Full wwPDB X-ray Structure Validation Report (i)

May 26, 2020 – 07:09 am BST

PDB ID	:	1TML
Title	:	CRYSTAL STRUCTURE OF THE CATALYTIC DOMAIN OF A THERMOPHILIC ENDOCELLULASE
Authors	:	Spezio, M.; Wilson, D.B.; Karplus, P.A.
Deposited on	:	1993-06-08
Resolution	:	1.80 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.

The following versions of software and data (see references (1)) were used in the production of this report:

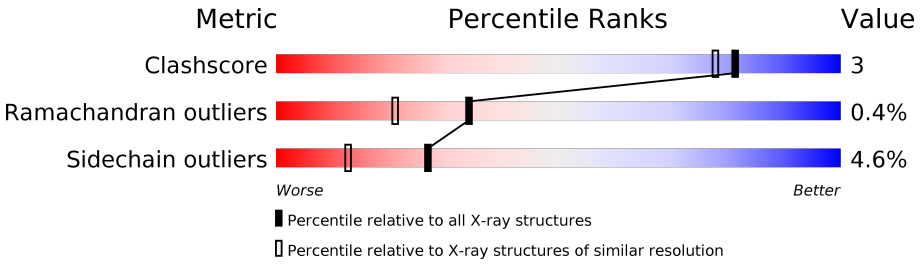
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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))
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)

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%

Mol	Chain	Length	Quality of chain
1	A	286	86% 12% •

2 Entry composition

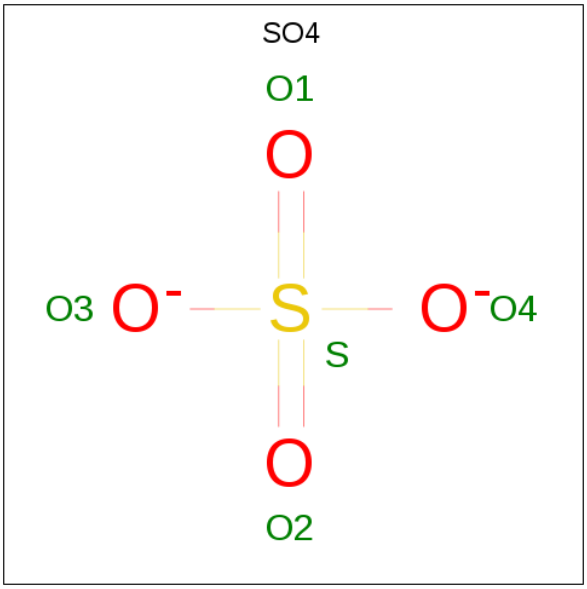
There are 3 unique types of molecules in this entry. The entry contains 2843 atoms, of which 625 are hydrogens and 0 are deuteriums.

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-1,4-BETA-D-GLUCANASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	286	Total	C	H	N	O	S	0	0	0
			2619	1332	479	383	413	12			

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 is water.

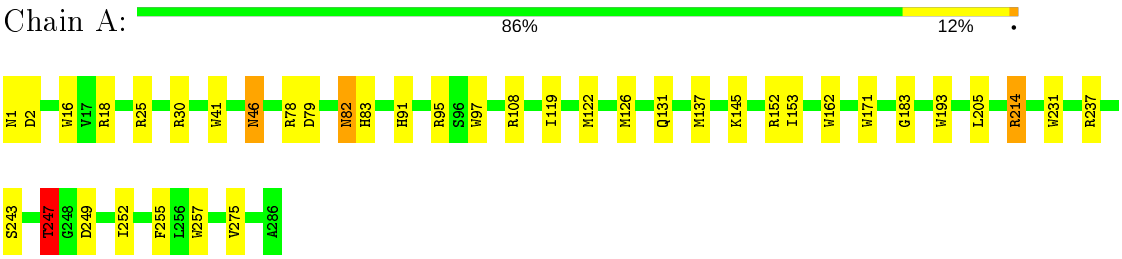
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	73	Total	H	O	0	0
			219	146	73		

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. 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. 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-1,4-BETA-D-GLUCANASE

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	43.35Å 65.94Å 43.41Å 90.00° 107.69° 90.00°	Depositor
Resolution (Å)	(Not available) – 1.80 41.36 – 2.58	Depositor EDS
% Data completeness (in resolution range)	(Not available) ((Not available)-1.80) 94.1 (41.36-2.58)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	9.99 (at 2.58Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.184 , (Not available) 0.157 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	16.1	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 30.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.031 for l,-k,h	Xtriage
F_o,F_c correlation	0.28	EDS
Total number of atoms	2843	wwPDB-VP
Average B, all atoms (Å^2)	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.44% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.80	0/2198	1.43	38/3002 (1.3%)

There are no bond length outliers.

All (38) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	18	ARG	NE-CZ-NH2	-10.93	114.83	120.30
1	A	16	TRP	CD1-CG-CD2	8.73	113.28	106.30
1	A	97	TRP	CD1-CG-CD2	8.53	113.13	106.30
1	A	171	TRP	CD1-CG-CD2	8.18	112.84	106.30
1	A	247	THR	N-CA-CB	-8.14	94.84	110.30
1	A	18	ARG	NE-CZ-NH1	8.10	124.35	120.30
1	A	152	ARG	NE-CZ-NH2	-8.07	116.26	120.30
1	A	193	TRP	CD1-CG-CD2	8.07	112.76	106.30
1	A	237	ARG	NE-CZ-NH1	7.97	124.28	120.30
1	A	231	TRP	CD1-CG-CD2	7.85	112.58	106.30
1	A	257	TRP	CD1-CG-CD2	7.46	112.27	106.30
1	A	152	ARG	NE-CZ-NH1	7.26	123.93	120.30
1	A	16	TRP	CE2-CD2-CG	-7.13	101.60	107.30
1	A	171	TRP	CE2-CD2-CG	-7.00	101.70	107.30
1	A	231	TRP	CE2-CD2-CG	-6.93	101.76	107.30
1	A	193	TRP	CE2-CD2-CG	-6.88	101.79	107.30
1	A	41	TRP	CD1-CG-CD2	6.73	111.69	106.30
1	A	162	TRP	CE2-CD2-CG	-6.45	102.14	107.30
1	A	41	TRP	CE2-CD2-CG	-6.34	102.22	107.30
1	A	97	TRP	CE2-CD2-CG	-6.25	102.30	107.30
1	A	30	ARG	NE-CZ-NH1	6.15	123.37	120.30
1	A	78	ARG	NE-CZ-NH1	6.09	123.34	120.30
1	A	97	TRP	CG-CD1-NE1	-6.05	104.05	110.10
1	A	257	TRP	CE2-CD2-CG	-6.04	102.47	107.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	108	ARG	NE-CZ-NH2	-6.02	117.29	120.30
1	A	162	TRP	CD1-CG-CD2	5.99	111.09	106.30
1	A	16	TRP	CG-CD1-NE1	-5.91	104.19	110.10
1	A	231	TRP	CG-CD1-NE1	-5.83	104.27	110.10
1	A	78	ARG	NE-CZ-NH2	-5.78	117.41	120.30
1	A	231	TRP	CG-CD2-CE3	5.72	139.05	133.90
1	A	231	TRP	CB-CG-CD1	-5.60	119.72	127.00
1	A	162	TRP	CG-CD2-CE3	5.56	138.91	133.90
1	A	171	TRP	CG-CD1-NE1	-5.45	104.65	110.10
1	A	193	TRP	CG-CD1-NE1	-5.40	104.70	110.10
1	A	25	ARG	NE-CZ-NH2	-5.27	117.66	120.30
1	A	30	ARG	NE-CZ-NH2	-5.26	117.67	120.30
1	A	214	ARG	NE-CZ-NH2	-5.22	117.69	120.30
1	A	16	TRP	CB-CG-CD1	-5.12	120.34	127.00

There are no chirality outliers.

There are no planarity outliers.

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	2140	479	2029	11	0
2	A	5	0	0	1	0
3	A	73	146	0	0	0
All	All	2218	625	2029	11	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 3.

All (11) 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:119:ILE:HD11	1:A:137:MET:HG3	1.73	0.69
1:A:247:THR:HG21	1:A:252:ILE:O	1.93	0.68
1:A:46:ASN:H	1:A:46:ASN:HD22	1.48	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:247:THR:HG22	1:A:249:ASP:H	1.67	0.59
1:A:79:ASP:HB3	1:A:82:ASN:HD21	1.77	0.50
1:A:145:LYS:HG3	1:A:153:ILE:CD1	2.44	0.47
1:A:183:GLY:HA3	1:A:214:ARG:O	2.17	0.44
1:A:91:HIS:HB3	1:A:95:ARG:NH1	2.34	0.42
1:A:79:ASP:HB2	2:A:360:SO4:O1	2.21	0.41
1:A:79:ASP:HB3	1:A:82:ASN:ND2	2.35	0.41
1:A:243:SER:HA	1:A:255:PHE:O	2.21	0.41

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	284/286 (99%)	274 (96%)	9 (3%)	1 (0%)	34 21

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	83	HIS

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	219/219 (100%)	209 (95%)	10 (5%)	27 13

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	ASN
1	A	2	ASP
1	A	46	ASN
1	A	82	ASN
1	A	122	MET
1	A	126	MET
1	A	131	GLN
1	A	205	LEU
1	A	247	THR
1	A	275	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1	ASN
1	A	46	ASN
1	A	49	GLN
1	A	53	GLN
1	A	82	ASN
1	A	83	HIS
1	A	107	ASN
1	A	131	GLN
1	A	273	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	SO4	A	360	-	4,4,4	0.87	0	6,6,6	0.42	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	360	SO4	1	0

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

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.