Full wwPDB X-ray Structure Validation Report (i)

Oct 31, 2023 – 01:45 PM JST

PDB ID	:	5F3Z
Title	:	Dengue serotype 3 RNA-dependent RNA polymerase bound to PC-79-SH52
Authors	:	Noble, C.G.
Deposited on	:	2015-12-03
Resolution	:	2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) 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 (1)) were used in the production of this report:

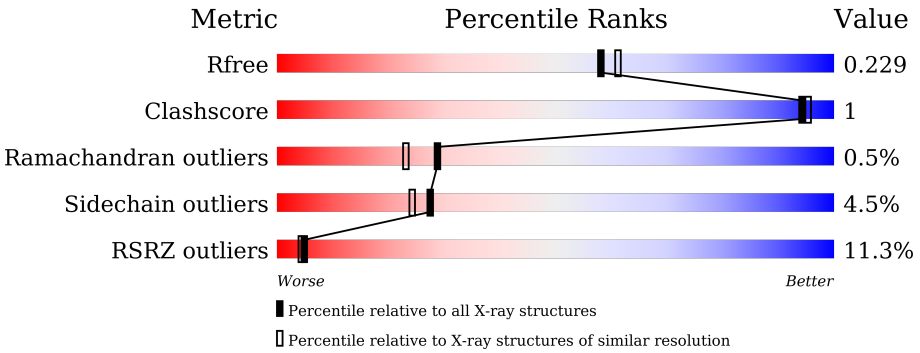
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 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}	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (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 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	635	10% 85% 7% 8%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 5236 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 Genome polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	585	4780	3020	856	873	31	0	3	0

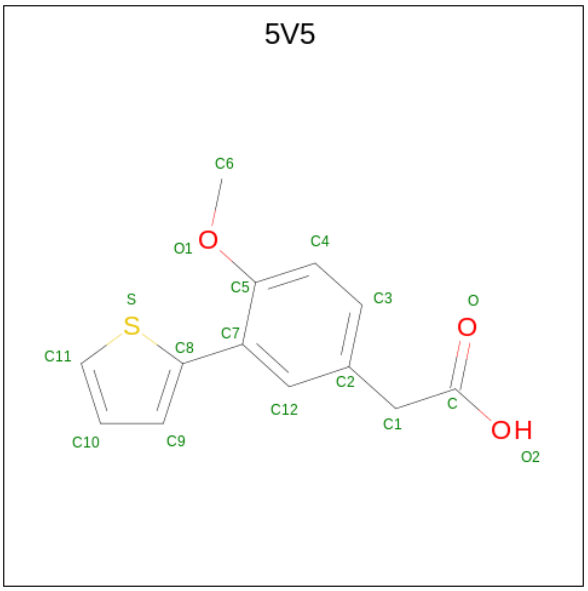
There are 7 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	266	GLY	-	expression tag	UNP Q6DLV0
A	267	SER	-	expression tag	UNP Q6DLV0
A	268	HIS	-	expression tag	UNP Q6DLV0
A	269	MET	-	expression tag	UNP Q6DLV0
A	270	LEU	-	expression tag	UNP Q6DLV0
A	271	ASP	-	expression tag	UNP Q6DLV0
A	374	GLU	GLY	variant	UNP Q6DLV0

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
2	A	2	2	2	0	0

- Molecule 3 is 2-(4-methoxy-3-thiophen-2-yl-phenyl)ethanoic acid (three-letter code: 5V5) (formula: $C_{13}H_{12}O_3S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	S	0	0
			17	13	3	1		

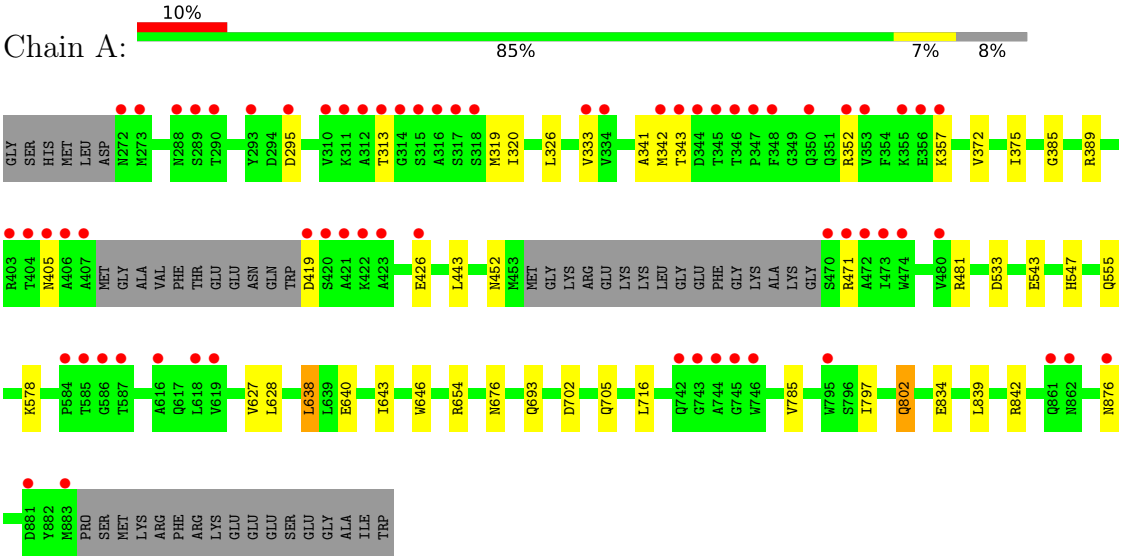
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	437	Total	O	0	0
			437	437		

3 Residue-property plots (i)

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: Genome polyprotein



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	161.30Å 176.22Å 57.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.28 – 2.00 28.28 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.6 (29.28-2.00) 98.6 (28.28-2.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 2.00Å)	Xtriage
Refinement program	BUSTER-TNT BUSTER 2.11.5	Depositor
R, R_{free}	0.184 , 0.219 0.191 , 0.229	Depositor DCC
R_{free} test set	2813 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	29.0	Xtriage
Anisotropy	0.628	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 60.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5236	wwPDB-VP
Average B, all atoms (Å^2)	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.69% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 5V5

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	0.50	0/4910	0.63	0/6650

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4780	0	4678	12	0
2	A	2	0	0	0	0
3	A	17	0	0	0	0
4	A	437	0	0	0	0
All	All	5236	0	4678	12	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 1.

All (12) 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:802:GLN:H	1:A:802:GLN:HE21	1.26	0.80
1:A:372:VAL:HG11	1:A:628:LEU:HD11	1.83	0.60
1:A:385:GLY:HA3	1:A:555:GLN:HE22	1.68	0.58
1:A:716:LEU:HD21	1:A:839:LEU:HD23	1.86	0.57
1:A:320:ILE:HD11	1:A:341:ALA:HB1	1.90	0.53
1:A:452:ASN:HD21	1:A:578:LYS:HD2	1.78	0.48
1:A:543:GLU:O	1:A:547[A]:HIS:HD2	1.97	0.47
1:A:375:ILE:HD11	1:A:640:GLU:HG2	1.99	0.45
1:A:646:TRP:CZ2	1:A:654:ARG:HG3	2.52	0.45
1:A:702:ASP:HB3	1:A:705:GLN:HG2	2.03	0.41
1:A:638:LEU:HD11	1:A:643:ILE:HD11	2.02	0.41
1:A:627:VAL:HG21	1:A:646:TRP:CD1	2.56	0.40

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	582/635 (92%)	565 (97%)	14 (2%)	3 (0%)	29 23

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	343	THR
1	A	342	MET
1	A	333	VAL

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	513/552 (93%)	490 (96%)	23 (4%)	27 24

All (23) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	295	ASP
1	A	313	THR
1	A	319	MET
1	A	326	LEU
1	A	352	ARG
1	A	357	LYS
1	A	389	ARG
1	A	405	ASN
1	A	419	ASP
1	A	426	GLU
1	A	443	LEU
1	A	471	ARG
1	A	481	ARG
1	A	533	ASP
1	A	638	LEU
1	A	676	ASN
1	A	693	GLN
1	A	785	VAL
1	A	797	ILE
1	A	802	GLN
1	A	834	GLU
1	A	842	ARG
1	A	876	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	297	ASN
1	A	339	GLN
1	A	452	ASN
1	A	548	ASN
1	A	555	GLN
1	A	621	GLN
1	A	682	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	704	GLN
1	A	768	HIS
1	A	802	GLN
1	A	835	ASN

5.3.3 RNA

There are no RNA molecules 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 monosaccharides in this entry.

5.6 Ligand geometry

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	5V5	A	1003	-	18,18,18	1.26	1 (5%)	20,24,24	2.31	3 (15%)

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	5V5	A	1003	-	-	2/9/10/10	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1003	5V5	C7-C5	4.30	1.49	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1003	5V5	C10-C11-S	-8.60	106.00	112.98
3	A	1003	5V5	C6-O1-C5	2.84	121.82	117.53
3	A	1003	5V5	O1-C5-C7	2.54	120.05	116.26

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1003	5V5	O2-C-C1-C2
3	A	1003	5V5	O-C-C1-C2

There are no ring outliers.

No monomer is involved in short contacts.

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

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, 95^{th} 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	585/635 (92%)	0.37	66 (11%)	5 4	19, 35, 77, 130	0

All (66) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	407	ALA	17.8
1	A	353	VAL	9.5
1	A	406	ALA	6.7
1	A	343	THR	6.6
1	A	745	GLY	6.4
1	A	743	GLY	6.4
1	A	348	PHE	5.9
1	A	419	ASP	5.8
1	A	316	ALA	5.7
1	A	344	ASP	5.4
1	A	346	THR	5.4
1	A	312	ALA	5.3
1	A	744	ALA	5.3
1	A	345	THR	5.0
1	A	342	MET	5.0
1	A	352	ARG	4.6
1	A	290	THR	4.5
1	A	584	PRO	4.5
1	A	746	TRP	4.3
1	A	315	SER	4.3
1	A	317	SER	4.3
1	A	272	ASN	3.9
1	A	314	GLY	3.8
1	A	313	THR	3.7
1	A	586	GLY	3.7
1	A	318	SER	3.5
1	A	423	ALA	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	404	THR	3.4
1	A	742	GLN	3.4
1	A	405	ASN	3.3
1	A	470	SER	3.3
1	A	420	SER	3.2
1	A	585	THR	3.1
1	A	295	ASP	2.9
1	A	333	VAL	2.9
1	A	289	SER	2.8
1	A	293	TYR	2.8
1	A	288	ASN	2.8
1	A	310	VAL	2.8
1	A	355	LYS	2.8
1	A	347	PRO	2.7
1	A	421	ALA	2.7
1	A	795	TRP	2.6
1	A	472	ALA	2.5
1	A	471	ARG	2.5
1	A	311	LYS	2.5
1	A	480	VAL	2.5
1	A	356	GLU	2.4
1	A	876	ASN	2.4
1	A	357	LYS	2.4
1	A	862	ASN	2.4
1	A	619	VAL	2.3
1	A	422	LYS	2.3
1	A	334	VAL	2.3
1	A	426	GLU	2.2
1	A	350	GLN	2.2
1	A	273	MET	2.2
1	A	473	ILE	2.2
1	A	883	MET	2.2
1	A	881	ASP	2.2
1	A	474	TRP	2.1
1	A	403	ARG	2.1
1	A	587	THR	2.1
1	A	861	GLN	2.1
1	A	618	LEU	2.1
1	A	616	ALA	2.0

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 monosaccharides 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. The B-factors column lists the minimum, median, 95^{th} 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
3	5V5	A	1003	17/17	0.96	0.12	25,30,38,42	0
2	ZN	A	1002	1/1	1.00	0.07	37,37,37,37	0
2	ZN	A	1001	1/1	1.00	0.08	26,26,26,26	0

6.5 Other polymers (i)

There are no such residues in this entry.