Full wwPDB X-ray Structure Validation Report (i)

May 13, 2020 – 12:33 am BST

PDB ID	:	5AY7
Title	:	A psychrophilic glycoside hydrolase family 10 endo-beta-1,4-xylanase
Authors	:	Zheng, Y.; Li, Y.; Liu, W.; Guo, R.T.
Deposited on	:	2015-08-10
Resolution	:	2.15 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
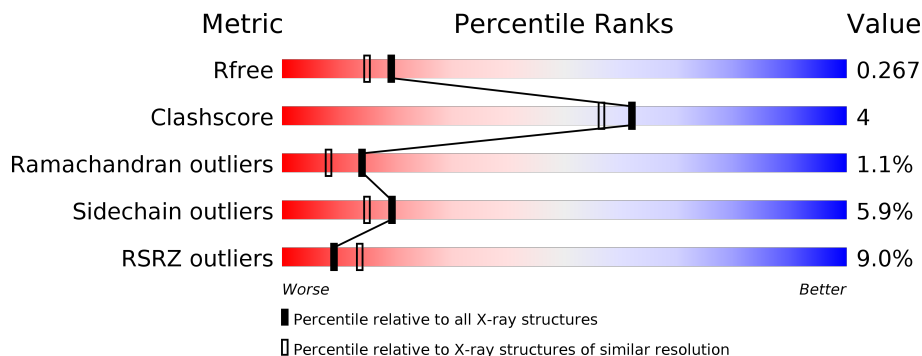with specific help available everywhere you see the (i) symbol.

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 2.15 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	Percentile Ranks	Value
Rfree		0.267
Clashscore		4
Ramachandran outliers		1.1%
Sidechain outliers		5.9%
RSRZ outliers		9.0%

Worse — Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	1479 (2.16-2.16)
Clashscore	141614	1585 (2.16-2.16)
Ramachandran outliers	138981	1560 (2.16-2.16)
Sidechain outliers	138945	1559 (2.16-2.16)
RSRZ outliers	127900	1456 (2.16-2.16)

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	355	10% / 75% 13% · 11%
1	B	355	6% / 75% 12% 5% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5441 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 xylanase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	315	Total	C	N	O	S	0	0	0
			2589	1652	440	485	12			
1	B	326	Total	C	N	O	S	0	0	0
			2670	1701	455	501	13			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	80	Total	O	0	0
			80	80		
2	B	102	Total	O	0	0
			102	102		

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: xylanase
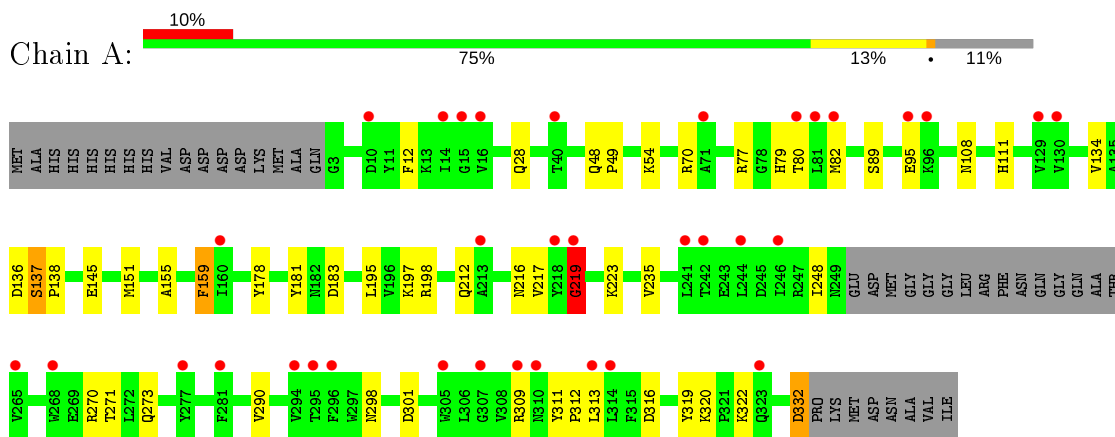
- Molecule 1: xylanase
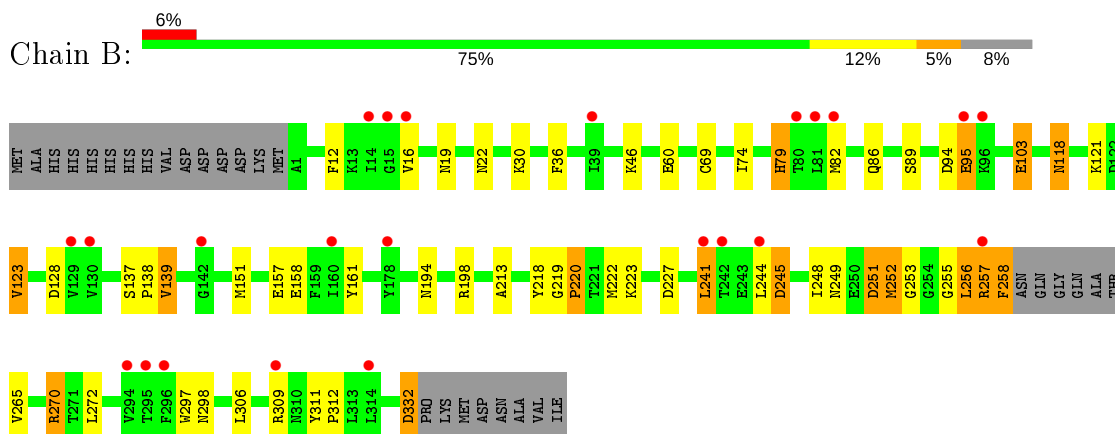
4   Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	59.43Å 78.35Å 89.30Å 90.00° 103.15° 90.00°	Depositor
Resolution (Å)	25.00 – 2.15 24.54 – 2.16	Depositor EDS
% Data completeness (in resolution range)	98.4 (25.00-2.15) 98.5 (24.54-2.16)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	4.95 (at 2.15Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.192 , 0.257 0.202 , 0.267	Depositor DCC
R_{free} test set	2069 reflections (4.86%)	wwPDB-VP
Wilson B-factor (Å^2)	44.5	Xtriage
Anisotropy	0.228	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 41.1	EDS
L-test for twinning[2]	$<|L|> = 0.50, <L^2> = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5441	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.88% 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)

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	0.63	0/2656	0.82	7/3599 (0.2%)
1	B	0.73	0/2738	0.87	4/3707 (0.1%)
All	All	0.68	0/5394	0.85	11/7306 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	B	0	7
All	All	0	12

There are no bond length outliers.

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	137	SER	C-N-CD	-8.14	102.69	120.60
1	A	77	ARG	NE-CZ-NH2	-6.63	116.98	120.30
1	B	123	VAL	CB-CA-C	-6.30	99.42	111.40
1	A	77	ARG	NE-CZ-NH1	6.07	123.33	120.30
1	B	256	LEU	CB-CG-CD2	5.96	121.13	111.00
1	A	70	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	A	136	ASP	CB-CG-OD2	-5.43	113.41	118.30
1	B	253	GLY	N-CA-C	-5.40	99.60	113.10
1	A	80	THR	N-CA-C	5.22	125.09	111.00
1	B	220	PRO	N-CA-C	-5.21	98.56	112.10
1	A	290	VAL	CB-CA-C	-5.02	101.87	111.40

There are no chirality outliers.

All (12) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	137	SER	Mainchain,Peptide
1	A	219	GLY	Peptide
1	A	79	HIS	Mainchain,Peptide
1	B	137	SER	Peptide
1	B	218	TYR	Peptide
1	B	219	GLY	Peptide
1	B	249	ASN	Peptide
1	B	257	ARG	Peptide
1	B	79	HIS	Mainchain,Peptide

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	2589	0	2489	17	0
1	B	2670	0	2566	28	0
2	A	80	0	0	0	0
2	B	102	0	0	2	0
All	All	5441	0	5055	45	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 (45) 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:265:VAL:N	2:B:401:HOH:O	2.25	0.69
1:B:265:VAL:HB	1:B:270:ARG:HD2	1.75	0.68
1:A:316:ASP:OD2	1:A:320:LYS:HB2	1.98	0.63
1:B:82:MET:HE2	1:B:151:MET:HG2	1.83	0.60
1:B:94:ASP:O	1:B:95:GLU:HB3	2.02	0.59
1:B:311:TYR:N	1:B:312:PRO:HD3	2.19	0.58
1:A:301:ASP:OD1	1:A:312:PRO:O	2.21	0.57
1:B:332:ASP:OD1	1:B:332:ASP:O	2.24	0.54
1:B:69:CYS:HB3	1:B:74:ILE:O	2.10	0.52
1:B:79:HIS:HD2	1:B:128:ASP:OD2	1.92	0.52
1:A:332:ASP:C	1:A:332:ASP:OD1	2.48	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:181:TYR:HB3	1:A:212:GLN:NE2	2.27	0.50
1:B:103:GLU:O	1:B:103:GLU:OE2	2.30	0.50
1:B:94:ASP:O	1:B:95:GLU:CB	2.58	0.50
1:B:213:ALA:HB2	1:B:241:LEU:HD13	1.93	0.50
1:A:155:ALA:HB3	1:A:159:PHE:HB3	1.94	0.49
1:B:244:LEU:HD12	1:B:245:ASP:H	1.78	0.49
1:A:301:ASP:OD2	1:A:322:LYS:NZ	2.46	0.49
1:B:256:LEU:HD12	1:B:258:PHE:CD1	2.48	0.48
1:A:311:TYR:N	1:A:312:PRO:HD3	2.29	0.48
1:B:297:TRP:O	1:B:298:ASN:HB2	2.14	0.47
1:A:82:MET:HE2	1:A:151:MET:HG2	1.95	0.47
1:B:223:LYS:HE2	1:B:227:ASP:OD1	2.15	0.47
1:B:158:GLU:HA	1:B:161:TYR:CD2	2.49	0.46
1:B:79:HIS:HE1	2:B:448:HOH:O	1.97	0.46
1:A:195:LEU:HA	1:A:198:ARG:NH1	2.32	0.44
1:B:46:LYS:HZ2	1:B:86:GLN:HE22	1.65	0.43
1:B:194:ASN:HB3	1:B:198:ARG:NH2	2.33	0.43
1:A:54:LYS:HE2	1:A:108:ASN:OD1	2.18	0.43
1:B:252:MET:HE3	1:B:306:LEU:HD21	2.02	0.42
1:B:16:VAL:HG11	1:B:36:PHE:CD2	2.55	0.42
1:A:235:VAL:O	1:A:235:VAL:HG22	2.19	0.42
1:B:118:ASN:HA	1:B:121:LYS:HD2	2.02	0.42
1:A:145:GLU:OE1	1:A:198:ARG:NH2	2.53	0.42
1:B:139:VAL:HG12	1:B:139:VAL:O	2.19	0.42
1:B:19:ASN:H	1:B:22:ASN:ND2	2.17	0.41
1:A:28:GLN:NE2	1:A:319:TYR:CZ	2.89	0.41
1:A:48:GLN:HB3	1:A:49:PRO:HD3	2.03	0.41
1:A:216:ASN:H	1:A:219:GLY:HA2	1.84	0.41
1:B:255:GLY:O	1:B:258:PHE:HB3	2.21	0.41
1:A:298:ASN:HB2	1:A:313:LEU:HD11	2.03	0.41
1:B:157:GLU:OE2	1:B:198:ARG:HD2	2.20	0.41
1:B:258:PHE:N	1:B:258:PHE:CD1	2.88	0.41
1:A:134:VAL:HG23	1:A:178:TYR:OH	2.21	0.40
1:B:251:ASP:CG	1:B:251:ASP:O	2.59	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	311/355 (88%)	293 (94%)	16 (5%)	2 (1%)	25	18
1	B	322/355 (91%)	301 (94%)	16 (5%)	5 (2%)	9	4
All	All	633/710 (89%)	594 (94%)	32 (5%)	7 (1%)	14	8

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	138	PRO
1	B	138	PRO
1	A	219	GLY
1	B	252	MET
1	B	95	GLU
1	B	139	VAL
1	B	220	PRO

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	277/309 (90%)	262 (95%)	15 (5%)	22	18
1	B	284/309 (92%)	266 (94%)	18 (6%)	18	13
All	All	561/618 (91%)	528 (94%)	33 (6%)	19	15

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	PHE
1	A	89	SER
1	A	95	GLU
1	A	111	HIS
1	A	159	PHE
1	A	183	ASP
1	A	197	LYS
1	A	217	VAL
1	A	223	LYS
1	A	248	ILE
1	A	270	ARG
1	A	271	THR
1	A	273	GLN
1	A	309	ARG
1	A	332	ASP
1	B	12	PHE
1	B	30	LYS
1	B	60	GLU
1	B	89	SER
1	B	103	GLU
1	B	118	ASN
1	B	123	VAL
1	B	222	MET
1	B	241	LEU
1	B	245	ASP
1	B	248	ILE
1	B	251	ASP
1	B	257	ARG
1	B	258	PHE
1	B	270	ARG
1	B	272	LEU
1	B	309	ARG
1	B	332	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (19) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	ASN
1	A	22	ASN
1	A	43	ASN
1	A	79	HIS
1	A	86	GLN
1	A	179	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	212	GLN
1	A	216	ASN
1	A	279	GLN
1	A	310	ASN
1	B	22	ASN
1	B	43	ASN
1	B	79	HIS
1	B	86	GLN
1	B	179	ASN
1	B	212	GLN
1	B	216	ASN
1	B	310	ASN
1	B	323	GLN

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 (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(Å²)	Q<0.9
1	A	315/355 (88%)	0.39	35 (11%)	5	7	39, 57, 89, 129	0
1	B	326/355 (91%)	0.26	23 (7%)	16	22	33, 50, 76, 104	0
All	All	641/710 (90%)	0.33	58 (9%)	9	14	33, 52, 85, 129	0

All (58) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	244	LEU	5.2
1	A	268	TRP	5.1
1	A	296	PHE	4.6
1	B	295	THR	4.5
1	B	96	LYS	4.4
1	B	296	PHE	4.3
1	B	130	VAL	4.2
1	B	129	VAL	4.2
1	A	241	LEU	4.2
1	B	294	VAL	4.2
1	A	295	THR	4.1
1	A	277	TYR	4.0
1	A	265	VAL	3.9
1	A	313	LEU	3.8
1	A	314	LEU	3.8
1	B	16	VAL	3.8
1	B	160	ILE	3.7
1	B	257	ARG	3.6
1	B	244	LEU	3.6
1	A	294	VAL	3.5
1	B	82	MET	3.5
1	B	81	LEU	3.4
1	B	241	LEU	3.4
1	A	81	LEU	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	218	TYR	3.2
1	A	96	LYS	3.2
1	A	246	ILE	3.1
1	A	213	ALA	3.0
1	A	307	GLY	3.0
1	A	16	VAL	2.8
1	B	242	THR	2.8
1	A	323	GLN	2.7
1	B	309	ARG	2.6
1	B	15	GLY	2.6
1	B	142	GLY	2.5
1	B	95	GLU	2.5
1	A	80	THR	2.4
1	B	14	ILE	2.4
1	A	10	ASP	2.4
1	A	82	MET	2.3
1	A	242	THR	2.3
1	A	309	ARG	2.3
1	B	314	LEU	2.2
1	A	219	GLY	2.2
1	A	281	PHE	2.2
1	B	178	TYR	2.2
1	B	39	ILE	2.2
1	A	130	VAL	2.2
1	A	40	THR	2.2
1	A	95	GLU	2.2
1	B	80	THR	2.1
1	A	15	GLY	2.1
1	A	14	ILE	2.1
1	A	129	VAL	2.1
1	A	310	ASN	2.1
1	A	71	ALA	2.1
1	A	160	ILE	2.0
1	A	305	TRP	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 carbohydrates 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.