Full wwPDB X-ray Structure Validation Report (i)

Dec 17, 2021 – 09:21 am GMT

PDB ID	:	7QJM
Title	:	Crystal structure of an alpha/beta-hydrolase enzyme from Chloroflexus sp. MS-G (202)
Deposited on	:	2021-12-17
Resolution	:	2.19 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:

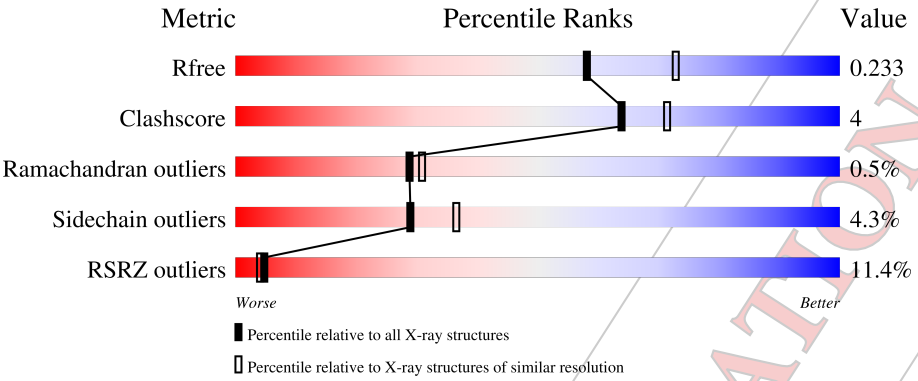
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.24
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.24

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 2.19 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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	388	3% 68% 9% · 22%
1	B	388	15% 64% 12% · 23%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4768 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 alpha/beta-hydrolase (202).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	301	Total	C	N	O	S	0	0	0
			2328	1485	415	421	7			
1	B	297	Total	C	N	O	S	0	0	0
			2301	1467	411	416	7			

- Molecule 2 is water.

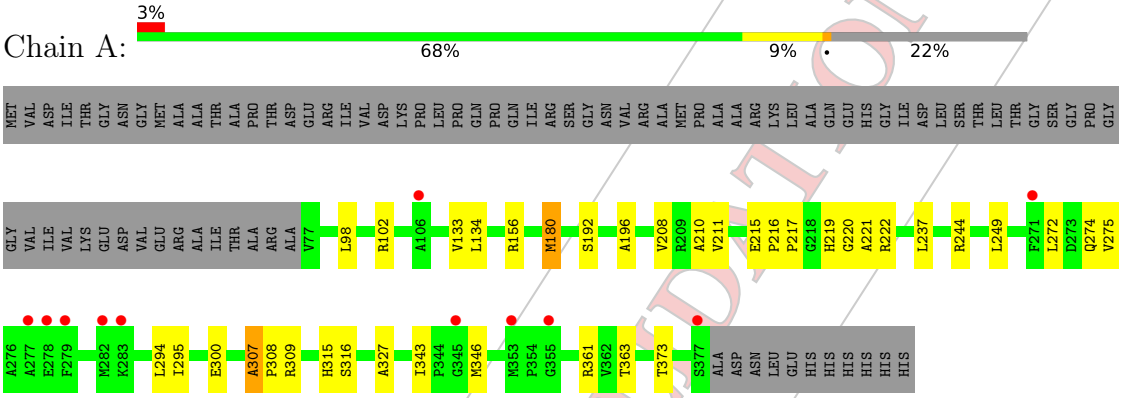
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	80	Total	O	0	0
			80	80		
2	B	59	Total	O	0	0
			59	59		

3 Residue-property plots

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.

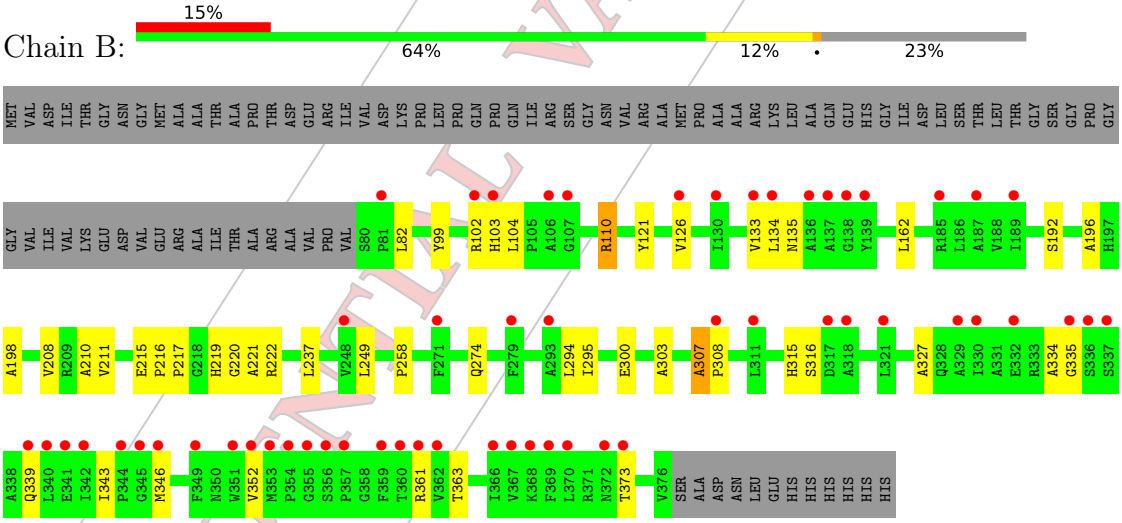
- Molecule 1: alpha/beta-hydrolase (202)

Chain A:



- Molecule 1: alpha/beta-hydrolase (202)

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	90.68Å 124.91Å 169.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.33 – 2.19 50.33 – 2.19	Depositor EDS
% Data completeness (in resolution range)	100.0 (50.33-2.19) 92.1 (50.33-2.19)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.57 (at 2.18Å)	Xtriage
Refinement program	BUSTER 2.10.4	Depositor
R, R_{free}	0.216 , 0.244 0.204 , 0.233	Depositor DCC
R_{free} test set	2537 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	55.1	Xtriage
Anisotropy	0.108	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4768	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.17% 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	0.43	0/2382	0.60	0/3246
1	B	0.42	0/2354	0.61	1/3206 (0.0%)
All	All	0.43	0/4736	0.60	1/6452 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	121	TYR	CA-CB-CG	5.61	124.06	113.40

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	2328	0	2355	18	0
1	B	2301	0	2325	21	0
2	A	80	0	0	0	0
2	B	59	0	0	0	0
All	All	4768	0	4680	39	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 4.

All (39) 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:B:102:ARG:NH1	1:B:103:HIS:NE2	2.42	0.67
1:B:307:ALA:HB1	1:B:308:PRO:HD3	1.78	0.66
1:A:307:ALA:HB1	1:A:308:PRO:HD3	1.79	0.65
1:A:210:ALA:HB2	1:A:373:THR:HG21	1.80	0.62
1:A:307:ALA:CB	1:A:308:PRO:CD	2.79	0.61
1:B:210:ALA:HB2	1:B:373:THR:HG21	1.83	0.60
1:B:307:ALA:CB	1:B:308:PRO:CD	2.79	0.60
1:B:133:VAL:HG11	1:B:363:THR:HB	1.85	0.58
1:A:133:VAL:HG11	1:A:363:THR:HB	1.85	0.58
1:B:315:HIS:ND1	1:B:327:ALA:HB2	2.22	0.55
1:A:315:HIS:ND1	1:A:327:ALA:HB2	2.22	0.55
1:B:222:ARG:NH2	1:B:300:GLU:OE1	2.34	0.53
1:B:220:GLY:HA3	1:B:294:LEU:O	2.11	0.51
1:A:307:ALA:HB1	1:A:308:PRO:CD	2.43	0.49
1:B:221:ALA:HB2	1:B:295:ILE:HG23	1.95	0.48
1:A:220:GLY:HA3	1:A:294:LEU:O	2.13	0.48
1:A:222:ARG:NH2	1:A:300:GLU:OE2	2.40	0.47
1:B:102:ARG:NH1	1:B:103:HIS:CE1	2.82	0.47
1:B:303:ALA:HB1	1:B:334:ALA:HB2	1.97	0.47
1:A:221:ALA:HB2	1:A:295:ILE:HG23	1.97	0.46
1:A:343:ILE:HG23	1:A:361:ARG:HD2	1.99	0.44
1:B:316:SER:HB3	1:B:346:MET:O	2.18	0.44
1:A:307:ALA:CB	1:A:308:PRO:HD3	2.43	0.43
1:B:126:VAL:HG22	1:B:352:VAL:HG13	2.00	0.43
1:B:343:ILE:HG23	1:B:361:ARG:HD2	2.00	0.43
1:B:104:LEU:HD11	1:B:110:ARG:HD3	2.00	0.42
1:B:192:SER:HA	1:B:215:GLU:O	2.20	0.42
1:B:216:PRO:HA	1:B:217:PRO:HD3	1.91	0.42
1:A:98:LEU:HD21	1:A:180:MET:HE3	2.01	0.42
1:A:216:PRO:HA	1:A:217:PRO:HD3	1.93	0.41
1:B:99:TYR:HB3	1:B:135:ASN:HD21	1.84	0.41
1:B:196:ALA:HB2	1:B:217:PRO:HD2	2.01	0.41
1:B:219:HIS:CE1	1:B:221:ALA:HB3	2.55	0.41
1:A:208:VAL:O	1:A:309:ARG:HD2	2.21	0.41
1:B:198:ALA:HB1	1:B:211:VAL:HG23	2.01	0.41
1:A:196:ALA:HB2	1:A:217:PRO:HD2	2.03	0.41
1:A:219:HIS:CE1	1:A:221:ALA:HB3	2.56	0.41
1:A:316:SER:HB3	1:A:346:MET:O	2.21	0.41
1:A:192:SER:HA	1:A:215:GLU:O	2.20	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	299/388 (77%)	291 (97%)	7 (2%)	1 (0%)	41 46
1	B	295/388 (76%)	283 (96%)	10 (3%)	2 (1%)	22 22
All	All	594/776 (76%)	574 (97%)	17 (3%)	3 (0%)	29 31

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	307	ALA
1	B	307	ALA
1	B	335	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	245/313 (78%)	234 (96%)	11 (4%)	27 34
1	B	241/313 (77%)	231 (96%)	10 (4%)	30 39
All	All	486/626 (78%)	465 (96%)	21 (4%)	29 36

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	102	ARG
1	A	134	LEU
1	A	156	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	180	MET
1	A	211	VAL
1	A	237	LEU
1	A	244	ARG
1	A	249	LEU
1	A	272	LEU
1	A	274	GLN
1	A	275	VAL
1	B	82	LEU
1	B	110	ARG
1	B	134	LEU
1	B	162	LEU
1	B	208	VAL
1	B	237	LEU
1	B	249	LEU
1	B	258	PRO
1	B	274	GLN
1	B	339	GLN

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

Mol	Chain	Res	Type
1	B	135	ASN

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 monosaccharides 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	301/388 (77%)	0.42	11 (3%) 41 39	45, 70, 99, 111	0
1	B	297/388 (76%)	1.15	57 (19%) 1 1	49, 83, 134, 142	0
All	All	598/776 (77%)	0.79	68 (11%) 5 4	45, 75, 131, 142	0

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	345	GLY	8.2
1	B	353	MET	7.7
1	B	362	VAL	7.2
1	B	366	ILE	7.0
1	A	345	GLY	6.1
1	B	351	TRP	5.9
1	B	107	GLY	5.8
1	B	335	GLY	4.9
1	B	106	ALA	4.7
1	B	134	LEU	4.6
1	B	308	PRO	4.5
1	B	370	LEU	4.5
1	B	271	PHE	4.4
1	B	340	LEU	4.3
1	B	138	GLY	4.1
1	B	342	ILE	4.1
1	A	377	SER	4.1
1	B	344	PRO	4.0
1	B	360	THR	3.9
1	B	355	GLY	3.8
1	B	369	PHE	3.8
1	B	372	ASN	3.7
1	B	357	PRO	3.7
1	B	346	MET	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	337	SER	3.6
1	B	185	ARG	3.5
1	B	359	PHE	3.5
1	A	271	PHE	3.4
1	B	126	VAL	3.4
1	B	137	ALA	3.2
1	B	332	GLU	3.1
1	A	355	GLY	3.1
1	A	279	PHE	3.0
1	B	133	VAL	3.0
1	B	354	PRO	3.0
1	B	130	ILE	2.9
1	B	356	SER	2.9
1	B	311	LEU	2.8
1	B	349	PHE	2.7
1	B	321	LEU	2.7
1	B	81	PRO	2.6
1	A	106	ALA	2.6
1	B	318	ALA	2.6
1	A	353	MET	2.6
1	B	330	ILE	2.5
1	B	361	ARG	2.5
1	B	336	SER	2.5
1	B	136	ALA	2.5
1	A	283	LYS	2.5
1	B	341	GLU	2.5
1	B	103	HIS	2.4
1	B	329	ALA	2.4
1	A	282	MET	2.4
1	B	317	ASP	2.4
1	B	373	THR	2.3
1	B	367	VAL	2.3
1	B	139	TYR	2.3
1	A	278	GLU	2.3
1	B	189	ILE	2.2
1	B	187	ALA	2.2
1	B	368	LYS	2.2
1	B	352	VAL	2.2
1	B	248	VAL	2.1
1	B	102	ARG	2.1
1	A	277	ALA	2.1
1	B	293	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	279	PHE	2.0
1	B	339	GLN	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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.