



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 16, 2025 – 01:16 pm BST

PDB ID : 9EUX / pdb_00009eux
Title : Glycoside hydrolase family 191 enzyme from *Thermotoga maritima*
Authors : Roth, C.
Deposited on : 2024-03-28
Resolution : 2.40 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	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

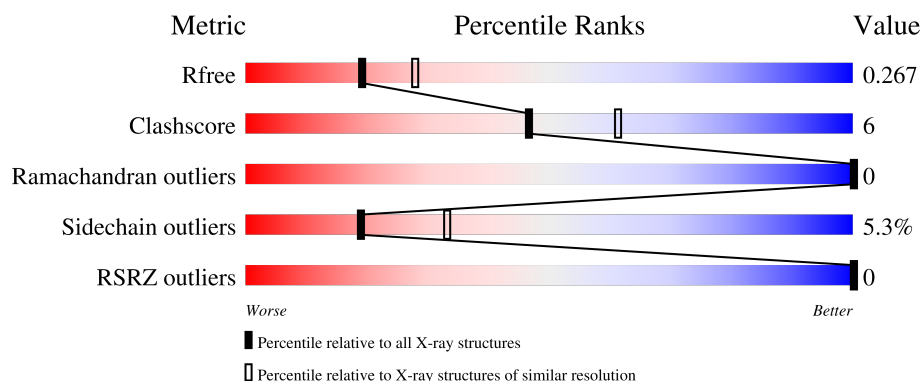
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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}	164625	4642 (2.40-2.40)
Clashscore	180529	5218 (2.40-2.40)
Ramachandran outliers	177936	5158 (2.40-2.40)
Sidechain outliers	177891	5159 (2.40-2.40)
RSRZ outliers	164620	4642 (2.40-2.40)

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	323	
1	B	323	
1	C	323	
1	D	323	
1	E	323	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	323	<div><div></div><div>76%</div><div>9% • 12%</div></div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 27992 atoms, of which 13354 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 Uncharacterized protein TM_1410.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	G	284	Total	C	H	N	O	S	47	4	0
			4606	1524	2240	376	460	6			
1	E	282	Total	C	H	N	O	S	46	2	0
			4560	1510	2217	372	455	6			
1	C	282	Total	C	H	N	O	S	46	4	0
			4592	1519	2235	376	456	6			
1	D	280	Total	C	H	N	O	S	46	2	0
			4538	1503	2208	370	451	6			
1	B	284	Total	C	H	N	O	S	46	0	0
			4573	1515	2223	373	456	6			
1	A	283	Total	C	H	N	O	S	46	2	0
			4583	1516	2231	376	454	6			


- Molecule 2 is water.

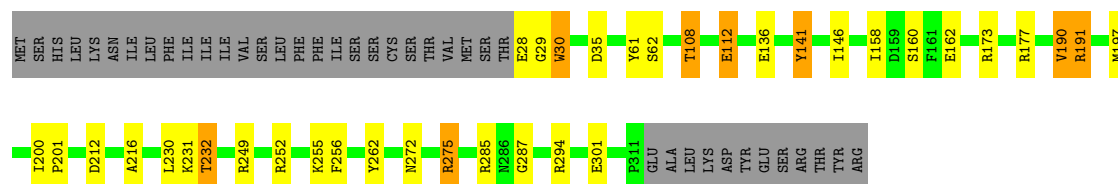
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	83	Total	O	0	0
			83	83		
2	E	92	Total	O	0	0
			92	92		
2	C	75	Total	O	0	0
			75	75		
2	D	93	Total	O	0	0
			93	93		
2	B	94	Total	O	0	0
			94	94		
2	A	103	Total	O	0	0
			103	103		

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.

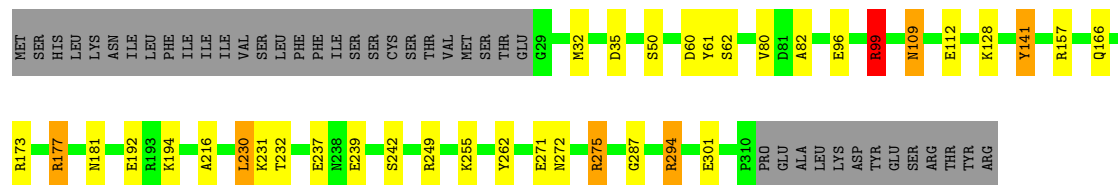
• Molecule 1: Uncharacterized protein TM_1410

Chain G: 



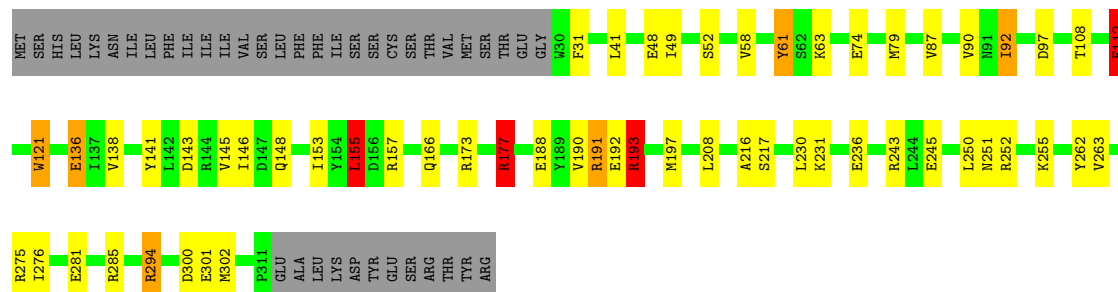
• Molecule 1: Uncharacterized protein TM_1410

Chain E: 



• Molecule 1: Uncharacterized protein TM_1410

Chain C: 



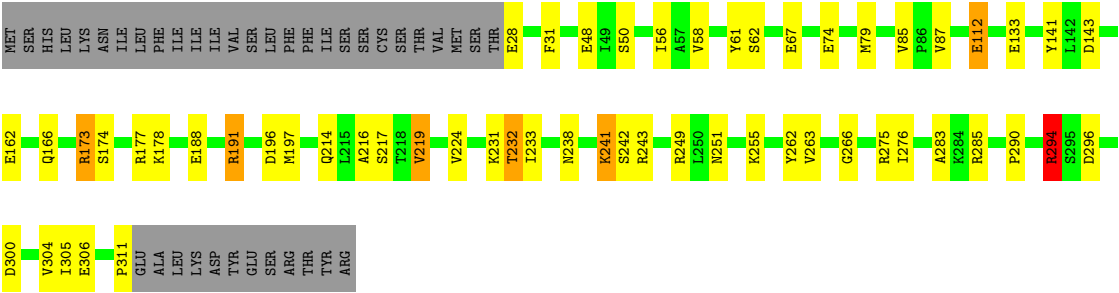
• Molecule 1: Uncharacterized protein TM_1410

Chain D: 

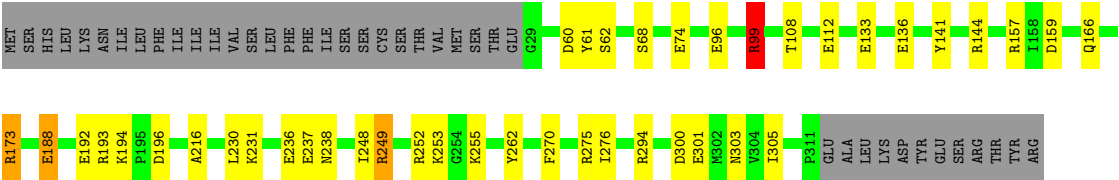




• Molecule 1: Uncharacterized protein TM_1410



• Molecule 1: Uncharacterized protein TM_1410



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	195.19Å 84.50Å 195.48Å 90.00° 119.91° 90.00°	Depositor
Resolution (Å)	48.95 – 2.40 48.95 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.95-2.40) 99.4 (48.95-2.40)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.23 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.8.0425, REFMAC 5.8.0425	Depositor
R, R_{free}	0.236 , 0.264 0.237 , 0.267	Depositor DCC
R_{free} test set	11154 reflections (5.16%)	wwPDB-VP
Wilson B-factor (Å ²)	45.9	Xtriage
Anisotropy	0.431	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 22.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.38$, $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	0.207 for -h-l,k,h 0.207 for l,k,-h-l 0.098 for h,-k,-h-l 0.096 for -h-l,-k,l 0.098 for l,-k,h	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	27992	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.14% 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

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	0.68	2/2423 (0.1%)	1.19	11/3290 (0.3%)
1	B	0.68	1/2415 (0.0%)	1.16	16/3280 (0.5%)
1	C	0.70	3/2444 (0.1%)	1.18	24/3319 (0.7%)
1	D	0.70	3/2399 (0.1%)	1.15	9/3258 (0.3%)
1	E	0.64	3/2423 (0.1%)	1.13	10/3291 (0.3%)
1	G	0.68	3/2448 (0.1%)	1.12	13/3326 (0.4%)
All	All	0.68	15/14552 (0.1%)	1.15	83/19764 (0.4%)

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	3
1	B	0	3
1	C	0	3
1	D	0	4
1	E	0	1
1	G	0	1
All	All	0	15

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	192	GLU	CD-OE2	10.04	1.36	1.25
1	D	192	GLU	CD-OE1	7.15	1.33	1.25
1	E	192	GLU	CD-OE1	7.03	1.33	1.25
1	G	136	GLU	CD-OE2	6.97	1.33	1.25
1	C	136	GLU	CD-OE2	6.87	1.33	1.25

The worst 5 of 83 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	99	ARG	NE-CZ-NH1	15.12	127.86	120.30
1	C	300	ASP	CB-CA-C	-13.09	84.22	110.40
1	D	309	GLN	N-CA-CB	12.43	132.98	110.60
1	A	300	ASP	CB-CA-C	-12.12	86.17	110.40
1	A	99	ARG	NE-CZ-NH2	-12.10	114.25	120.30

There are no chirality outliers.

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	193[A]	ARG	Sidechain
1	C	243	ARG	Sidechain
1	C	294	ARG	Sidechain
1	E	177	ARG	Sidechain
1	G	249	ARG	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	2352	2231	2224	22	0
1	B	2350	2223	2222	34	0
1	C	2357	2235	2210	42	0
1	D	2330	2208	2201	27	0
1	E	2343	2217	2205	18	0
1	G	2366	2240	2220	15	0
2	A	103	0	0	8	0
2	B	94	0	0	4	0
2	C	75	0	0	8	0
2	D	93	0	0	3	0
2	E	92	0	0	4	0
2	G	83	0	0	0	0
All	All	14638	13354	13282	154	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 6.

The worst 5 of 154 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:G:191:ARG:NH1	1:G:197:MET:O	1.91	1.04
1:C:191:ARG:NH1	1:C:197:MET:O	1.90	1.02
1:B:191:ARG:NH1	1:B:197:MET:O	1.92	1.02
1:B:58:VAL:HG22	1:B:87:VAL:CG2	2.04	0.87
1:B:56:ILE:HG13	1:B:85:VAL:HG23	1.56	0.85

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	283/323 (88%)	273 (96%)	10 (4%)	0	100	100
1	B	282/323 (87%)	271 (96%)	11 (4%)	0	100	100
1	C	284/323 (88%)	275 (97%)	9 (3%)	0	100	100
1	D	280/323 (87%)	269 (96%)	11 (4%)	0	100	100
1	E	282/323 (87%)	273 (97%)	9 (3%)	0	100	100
1	G	286/323 (88%)	273 (96%)	13 (4%)	0	100	100
All	All	1697/1938 (88%)	1634 (96%)	63 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	252/289 (87%)	238 (94%)	14 (6%)	17	30
1	B	251/289 (87%)	237 (94%)	14 (6%)	17	30
1	C	254/289 (88%)	239 (94%)	15 (6%)	16	28
1	D	249/289 (86%)	237 (95%)	12 (5%)	21	37
1	E	251/289 (87%)	236 (94%)	15 (6%)	16	27
1	G	255/289 (88%)	241 (94%)	14 (6%)	18	31
All	All	1512/1734 (87%)	1428 (94%)	84 (6%)	19	30

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	62	SER
1	A	62	SER
1	B	112	GLU
1	B	232	THR
1	A	141	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	94	GLN
1	D	203	ASN
1	A	303	ASN
1	B	251	ASN
1	C	42	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides 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	283/323 (87%)	-1.48	0 100 100	20, 47, 76, 100	1 (0%)
1	B	284/323 (87%)	-1.46	0 100 100	28, 47, 75, 100	0
1	C	282/323 (87%)	-1.36	0 100 100	23, 53, 81, 105	2 (0%)
1	D	280/323 (86%)	-1.38	0 100 100	26, 53, 86, 109	1 (0%)
1	E	282/323 (87%)	-1.39	0 100 100	23, 53, 86, 101	1 (0%)
1	G	284/323 (87%)	-1.37	0 100 100	25, 52, 83, 138	2 (0%)
All	All	1695/1938 (87%)	-1.41	0 100 100	20, 51, 82, 138	7 (0%)

There are no RSRZ outliers to report.

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.