	0	0
			47	28	7	6	4	2		
2	B	1	Total	C	F	N	O	S	0	0
			47	28	7	6	4	2		
2	C	1	Total	C	F	N	O	S	0	0
			47	28	7	6	4	2		
2	D	1	Total	C	F	N	O	S	0	0
			47	28	7	6	4	2		

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	Cl 1	0	0
3	C	1	Total 1	Cl 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	367	Total 367	O 367	0	0
4	B	374	Total 374	O 374	0	0
4	C	336	Total 336	O 336	0	0
4	D	388	Total 388	O 388	0	0

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	101.12Å 101.18Å 250.06Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	93.97 – 1.90 93.79 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (93.97-1.90) 100.0 (93.79-1.90)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.08 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.7.0017	Depositor
R, R_{free}	0.170 , 0.204 0.177 , 0.211	Depositor DCC
R_{free} test set	10170 reflections (5.31%)	DCC
Wilson B-factor (Å ²)	31.1	Xtriage
Anisotropy	0.330	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 47.4	EDS
Estimated twinning fraction	0.065 for k,h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	1 of 201846 reflections (0.000%)	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	19445	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 14.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: JT1, CL

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	1.05	9/4544 (0.2%)	0.86	2/6168 (0.0%)
1	B	1.03	5/4519 (0.1%)	0.85	2/6134 (0.0%)
1	C	1.02	2/4571 (0.0%)	0.85	1/6207 (0.0%)
1	D	1.07	4/4531 (0.1%)	0.89	3/6149 (0.0%)
All	All	1.04	20/18165 (0.1%)	0.86	8/24658 (0.0%)

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	6	TRP	CD2-CE2	5.83	1.48	1.41
1	A	408	TRP	CG-CD1	5.76	1.44	1.36
1	B	550	TRP	CD2-CE2	5.63	1.48	1.41
1	A	500	TRP	CD2-CE2	5.51	1.48	1.41
1	A	408	TRP	CD2-CE2	5.48	1.48	1.41

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	280	ARG	NE-CZ-NH2	-5.87	117.36	120.30
1	D	345	ARG	NE-CZ-NH1	5.83	123.21	120.30
1	B	277	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	C	517	ARG	NE-CZ-NH2	-5.62	117.49	120.30
1	B	318	ASP	CB-CG-OD1	5.28	123.05	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	4451	0	0	9	0
1	B	4425	0	0	6	0
1	C	4477	0	0	4	0
1	D	4437	0	0	7	0
2	A	47	0	22	0	0
2	B	47	0	22	0	0
2	C	47	0	22	0	0
2	D	47	0	22	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	367	0	0	2	0
4	B	374	0	0	0	0
4	C	336	0	0	1	0
4	D	388	0	0	0	0
All	All	19445	0	88	26	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

The worst 5 of 26 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:303[B]:CYS:SG	1:A:313:MET:CE	2.43	1.06
1:D:303[B]:CYS:SG	1:D:313:MET:CE	2.75	0.75
1:C:236[A]:GLU:OE2	1:C:280:ARG:NH2	2.21	0.72
1:C:303[B]:CYS:SG	1:C:313:MET:CE	2.80	0.70
1:A:236[A]:GLU:OE2	1:A:280:ARG:NH2	2.26	0.68

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	572/578 (99%)	565 (99%)	7 (1%)	0	100	100
1	B	569/578 (98%)	563 (99%)	6 (1%)	0	100	100
1	C	576/578 (100%)	568 (99%)	8 (1%)	0	100	100
1	D	570/578 (99%)	566 (99%)	4 (1%)	0	100	100
All	All	2287/2312 (99%)	2262 (99%)	25 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	493/492 (100%)	493 (100%)	0	100	100
1	B	488/492 (99%)	486 (100%)	2 (0%)	95	95
1	C	496/492 (101%)	495 (100%)	1 (0%)	96	96
1	D	491/492 (100%)	490 (100%)	1 (0%)	96	96
All	All	1968/1968 (100%)	1964 (100%)	4 (0%)	96	96

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	14	CYS
1	B	56	ARG
1	C	56	ARG
1	D	56	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 6 ligands modelled in this entry, 2 are monoatomic - leaving 4 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
2	JT1	A	1000	-	52,52,52	1.95	11 (21%)	78,80,80	21.45	27 (34%)
2	JT1	B	601	-	52,52,52	2.23	19 (36%)	78,80,80	14.75	21 (26%)
2	JT1	C	601	-	52,52,52	1.96	16 (30%)	78,80,80	21.14	20 (25%)
2	JT1	D	1000	-	52,52,52	1.89	11 (21%)	78,80,80	18.83	21 (26%)

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	JT1	A	1000	-	-	0/38/55/55	0/3/6/6
2	JT1	B	601	-	-	0/38/55/55	0/3/6/6

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	JT1	C	601	-	-	0/38/55/55	0/3/6/6
2	JT1	D	1000	-	-	0/38/55/55	0/3/6/6

The worst 5 of 57 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1000	JT1	C30-S27	-7.03	1.67	1.74
2	C	601	JT1	C4-S3	-6.07	1.67	1.76
2	A	1000	JT1	C30-S27	-5.85	1.68	1.74
2	B	601	JT1	C30-S27	-5.52	1.68	1.74
2	B	601	JT1	O5-S3	-5.24	1.37	1.43

The worst 5 of 89 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1000	JT1	N28-C29-N34	-188.03	120.54	128.02
2	C	601	JT1	N28-C29-N34	-185.67	120.63	128.02
2	D	1000	JT1	N28-C29-N34	-165.06	121.45	128.02
2	B	601	JT1	N28-C29-N34	-128.64	122.90	128.02
2	A	1000	JT1	C26-N28-C29	-13.82	107.68	118.71

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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(Å ²)	Q<0.9
1	A	553/578 (95%)	-0.00	17 (3%) 47 48	20, 35, 72, 123	0
1	B	557/578 (96%)	-0.00	13 (2%) 57 59	22, 37, 70, 122	0
1	C	556/578 (96%)	0.07	22 (3%) 36 37	21, 37, 71, 128	0
1	D	553/578 (95%)	-0.01	20 (3%) 41 42	21, 33, 69, 123	0
All	All	2219/2312 (95%)	0.01	72 (3%) 44 46	20, 36, 71, 128	0

The worst 5 of 72 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	23	ILE	7.6
1	D	544	GLN	7.2
1	D	23	ILE	7.1
1	B	544	GLN	6.8
1	D	152	GLY	5.9

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 carbohydrates 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. 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	JT1	D	1000	47/47	0.10	-0.12	22,25,29,31	0
2	JT1	C	601	47/47	0.11	-0.21	24,30,37,40	0
2	JT1	B	601	47/47	0.10	-0.26	23,28,43,52	0
2	JT1	A	1000	47/47	0.09	-0.51	21,25,31,34	0
3	CL	B	602	1/1	0.08	-1.27	43,43,43,43	0
3	CL	C	602	1/1	0.07	-2.07	47,47,47,47	0

6.5 Other polymers ⓘ

There are no such residues in this entry.