



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 06:28 pm BST

PDB ID : 2ZWY
Title : alpha-L-fucosidase
Authors : Wu, H.-J.; Ko, T.-P.; Ho, C.-W.; Lin, C.-H.; Wang, A.H.-J.
Deposited on : 2008-12-18
Resolution : 2.75 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
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)	:	2.11

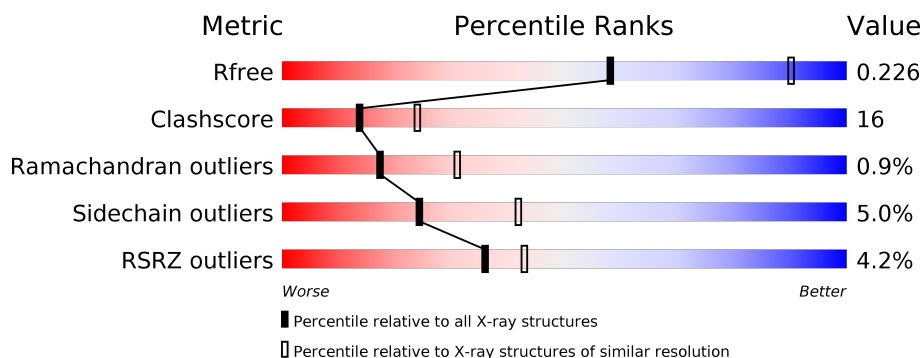
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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 $\leq 5\%$. 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	455	<div> <div>4%</div> <div> <div></div> <div>67%</div> <div>29%</div> <div>• •</div> </div> </div>
1	B	455	<div> <div>4%</div> <div> <div></div> <div>64%</div> <div>29%</div> <div>• •</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 7581 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-L-fucosidase, putative.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	442	Total	C	N	O	S	0	0	0
			3648	2377	600	663	8			
1	B	442	Total	C	N	O	S	0	0	0
			3648	2377	600	663	8			

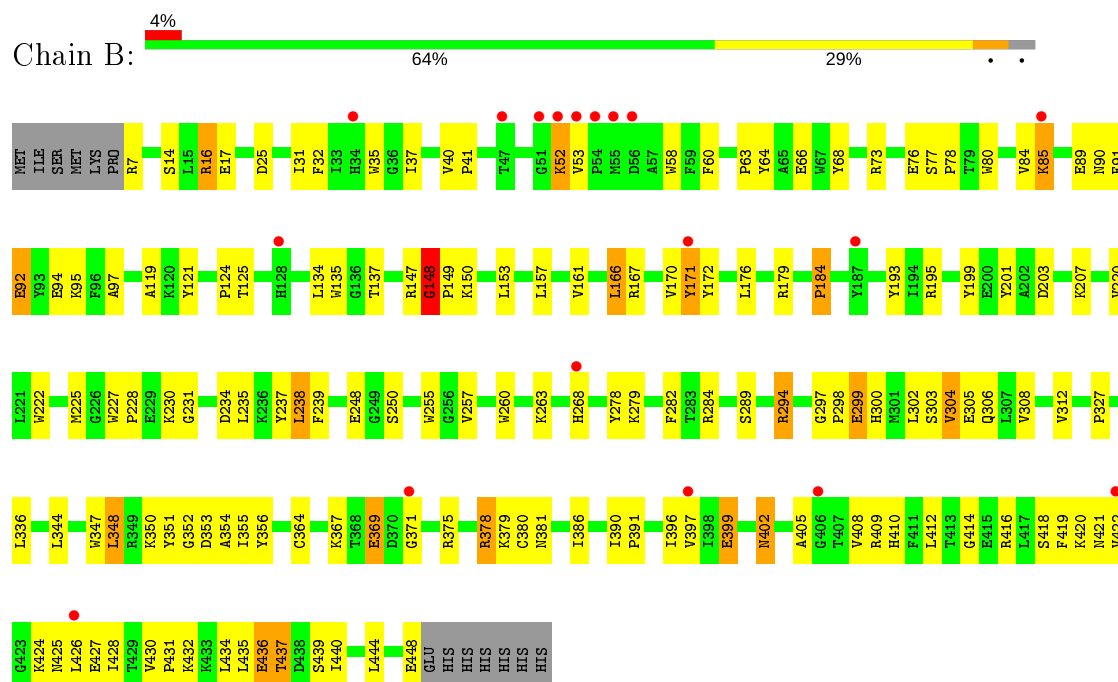
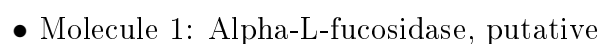
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	450	HIS	-	EXPRESSION TAG	UNP Q9WYE2
A	451	HIS	-	EXPRESSION TAG	UNP Q9WYE2
A	452	HIS	-	EXPRESSION TAG	UNP Q9WYE2
A	453	HIS	-	EXPRESSION TAG	UNP Q9WYE2
A	454	HIS	-	EXPRESSION TAG	UNP Q9WYE2
A	455	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	450	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	451	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	452	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	453	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	454	HIS	-	EXPRESSION TAG	UNP Q9WYE2
B	455	HIS	-	EXPRESSION TAG	UNP Q9WYE2

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	145	Total	O	0	0
			145	145		
2	B	140	Total	O	0	0
			140	140		

- Molecule 1: Alpha-L-fucosidase, putative



4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α , β , γ	180.59Å 180.59Å 169.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.75 29.40 – 2.75	Depositor EDS
% Data completeness (in resolution range)	95.3 (30.00-2.75) 95.2 (29.40-2.75)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	0.61	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.61 (at 2.76Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.169 , 0.226 0.169 , 0.226	Depositor DCC
R_{free} test set	1277 reflections (4.83%)	wwPDB-VP
Wilson B-factor (Å ²)	46.1	Xtriage
Anisotropy	0.154	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 62.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7581	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	0.99	3/3768 (0.1%)	0.97	1/5120 (0.0%)
1	B	1.00	4/3768 (0.1%)	0.98	5/5120 (0.1%)
All	All	0.99	7/7536 (0.1%)	0.98	6/10240 (0.1%)

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	B	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	248	GLU	CG-CD	5.58	1.60	1.51
1	B	193	TYR	CD1-CE1	5.58	1.47	1.39
1	A	305	GLU	CG-CD	5.51	1.60	1.51
1	B	170	VAL	CA-CB	5.28	1.65	1.54
1	B	60	PHE	CE1-CZ	5.17	1.47	1.37

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	179	ARG	NE-CZ-NH2	8.32	124.46	120.30
1	B	380	CYS	CA-CB-SG	-6.37	102.53	114.00
1	B	195	ARG	NE-CZ-NH1	-6.14	117.23	120.30
1	A	203	ASP	CB-CG-OD2	-5.71	113.16	118.30
1	B	148	GLY	N-CA-C	5.47	126.77	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	237	TYR	Sidechain

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	3648	0	3512	121	0
1	B	3648	0	3512	114	0
2	A	145	0	0	0	1
2	B	140	0	0	2	0
All	All	7581	0	7024	227	1

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 16.

The worst 5 of 227 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:378:ARG:HG2	1:B:378:ARG:HH11	1.24	1.03
1:A:269:VAL:HG12	1:A:270:ASN:ND2	1.84	0.93
1:B:297:GLY:H	1:B:300:HIS:HD2	1.17	0.93
1:B:409:ARG:NH1	1:B:414:GLY:O	2.02	0.93
1:A:381:ASN:HD22	1:A:381:ASN:H	0.92	0.91

All (1) 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:535:HOH:O	2:A:535:HOH:O[2_555]	1.72	0.48

5.3 Torsion angles

5.3.1 Protein backbone

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	440/455 (97%)	395 (90%)	42 (10%)	3 (1%)	22	39
1	B	440/455 (97%)	401 (91%)	34 (8%)	5 (1%)	14	25
All	All	880/910 (97%)	796 (90%)	76 (9%)	8 (1%)	17	31

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	267	TYR
1	A	148	GLY
1	B	148	GLY
1	B	90	ASN
1	B	352	GLY

5.3.2 Protein sidechains

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	382/395 (97%)	370 (97%)	12 (3%)	40	60
1	B	382/395 (97%)	356 (93%)	26 (7%)	16	28
All	All	764/790 (97%)	726 (95%)	38 (5%)	24	42

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	92	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	184	PRO
1	B	418	SER
1	B	166	LEU
1	B	222	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	270	ASN
1	B	300	HIS
1	A	381	ASN
1	A	268	HIS
1	B	108	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(Å ²)	Q<0.9
1	A	442/455 (97%)	-0.14	19 (4%) 35 42	21, 39, 70, 83	0
1	B	442/455 (97%)	-0.10	18 (4%) 37 44	21, 41, 72, 87	0
All	All	884/910 (97%)	-0.12	37 (4%) 36 43	21, 40, 72, 87	0

The worst 5 of 37 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	55	MET	6.4
1	A	51	GLY	5.1
1	A	48	GLY	5.0
1	A	54	PRO	4.7
1	A	50	LEU	3.9

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.