



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 09:14 PM UTC

PDB ID : 2YIK / pdb_00002yik
Title : Catalytic domain of Clostridium thermocellum CelT
Authors : Tsai, J.-Y.; Kesavulu, M.M.; Hsiao, C.-D.
Deposited on : 2011-05-16
Resolution : 2.10 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 ⓘ 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

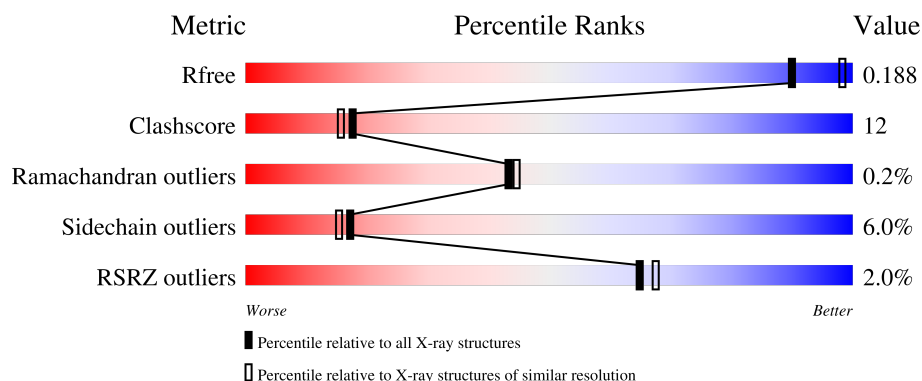
1 Overall quality at a glance

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	611	 2% 64% 14% 19%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4281 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 ENDOGLUCANASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	494	Total	C	N	O	S	0	0	1
			3934	2501	661	754	18			

- Molecule 2 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Ca	0	0
			3	3		

- Molecule 3 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	343	Total	O	0	0
			343	343		

4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	96.86Å 96.86Å 159.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.38 – 2.10 29.38 – 2.10	Depositor EDS
% Data completeness (in resolution range)	85.9 (29.38-2.10) 90.5 (29.38-2.10)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.42 (at 2.10Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.182 , 0.199 0.178 , 0.188	Depositor DCC
R_{free} test set	2136 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å ²)	27.4	Xtriage
Anisotropy	0.572	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 38.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	4281	wwPDB-VP
Average B, all atoms (Å ²)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.40% 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

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

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	0.54	4/4059 (0.1%)	1.04	24/5521 (0.4%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	270	LYS	CA-CB	6.55	1.64	1.53
1	A	315	ILE	CA-CB	6.14	1.64	1.54
1	A	269	PRO	C-O	5.22	1.31	1.24
1	A	530	HIS	C-N	-5.09	1.26	1.33

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	268	GLY	CA-C-N	-11.89	106.66	121.04
1	A	268	GLY	C-N-CA	-11.89	106.66	121.04
1	A	311	PRO	N-CA-C	8.40	120.95	110.70
1	A	136	TRP	CA-CB-CG	7.86	128.54	113.60
1	A	60	VAL	N-CA-C	7.32	118.12	110.72
1	A	124	GLY	N-CA-C	6.84	120.94	112.73
1	A	519	SER	N-CA-C	-6.56	104.76	112.89
1	A	389	GLN	N-CA-C	6.50	120.47	112.54
1	A	270	LYS	CA-C-O	-6.20	111.64	120.51
1	A	116	ASP	N-CA-C	6.18	118.94	111.40
1	A	269	PRO	CA-C-N	6.15	133.28	121.54
1	A	269	PRO	C-N-CA	6.15	133.28	121.54
1	A	518	ASP	CB-CA-C	-6.15	99.72	112.44
1	A	115	HIS	N-CA-C	-6.09	101.91	110.50
1	A	268	GLY	N-CA-C	5.97	124.52	112.34
1	A	339	ASP	N-CA-C	5.76	119.29	112.72
1	A	503	ALA	N-CA-C	5.75	118.04	111.02

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	293	THR	N-CA-C	5.57	121.49	114.31
1	A	270	LYS	CA-C-N	-5.54	112.86	122.83
1	A	270	LYS	C-N-CA	-5.54	112.86	122.83
1	A	353	LYS	N-CA-C	-5.38	101.81	110.14
1	A	257	LYS	N-CA-C	5.30	117.97	111.82
1	A	135	GLY	CA-C-N	5.03	126.98	120.44
1	A	135	GLY	C-N-CA	5.03	126.98	120.44

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	3934	0	3595	93	0
2	A	3	0	0	0	0
3	A	1	0	0	0	0
4	A	343	0	0	10	1
All	All	4281	0	3595	93	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 12.

All (93) 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:A:210:THR:HG22	1:A:212:GLU:H	1.29	0.96
1:A:445:ASN:HD22	1:A:445:ASN:H	1.13	0.91
1:A:54:ASN:HD21	1:A:452:HIS:HE1	1.20	0.89
1:A:396:ASN:HD21	1:A:424:GLN:HE21	1.23	0.86
1:A:60:VAL:HG13	1:A:71:ASP:HA	1.58	0.86
1:A:436:ARG:HH11	1:A:436:ARG:HG3	1.37	0.86
1:A:462:ASN:HD22	1:A:490:ASN:HD22	1.20	0.86
1:A:523:ILE:HD12	4:A:2340:HOH:O	1.76	0.85
1:A:372:TRP:HA	1:A:377:ILE:HD13	1.59	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:136:TRP:HZ3	1:A:140:GLU:OE2	1.61	0.84
1:A:475:LEU:H	1:A:502:ASN:HD21	1.27	0.82
1:A:452:HIS:HD2	1:A:455:ALA:H	1.26	0.80
1:A:315:ILE:HD11	1:A:399:ALA:HA	1.65	0.78
1:A:215:GLY:O	1:A:219:ILE:HD13	1.84	0.77
1:A:269:PRO:HG2	4:A:2196:HOH:O	1.83	0.77
1:A:153:HIS:O	1:A:157:ILE:HD13	1.85	0.76
1:A:358:GLN:C	1:A:359:ILE:HD12	2.11	0.74
1:A:48:MET:HE3	4:A:2006:HOH:O	1.89	0.73
1:A:499:ILE:HD12	1:A:500:ASP:N	2.04	0.72
1:A:54:ASN:ND2	1:A:452:HIS:HE1	1.87	0.71
1:A:210:THR:CG2	1:A:212:GLU:H	2.03	0.71
1:A:81:LEU:HD11	1:A:93:ILE:HD12	1.72	0.71
1:A:377:ILE:HD12	1:A:377:ILE:N	2.07	0.69
1:A:385:VAL:H	1:A:424:GLN:HE22	1.39	0.69
1:A:348:LYS:HE2	1:A:359:ILE:HD13	1.75	0.68
1:A:385:VAL:H	1:A:424:GLN:NE2	1.92	0.67
1:A:371:LYS:HG2	1:A:377:ILE:CD1	2.25	0.67
1:A:518:ASP:HB2	1:A:520:SER:H	1.59	0.66
1:A:119:ASP:OD2	1:A:191:HIS:HE1	1.80	0.65
1:A:436:ARG:HG3	1:A:436:ARG:NH1	2.11	0.64
1:A:315:ILE:HG13	1:A:316:HIS:H	1.65	0.62
1:A:136:TRP:CZ3	1:A:140:GLU:OE2	2.47	0.62
1:A:315:ILE:HD13	1:A:399:ALA:HB2	1.82	0.61
1:A:239:GLN:NE2	4:A:2179:HOH:O	2.33	0.60
1:A:452:HIS:CD2	1:A:455:ALA:H	2.16	0.60
1:A:54:ASN:HD21	1:A:452:HIS:CE1	2.10	0.59
1:A:210:THR:HG22	1:A:212:GLU:N	2.10	0.59
1:A:277:LYS:N	1:A:277:LYS:HD3	2.18	0.58
1:A:523:ILE:CD1	1:A:523:ILE:H	2.17	0.58
1:A:371:LYS:HG2	1:A:377:ILE:HD11	1.85	0.57
1:A:217:ASP:CG	1:A:218:ILE:HD12	2.29	0.57
1:A:366:GLN:NE2	4:A:2273:HOH:O	2.38	0.56
1:A:519:SER:O	1:A:522:GLN:NE2	2.38	0.55
1:A:372:TRP:CA	1:A:377:ILE:HD13	2.34	0.55
1:A:130:ALA:HB2	1:A:499:ILE:HD13	1.89	0.54
1:A:371:LYS:HG2	1:A:377:ILE:HD12	1.88	0.54
1:A:209:ILE:HG21	1:A:219:ILE:HD11	1.89	0.54
1:A:120:HIS:HE1	4:A:2084:HOH:O	1.91	0.52
1:A:440:VAL:HG13	1:A:474:ALA:HB3	1.92	0.52
1:A:523:ILE:N	1:A:523:ILE:HD13	2.25	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:218:ILE:HD13	1:A:263:VAL:CG2	2.41	0.51
1:A:437:CYS:HB3	1:A:446:SER:HA	1.92	0.51
1:A:359:ILE:HD12	1:A:359:ILE:N	2.26	0.51
1:A:445:ASN:HD22	1:A:445:ASN:N	1.90	0.51
1:A:117:ALA:HA	1:A:498:ALA:HA	1.93	0.51
1:A:396:ASN:ND2	1:A:424:GLN:HE21	2.00	0.50
1:A:52:ASP:OD1	1:A:153:HIS:HE1	1.94	0.50
1:A:445:ASN:H	1:A:445:ASN:ND2	1.93	0.50
1:A:523:ILE:CD1	1:A:523:ILE:N	2.75	0.49
1:A:188:ASP:OD2	1:A:270:LYS:HE3	2.12	0.49
1:A:436:ARG:HD3	1:A:447:VAL:HG12	1.95	0.48
1:A:348:LYS:CE	1:A:359:ILE:HD13	2.42	0.48
1:A:462:ASN:ND2	1:A:490:ASN:HD22	1.99	0.48
1:A:182:HIS:HE1	4:A:2051:HOH:O	1.97	0.47
1:A:153:HIS:CE1	1:A:157:ILE:HD11	2.49	0.47
1:A:182:HIS:HD2	4:A:2087:HOH:O	1.98	0.47
1:A:450:PRO:O	1:A:467:HIS:HE1	1.97	0.47
1:A:81:LEU:CD1	1:A:93:ILE:HD12	2.44	0.47
1:A:371:LYS:C	1:A:377:ILE:CD1	2.89	0.46
1:A:371:LYS:C	1:A:377:ILE:HD13	2.41	0.45
1:A:436:ARG:HD3	1:A:447:VAL:CG1	2.47	0.45
1:A:436:ARG:NH1	4:A:2315:HOH:O	2.49	0.45
1:A:427:TYR:C	1:A:427:TYR:CD1	2.97	0.43
1:A:377:ILE:N	1:A:377:ILE:CD1	2.77	0.43
1:A:491:ASP:OD2	1:A:494:TYR:HB2	2.19	0.43
1:A:130:ALA:CB	1:A:499:ILE:HD13	2.48	0.43
1:A:315:ILE:HG13	1:A:316:HIS:N	2.31	0.43
1:A:523:ILE:HD13	4:A:2341:HOH:O	2.17	0.42
1:A:209:ILE:HD13	1:A:219:ILE:CD1	2.49	0.42
1:A:219:ILE:CD1	1:A:219:ILE:N	2.82	0.42
1:A:157:ILE:HD12	1:A:157:ILE:N	2.34	0.42
1:A:277:LYS:HE3	1:A:280:ASP:OD1	2.19	0.42
1:A:495:ASN:HD22	1:A:495:ASN:HA	1.54	0.42
1:A:48:MET:HE1	1:A:150:GLN:HB3	2.02	0.42
1:A:260:LYS:HB3	1:A:278:TRP:CE2	2.55	0.41
1:A:453:ARG:HH11	1:A:453:ARG:HG3	1.86	0.41
1:A:52:ASP:OD1	1:A:153:HIS:CE1	2.73	0.41
1:A:218:ILE:HD12	1:A:218:ILE:N	2.36	0.41
1:A:46:TYR:O	1:A:473:GLY:HA3	2.21	0.41
1:A:440:VAL:HG22	1:A:471:LEU:O	2.20	0.41
1:A:120:HIS:CD2	1:A:200:ASP:OD2	2.74	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:408:LYS:HA	1:A:412:ASP:HA	2.03	0.40
1:A:475:LEU:H	1:A:502:ASN:ND2	2.08	0.40

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)
4:A:2172:HOH:O	4:A:2172:HOH:O[7_555]	1.91	0.29

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	492/611 (80%)	477 (97%)	14 (3%)	1 (0%)	43 44

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	270	LYS

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	402/505 (80%)	378 (94%)	24 (6%)	17 15

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	60	VAL
1	A	61	ASP
1	A	78	ARG
1	A	136	TRP
1	A	142	LYS
1	A	209	ILE
1	A	210	THR
1	A	219	ILE
1	A	229	ASN
1	A	236	THR
1	A	270	LYS
1	A	277	LYS
1	A	294	GLN
1	A	315	ILE
1	A	335	LEU
1	A	377	ILE
1	A	379	VAL
1	A	440	VAL
1	A	445	ASN
1	A	468	LYS
1	A	495	ASN
1	A	518	ASP
1	A	523	ILE
1	A	529	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	96	ASN
1	A	120	HIS
1	A	150	GLN
1	A	153	HIS
1	A	176	ASN
1	A	182	HIS
1	A	191	HIS
1	A	229	ASN
1	A	232	ASN
1	A	258	ASN
1	A	354	ASN
1	A	366	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	424	GLN
1	A	445	ASN
1	A	451	HIS
1	A	452	HIS
1	A	462	ASN
1	A	495	ASN
1	A	502	ASN
1	A	530	HIS

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 4 ligands modelled in this entry, 4 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 ⓘ

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	494/611 (80%)	-0.17	10 (2%) 65 67	20, 28, 42, 54	0

All (10) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	313	GLY	5.4
1	A	136	TRP	3.9
1	A	518	ASP	3.1
1	A	269	PRO	3.0
1	A	270	LYS	2.6
1	A	294	GLN	2.6
1	A	359	ILE	2.3
1	A	352	ASN	2.1
1	A	468	LYS	2.1
1	A	494	TYR	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 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	CA	A	1531	1/1	0.94	0.06	32,32,32,32	0
2	CA	A	1333	1/1	0.97	0.04	30,30,30,30	0
2	CA	A	1332	1/1	0.98	0.05	29,29,29,29	0
3	ZN	A	1533	1/1	0.99	0.02	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.