



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 03:36 PM UTC

PDB ID : 6VE1 / pdb_00006ve1
Title : Crystal structure of endo-beta-N-acetylglucosaminidase H at high pH
Authors : Stachowski, T.R.; Snell, M.E.; Snell, E.S.
Deposited on : 2019-12-28
Resolution : 2.10 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

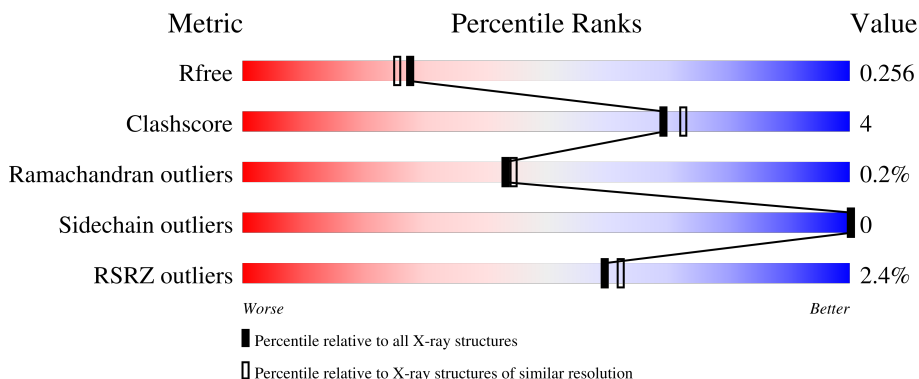
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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}	180053	6658 (2.10-2.10)
Clashscore	190562	7164 (2.10-2.10)
Ramachandran outliers	187476	7099 (2.10-2.10)
Sidechain outliers	187428	7100 (2.10-2.10)
RSRZ outliers	180081	6662 (2.10-2.10)

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 $\leq 5\%$. 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	281	
1	B	281	
1	C	281	
1	D	281	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 17020 atoms, of which 7767 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 Endo-beta-N-acetylglucosaminidase H.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	267	Total	C	H	N	O	S	0	0	0
			3969	1274	1942	349	402	2			
1	B	267	Total	C	H	N	O	S	0	0	0
			3969	1274	1942	349	402	2			
1	C	268	Total	C	H	N	O	S	0	0	0
			3971	1276	1940	349	403	3			
1	D	267	Total	C	H	N	O	S	0	0	0
			3969	1274	1943	348	401	3			

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P04067
A	2	LEU	-	expression tag	UNP P04067
A	3	SER	-	expression tag	UNP P04067
A	4	THR	-	expression tag	UNP P04067
A	5	GLY	-	expression tag	UNP P04067
A	6	CYS	-	expression tag	UNP P04067
A	7	TYR	-	expression tag	UNP P04067
A	8	MET	-	expression tag	UNP P04067
A	276	HIS	-	expression tag	UNP P04067
A	277	HIS	-	expression tag	UNP P04067
A	278	HIS	-	expression tag	UNP P04067
A	279	HIS	-	expression tag	UNP P04067
A	280	HIS	-	expression tag	UNP P04067
A	281	HIS	-	expression tag	UNP P04067
B	1	SER	-	expression tag	UNP P04067
B	2	LEU	-	expression tag	UNP P04067
B	3	SER	-	expression tag	UNP P04067
B	4	THR	-	expression tag	UNP P04067
B	5	GLY	-	expression tag	UNP P04067
B	6	CYS	-	expression tag	UNP P04067
B	7	TYR	-	expression tag	UNP P04067

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	8	MET	-	expression tag	UNP P04067
B	276	HIS	-	expression tag	UNP P04067
B	277	HIS	-	expression tag	UNP P04067
B	278	HIS	-	expression tag	UNP P04067
B	279	HIS	-	expression tag	UNP P04067
B	280	HIS	-	expression tag	UNP P04067
B	281	HIS	-	expression tag	UNP P04067
C	1	SER	-	expression tag	UNP P04067
C	2	LEU	-	expression tag	UNP P04067
C	3	SER	-	expression tag	UNP P04067
C	4	THR	-	expression tag	UNP P04067
C	5	GLY	-	expression tag	UNP P04067
C	6	CYS	-	expression tag	UNP P04067
C	7	TYR	-	expression tag	UNP P04067
C	8	MET	-	expression tag	UNP P04067
C	276	HIS	-	expression tag	UNP P04067
C	277	HIS	-	expression tag	UNP P04067
C	278	HIS	-	expression tag	UNP P04067
C	279	HIS	-	expression tag	UNP P04067
C	280	HIS	-	expression tag	UNP P04067
C	281	HIS	-	expression tag	UNP P04067
D	1	SER	-	expression tag	UNP P04067
D	2	LEU	-	expression tag	UNP P04067
D	3	SER	-	expression tag	UNP P04067
D	4	THR	-	expression tag	UNP P04067
D	5	GLY	-	expression tag	UNP P04067
D	6	CYS	-	expression tag	UNP P04067
D	7	TYR	-	expression tag	UNP P04067
D	8	MET	-	expression tag	UNP P04067
D	276	HIS	-	expression tag	UNP P04067
D	277	HIS	-	expression tag	UNP P04067
D	278	HIS	-	expression tag	UNP P04067
D	279	HIS	-	expression tag	UNP P04067
D	280	HIS	-	expression tag	UNP P04067
D	281	HIS	-	expression tag	UNP P04067

- Molecule 2 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total Mg 2 2	0	0
2	B	1	Total Mg 1 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	C	2	Total 2	Mg 2	0	0
2	D	1	Total 1	Mg 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	288	Total 288	O 288	0	0
3	B	300	Total 300	O 300	0	0
3	C	285	Total 285	O 285	0	0
3	D	263	Total 263	O 263	0	0

3 Residue-property plots

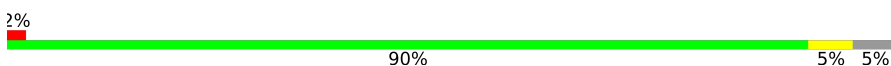
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Endo-beta-N-acetylglucosaminidase H

Chain A: 



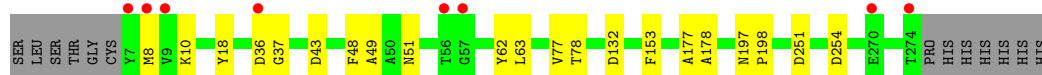
• Molecule 1: Endo-beta-N-acetylglucosaminidase H

Chain B: 



• Molecule 1: Endo-beta-N-acetylglucosaminidase H

Chain C: 



• Molecule 1: Endo-beta-N-acetylglucosaminidase H

Chain D: 



HIS
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2 ₁	Depositor
Cell constants a, b, c, α , β , γ	93.48Å 99.48Å 135.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.03 – 2.10 41.03 – 2.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (41.03-2.10) 99.8 (41.03-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.57 (at 1.82Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.221 , 0.258 0.222 , 0.256	Depositor DCC
R_{free} test set	5666 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	9.4	Xtriage
Anisotropy	1.160	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 42.9	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	0.92	EDS
Total number of atoms	17020	wwPDB-VP
Average B, all atoms (Å ²)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 53.52 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.1666e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG

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	0.41	0/2069	0.48	0/2820
1	B	0.40	0/2069	0.48	0/2820
1	C	0.37	0/2073	0.43	0/2826
1	D	0.41	1/2068 (0.0%)	0.48	0/2818
All	All	0.40	1/8279 (0.0%)	0.47	0/11284

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	273	ARG	CA-C	-5.49	1.50	1.53

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2027	1942	1941	13	0
1	B	2027	1942	1942	11	0
1	C	2031	1940	1940	15	0
1	D	2026	1943	1943	17	0
2	A	2	0	0	0	0
2	B	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	2	0	0	0	0
2	D	1	0	0	0	0
3	A	288	0	0	0	0
3	B	300	0	0	0	0
3	C	285	0	0	3	0
3	D	263	0	0	2	0
All	All	9253	7767	7766	56	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.

The worst 5 of 56 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:234:ALA:HB1	1:B:266:LEU:CD1	2.15	0.75
1:B:234:ALA:HB1	1:B:266:LEU:HD13	1.69	0.73
1:D:160:LEU:HD11	1:D:164:MET:HE2	1.77	0.66
1:D:236:ARG:O	1:D:239:ASP:O	2.19	0.60
1:B:262:PHE:HD1	1:B:266:LEU:HD22	1.66	0.59

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	265/281 (94%)	253 (96%)	11 (4%)	1 (0%)	30	28
1	B	265/281 (94%)	258 (97%)	7 (3%)	0	100	100
1	C	266/281 (95%)	256 (96%)	10 (4%)	0	100	100
1	D	265/281 (94%)	252 (95%)	12 (4%)	1 (0%)	30	28
All	All	1061/1124 (94%)	1019 (96%)	40 (4%)	2 (0%)	43	44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	55	ASP
1	A	12	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	209/223 (94%)	209 (100%)	0	100	100
1	B	209/223 (94%)	209 (100%)	0	100	100
1	C	209/223 (94%)	209 (100%)	0	100	100
1	D	209/223 (94%)	209 (100%)	0	100	100
All	All	836/892 (94%)	836 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	214	GLN
1	B	68	ASN
1	B	98	HIS
1	B	147	GLN
1	D	204	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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	267/281 (95%)	0.19	7 (2%) 57 60	5, 14, 29, 54	0
1	B	267/281 (95%)	0.08	6 (2%) 62 65	5, 13, 28, 56	0
1	C	268/281 (95%)	0.22	8 (2%) 52 55	5, 15, 32, 47	0
1	D	267/281 (95%)	0.23	5 (1%) 66 69	4, 16, 31, 51	0
All	All	1069/1124 (95%)	0.18	26 (2%) 59 62	4, 15, 31, 56	0

The worst 5 of 26 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	275	PRO	7.3
1	A	274	THR	5.3
1	C	7	TYR	5.0
1	B	8	MET	4.9
1	D	274	THR	4.8

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
2	MG	A	301	1/1	0.79	0.15	24,24,24,24	1
2	MG	C	301	1/1	0.85	0.18	26,26,26,26	0
2	MG	B	301	1/1	0.91	0.12	25,25,25,25	0
2	MG	C	302	1/1	0.91	0.15	27,27,27,27	0
2	MG	A	302	1/1	0.95	0.12	30,30,30,30	0
2	MG	D	301	1/1	0.96	0.08	22,22,22,22	0

6.5 Other polymers [i](#)

There are no such residues in this entry.