wwPDB X-ray Structure Validation Summary Report (i)

Jul 13, 2016 – 11:28 AM EDT

PDB ID	:	5JXS
Title	:	Mutant GC216/7AA of 3D polymerase from Foot-and-Mouth Disease Virus
Authors	:	Verdaguer, N.; Ferrer-Orta, C.
Deposited on	:	2016-05-13
Resolution	:	2.80 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 (1)) were used in the production of this report:

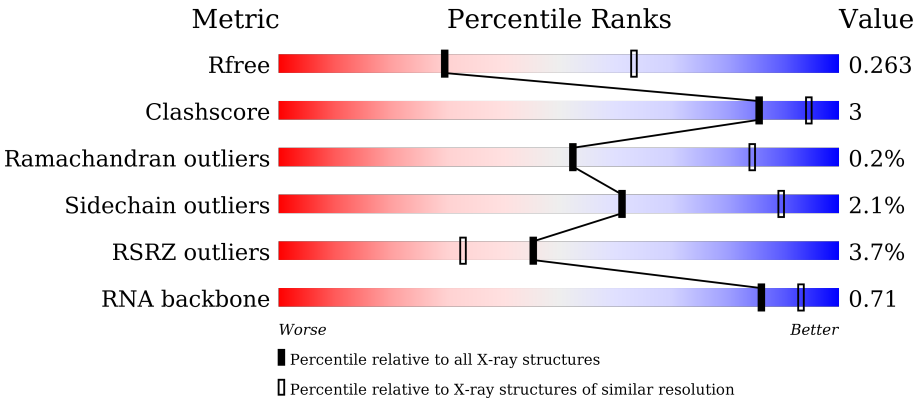
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027790
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027790

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 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)
RNA backbone	2183	1091 (3.20-2.40)

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.

Mol	Chain	Length	Quality of chain
1	A	481	2%, 91%, 8%
2	B	8	50%, 38%, 63%
3	C	6	33%, 100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MG	A	501	-	-	-	X

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4038 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			
1	A	476	3728	2372	644	692	20	0	1	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	216	ALA	GLY	engineered mutation	UNP P03311
A	217	ALA	CYS	engineered mutation	UNP P03311
A	471	ALA	-	expression tag	UNP P03311
A	472	ALA	-	expression tag	UNP P03311
A	473	ALA	-	expression tag	UNP P03311
A	474	LEU	-	expression tag	UNP P03311
A	475	GLU	-	expression tag	UNP P03311
A	476	HIS	-	expression tag	UNP P03311
A	477	HIS	-	expression tag	UNP P03311
A	478	HIS	-	expression tag	UNP P03311
A	479	HIS	-	expression tag	UNP P03311
A	480	HIS	-	expression tag	UNP P03311
A	481	HIS	-	expression tag	UNP P03311

- Molecule 2 is a RNA chain called RNA Template.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	B	8	171	76	31	56	8	0	0	0

- Molecule 3 is a RNA chain called RNA Primer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	6	129	57	24	42	6	0	0	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Mg 1 1	0	0

- Molecule 5 is water.

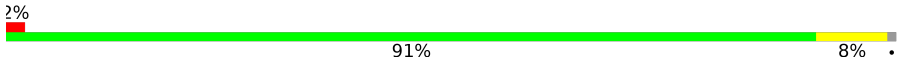
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	8	Total O 8 8	0	0
5	B	1	Total O 1 1	0	0

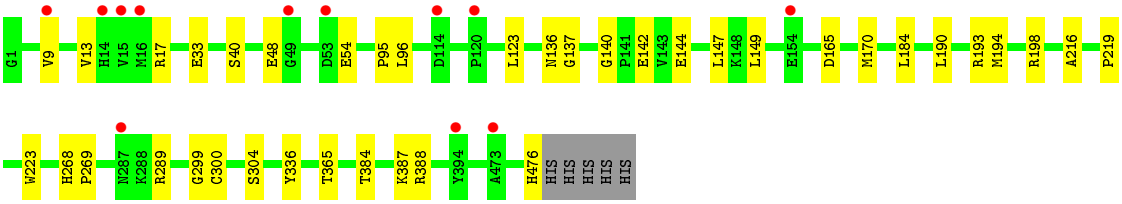
3 Residue-property plots (i)

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.

- Molecule 1: RNA dependent RNA polymerase

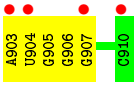
Chain A:





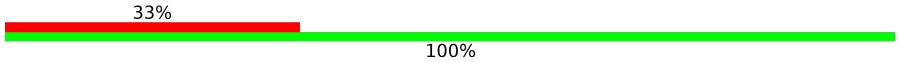
- Molecule 2: RNA Template

Chain B:



- Molecule 3: RNA Primer

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	95.78Å 95.78Å 100.88Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.10 – 2.80 43.10 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.0 (43.10-2.80) 95.0 (43.10-2.80)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.42 (at 2.81Å)	Xtriage
Refinement program	PHENIX (dev_2341: ???)	Depositor
R, R_{free}	0.205 , 0.264 0.203 , 0.263	Depositor DCC
R_{free} test set	672 reflections (5.21%)	DCC
Wilson B-factor (Å^2)	65.4	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 49.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.031 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4038	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.68% 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: MG

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	0.24	0/3818	0.41	0/5179
2	B	0.25	0/190	0.82	0/294
3	C	0.12	0/143	0.66	0/221
All	All	0.24	0/4151	0.45	0/5694

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 (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	3728	0	3620	19	0
2	B	171	0	88	7	0
3	C	129	0	67	0	0
4	A	1	0	0	0	0
5	A	8	0	0	0	0
5	B	1	0	0	0	0
All	All	4038	0	3775	22	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 22 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)
2:B:904:U:H2'	2:B:905:G:C8	2.36	0.60
1:A:96:LEU:HD13	1:A:198:ARG:HG2	1.85	0.58
2:B:904:U:H2'	2:B:905:G:H8	1.70	0.57
1:A:223:TRP:CH2	1:A:388:ARG:HB2	2.42	0.54
1:A:96:LEU:HB3	1:A:194:MET:HA	1.91	0.52

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	475/481 (99%)	457 (96%)	17 (4%)	1 (0%)	52 84

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	299	GLY

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	389/402 (97%)	381 (98%)	8 (2%)	61 90

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	336	TYR
1	A	476	HIS
1	A	384	THR
1	A	149	LEU
1	A	365	THR

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

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	B	7/8 (87%)	0	0
3	C	5/6 (83%)	0	0
All	All	12/14 (85%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

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 (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)

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	476/481 (98%)	0.02	12 (2%)	61	48	38, 70, 122, 153	0
2	B	8/8 (100%)	2.49	4 (50%)	0	0	120, 124, 169, 186	1 (12%)
3	C	6/6 (100%)	1.46	2 (33%)	0	0	93, 101, 123, 143	1 (16%)
All	All	490/495 (98%)	0.08	18 (3%)	45	33	38, 71, 124, 186	2 (0%)

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	903	A	6.2
1	A	287	ASN	4.5
1	A	53	ASP	3.6
2	B	910	C	3.6
3	C	915	G	3.4

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(Å^2)	Q<0.9
4	MG	A	501	1/1	0.86	0.30	5.39	61,61,61,61	0

6.5 Other polymers (i)

There are no such residues in this entry.