Preliminary Full wwPDB X-ray Structure Validation Report (i)

Feb 2, 2020 – 12:06 PM JST

Deposition ID : D_1300015506

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

This report is produced by the wwPDB Deposition System during initial deposition but before 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 (i)) were used in the production of this report:

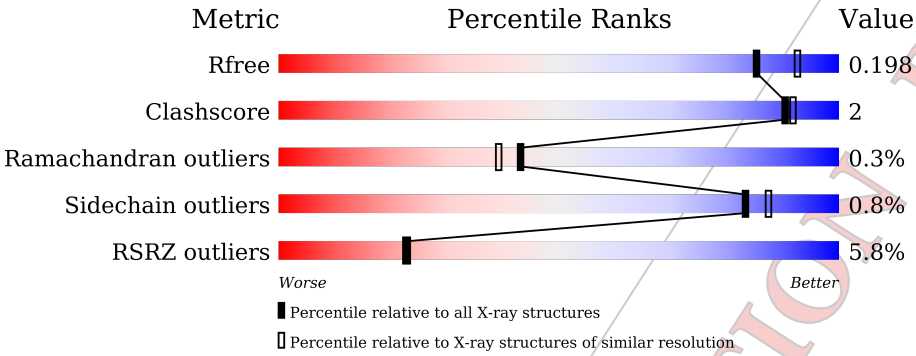
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.8
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.8

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}	111664	7193 (2.00-2.00)
Clashscore	122126	8267 (2.00-2.00)
Ramachandran outliers	120053	8166 (2.00-2.00)
Sidechain outliers	120020	8165 (2.00-2.00)
RSRZ outliers	108989	7011 (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 on 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	335	4% 92% · 5%
1	B	335	7% 88% 7% 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5683 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 multiple peptide resistance factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2478	1590	428	448	12			
1	B	319	Total	C	N	O	S	0	0	0
			2478	1590	428	448	12			

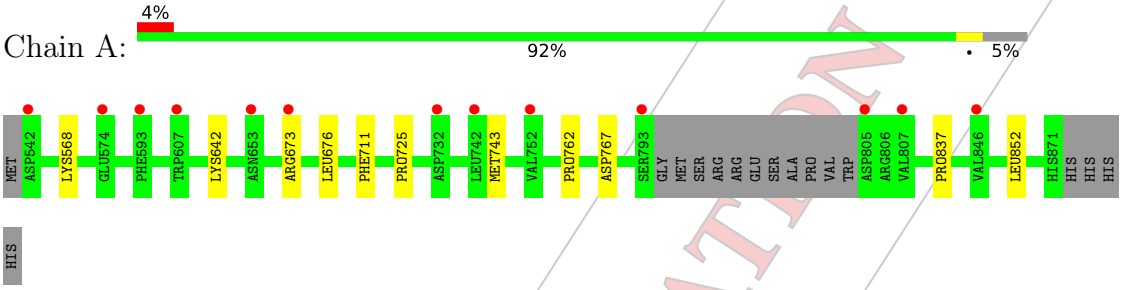
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	S	727	Total	O	0	0
			727	727		

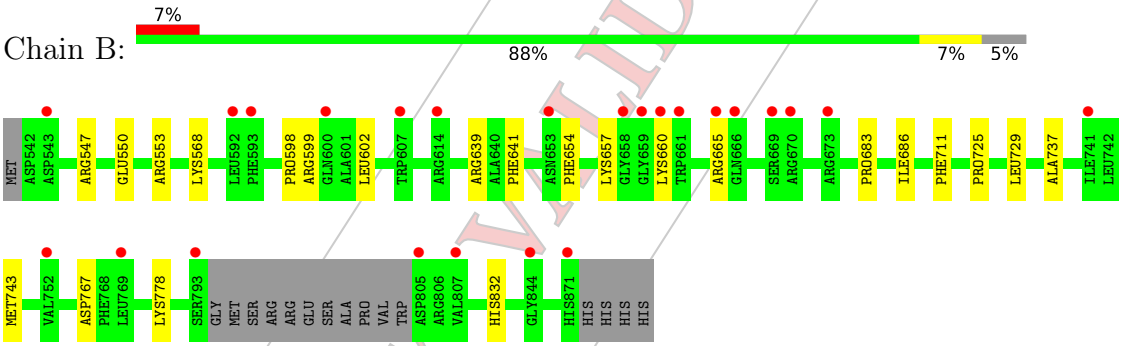
3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: multiple peptide resistance factor



• Molecule 1: multiple peptide resistance factor



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	83.59Å 95.79Å 211.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.21 – 2.00 35.21 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (35.21-2.00) 99.8 (35.21-2.00)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.15 (at 2.00Å)	Xtriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, R_{free}	0.168 , 0.198 0.168 , 0.198	Depositor DCC
R_{free} test set	2921 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	24.2	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 49.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5683	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% 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	0.40	0/2535	0.54	0/3421
1	B	0.40	0/2535	0.55	0/3421
All	All	0.40	0/5070	0.55	0/6842

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	2478	0	2458	6	0
1	B	2478	0	2458	13	0
2	S	727	0	0	0	0
All	All	5683	0	4916	19	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 2.

All (19) 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:673:ARG:HG2	1:A:673:ARG:HH11	1.59	0.67

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:550:GLU:OE2	1:B:553:ARG:NH1	2.29	0.66
1:A:642:LYS:HE3	1:A:837:PRO:HB3	1.82	0.62
1:B:654:PHE:O	1:B:778:LYS:NZ	2.37	0.58
1:A:673:ARG:HG2	1:A:673:ARG:NH1	2.26	0.50
1:B:660:LYS:HE2	1:B:832:HIS:CD2	2.47	0.50
1:B:547:ARG:HH11	1:B:547:ARG:HG3	1.78	0.49
1:B:598:PRO:O	1:B:602:LEU:HD13	2.13	0.48
1:B:725:PRO:HG2	1:B:743:MET:CE	2.44	0.47
1:B:657:LYS:HB3	1:B:665:ARG:CZ	2.45	0.47
1:A:725:PRO:HG2	1:A:743:MET:CE	2.46	0.46
1:A:676:LEU:HD11	1:A:762:PRO:HG2	1.97	0.45
1:A:852:LEU:HA	1:A:852:LEU:HD12	1.74	0.45
1:B:683:PRO:HA	1:B:686:ILE:HD12	1.98	0.45
1:B:725:PRO:HG2	1:B:743:MET:HE1	1.98	0.45
1:B:639:ARG:HG2	1:B:641:PHE:CE1	2.52	0.44
1:B:729:LEU:HB2	1:B:737:ALA:HB3	2.03	0.41
1:B:599:ARG:HA	1:B:602:LEU:HD22	2.03	0.41
1:B:657:LYS:HB3	1:B:665:ARG:NH1	2.36	0.41

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	315/335 (94%)	309 (98%)	5 (2%)	1 (0%)	43	39
1	B	315/335 (94%)	306 (97%)	8 (2%)	1 (0%)	43	39
All	All	630/670 (94%)	615 (98%)	13 (2%)	2 (0%)	43	39

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	711	PHE
1	B	711	PHE

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	254/268 (95%)	252 (99%)	2 (1%)	83	87
1	B	254/268 (95%)	252 (99%)	2 (1%)	83	87
All	All	508/536 (95%)	504 (99%)	4 (1%)	83	87

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

Mol	Chain	Res	Type
1	A	568	LYS
1	A	767	ASP
1	B	568	LYS
1	B	767	ASP

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	832	HIS

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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	319/335 (95%)	0.14	13 (4%)	37	37	14, 24, 41, 57	0
1	B	319/335 (95%)	0.30	24 (7%)	14	13	13, 24, 47, 58	0
All	All	638/670 (95%)	0.22	37 (5%)	23	23	13, 24, 45, 58	0

All (37) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	844	GLY	4.4
1	A	805	ASP	3.8
1	A	793	SER	3.4
1	A	846	VAL	3.2
1	B	614	ARG	3.1
1	B	871	HIS	3.0
1	B	670	ARG	2.9
1	B	741	ILE	2.9
1	B	666	GLN	2.9
1	B	659	GLY	2.8
1	B	752	VAL	2.8
1	B	653	ASN	2.7
1	B	660	LYS	2.7
1	B	673	ARG	2.7
1	B	805	ASP	2.6
1	A	673	ARG	2.5
1	B	607	TRP	2.4
1	B	793	SER	2.4
1	A	807	VAL	2.4
1	A	574	GLU	2.3
1	B	669	SER	2.3
1	B	807	VAL	2.3
1	A	653	ASN	2.3
1	B	543	ASP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	665	ARG	2.3
1	B	661	TRP	2.3
1	B	592	LEU	2.2
1	A	607	TRP	2.2
1	A	742	LEU	2.2
1	A	542	ASP	2.2
1	B	769	LEU	2.1
1	B	658	GLY	2.1
1	B	600	GLN	2.1
1	B	593	PHE	2.1
1	A	752	VAL	2.1
1	A	593	PHE	2.0
1	A	732	ASP	2.0

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.