Full wwPDB X-ray Structure Validation Report (i)

Jul 2, 2018 – 12:02 PM EDT

PDB ID	:	6DX2
Title	:	Crystal structure of the viral OTU domain protease from Dera Ghazi Khan virus
Deposited on	:	2018-06-28
Resolution	:	1.61 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report.

This report is produced by the wwPDB biocuration pipeline after annotation of the structure.

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 following versions of software and data (see references (1)) were used in the production of this report:

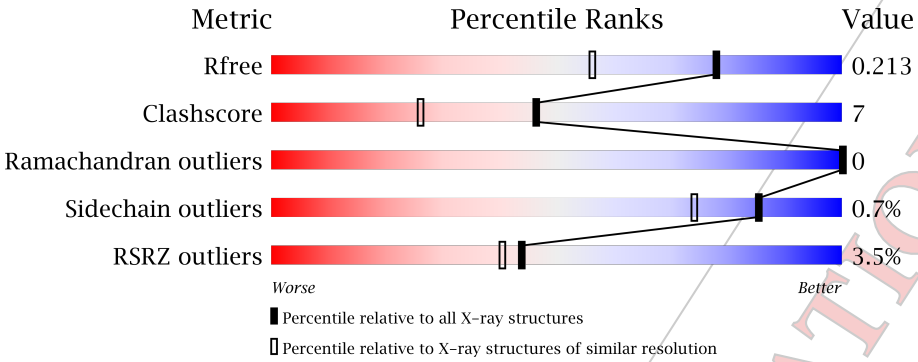
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	rb-20031172
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20031172

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 1.61 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}	111664	3975 (1.64-1.60)
Clashscore	122126	4258 (1.64-1.60)
Ramachandran outliers	120053	4162 (1.64-1.60)
Sidechain outliers	120020	4161 (1.64-1.60)
RSRZ outliers	108989	3894 (1.64-1.60)

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	176	3%, 70%, 18%, ·, 11%
1	B	176	3%, 80%, 9%, 11%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2675 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
1	A	156	Total	C	N	O	S	0	0	0
			1246	798	204	239	5			
1	B	156	Total	C	N	O	S	0	0	0
			1246	798	204	239	5			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	170	SER	-	expression tag	UNP A0A191KW82
A	171	HIS	-	expression tag	UNP A0A191KW82
A	172	HIS	-	expression tag	UNP A0A191KW82
A	173	HIS	-	expression tag	UNP A0A191KW82
A	174	HIS	-	expression tag	UNP A0A191KW82
A	175	HIS	-	expression tag	UNP A0A191KW82
A	176	HIS	-	expression tag	UNP A0A191KW82
B	170	SER	-	expression tag	UNP A0A191KW82
B	171	HIS	-	expression tag	UNP A0A191KW82
B	172	HIS	-	expression tag	UNP A0A191KW82
B	173	HIS	-	expression tag	UNP A0A191KW82
B	174	HIS	-	expression tag	UNP A0A191KW82
B	175	HIS	-	expression tag	UNP A0A191KW82
B	176	HIS	-	expression tag	UNP A0A191KW82

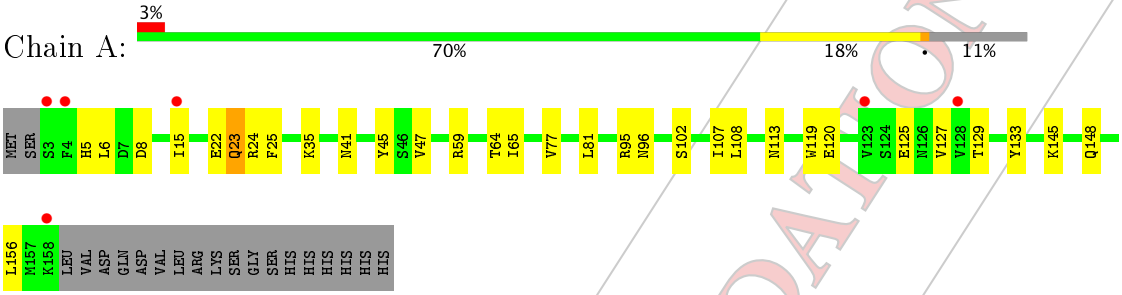
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	96	Total	O	0	0
			96	96		
2	B	87	Total	O	0	0
			87	87		

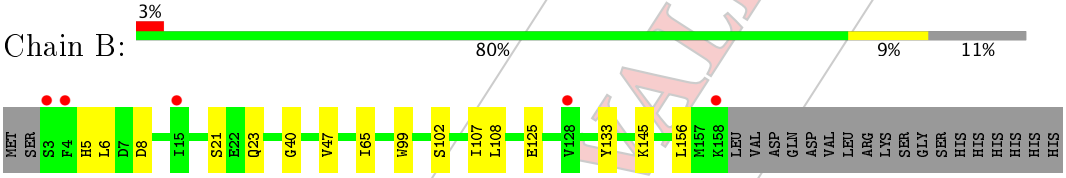
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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 geometryand 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



- Molecule 1: RNA-dependent RNA polymerase



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α, β, γ	68.42Å 68.42Å 69.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	28.02 – 1.61 28.02 – 1.61	Depositor EDS
% Data completeness (in resolution range)	98.6 (28.02-1.61) 92.5 (28.02-1.61)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.39 (at 1.62Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.179 , 0.213 0.179 , 0.213	Depositor DCC
R_{free} test set	2008 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	30.3	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 50.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.018 for -h,-l,-k 0.015 for -h,l,k 0.019 for l,-k,h 0.022 for -l,-k,-h 0.487 for h,-k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	2675	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.18% 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

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	0.68	0/1277	0.73	0/1736
1	B	0.70	0/1277	0.74	0/1736
All	All	0.69	0/2554	0.74	0/3472

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	1246	0	1191	23	0
1	B	1246	0	1191	10	0
2	A	96	0	0	4	2
2	B	87	0	0	0	2
All	All	2675	0	2382	33	3

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.

All (33) 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:95:ARG:NH1	2:A:201:HOH:O	1.99	0.94

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:148:GLN:NE2	2:A:202:HOH:O	2.15	0.78
1:A:6:LEU:HD12	1:A:156:LEU:HD21	1.70	0.72
1:B:47:VAL:HG11	1:B:108:LEU:HD21	1.75	0.68
1:B:6:LEU:HD12	1:B:156:LEU:HD21	1.79	0.65
1:A:47:VAL:HG11	1:A:108:LEU:HD21	1.81	0.63
1:B:21:SER:OG	1:B:23:GLN:HG2	2.02	0.59
1:A:23:GLN:HE21	1:A:25:PHE:HE1	1.53	0.55
1:A:23:GLN:NE2	1:A:119:TRP:HZ3	2.07	0.52
1:A:22:GLU:HA	1:A:24:ARG:NH2	2.26	0.50
1:A:113:ASN:ND2	2:A:209:HOH:O	2.48	0.45
1:A:15:ILE:HA	1:A:15:ILE:HD12	1.85	0.45
1:A:5:HIS:HB2	1:A:8:ASP:OD2	2.17	0.45
1:A:77:VAL:O	1:A:81:LEU:CD2	2.65	0.44
1:A:22:GLU:O	1:A:24:ARG:HG2	2.18	0.44
1:B:125:GLU:OE1	1:B:145:LYS:HD2	2.16	0.44
1:B:5:HIS:HB2	1:B:8:ASP:OD2	2.17	0.44
1:A:47:VAL:HG11	1:A:108:LEU:CD2	2.47	0.44
1:A:41:ASN:HB2	1:A:45:TYR:CZ	2.53	0.44
1:A:102:SER:HB2	1:A:133:TYR:OH	2.18	0.43
1:B:102:SER:HB2	1:B:133:TYR:OH	2.17	0.43
1:A:64:THR:HA	2:A:274:HOH:O	2.19	0.43
1:B:156:LEU:HD23	1:B:156:LEU:HA	1.68	0.43
1:B:65:ILE:HG23	1:B:107:ILE:HG21	2.02	0.42
1:B:40:GLY:HA2	1:B:99:TRP:CE2	2.55	0.41
1:A:23:GLN:HE22	1:A:127:VAL:HG11	1.85	0.41
1:A:59:ARG:NH2	1:A:96:ASN:OD1	2.53	0.41
1:A:65:ILE:HG23	1:A:107:ILE:HG21	2.03	0.41
1:A:6:LEU:HA	1:A:6:LEU:HD23	1.74	0.41
1:A:125:GLU:OE1	1:A:145:LYS:HD2	2.21	0.41
1:B:47:VAL:HG11	1:B:108:LEU:CD2	2.47	0.40
1:A:23:GLN:O	1:A:24:ARG:HD3	2.20	0.40
1:A:120:GLU:HB2	1:A:129:THR:OG1	2.21	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:232:HOH:O	2:A:282:HOH:O[3_555]	2.10	0.10
2:B:220:HOH:O	2:B:264:HOH:O[4_544]	2.13	0.07
2:A:228:HOH:O	2:B:214:HOH:O[4_554]	2.18	0.02

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	154/176 (88%)	153 (99%)	1 (1%)	0	100	100
1	B	154/176 (88%)	154 (100%)	0	0	100	100
All	All	308/352 (88%)	307 (100%)	1 (0%)	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	135/154 (88%)	133 (98%)	2 (2%)	67	45
1	B	135/154 (88%)	135 (100%)	0	100	100
All	All	270/308 (88%)	268 (99%)	2 (1%)	85	74

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	GLN
1	A	35	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	23	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

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, 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	156/176 (88%)	-0.11	6 (3%) 40 37	29, 42, 76, 110	0
1	B	156/176 (88%)	-0.10	5 (3%) 47 44	29, 42, 72, 125	0
All	All	312/352 (88%)	-0.10	11 (3%) 44 40	29, 42, 76, 125	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	158	LYS	5.8
1	B	4	PHE	5.5
1	A	3	SER	5.0
1	B	3	SER	4.0
1	A	4	PHE	3.7
1	B	158	LYS	3.0
1	B	15	ILE	2.9
1	A	15	ILE	2.8
1	B	128	VAL	2.6
1	A	128	VAL	2.2
1	A	123	VAL	2.2

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.