Full wwPDB X-ray Structure Validation Report ⓘ

Nov 2, 2024 – 02:37 PM EDT

PDB ID	:	1WZZ
Title	:	Structure of endo-beta-1,4-glucanase CMCax from Acetobacter xylinum
Authors	:	Yasutake, Y.; Kawano, S.; Tajima, K.; Yao, M.; Satoh, Y.; Munekata, M.; Tanaka, I.; Structural Genomics Consortium (SGC)
Deposited on	:	2005-03-10
Resolution	:	1.65 Å(reported)

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

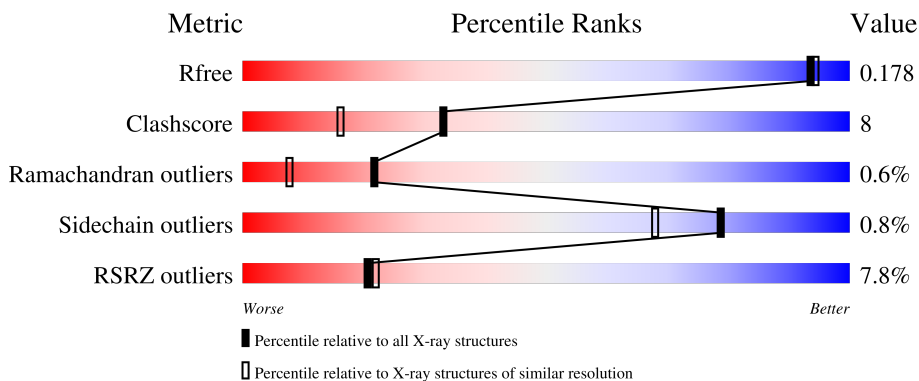|		
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 1.65 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	Percentile Ranks	Value
Rfree		0.178
Clashscore		8
Ramachandran outliers		0.6%
Sidechain outliers		0.8%
RSRZ outliers		7.8%

Worse   *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	2328 (1.66-1.66)
Clashscore	180529	2515 (1.66-1.66)
Ramachandran outliers	177936	2475 (1.66-1.66)
Sidechain outliers	177891	2475 (1.66-1.66)
RSRZ outliers	164620	2328 (1.66-1.66)

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 <=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	334	7% ... 84% ... 11% •

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 2808 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 Probable endoglucanase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	319	Total 2491	C 1605	N 429	O 446	S 1	Se 10	0	0	0

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	9	MET	-	expression tag	UNP P37696
A	10	ARG	-	expression tag	UNP P37696
A	11	GLY	-	expression tag	UNP P37696
A	12	SER	-	expression tag	UNP P37696
A	13	HIS	-	expression tag	UNP P37696
A	14	HIS	-	expression tag	UNP P37696
A	15	HIS	-	expression tag	UNP P37696
A	16	HIS	-	expression tag	UNP P37696
A	17	HIS	-	expression tag	UNP P37696
A	18	HIS	-	expression tag	UNP P37696
A	19	GLY	-	expression tag	UNP P37696
A	20	SER	-	expression tag	UNP P37696
A	63	MSE	MET	modified residue	UNP P37696
A	78	MSE	MET	modified residue	UNP P37696
A	80	MSE	MET	modified residue	UNP P37696
A	140	MSE	MET	modified residue	UNP P37696
A	150	MSE	MET	modified residue	UNP P37696
A	152	MSE	MET	modified residue	UNP P37696
A	161	MSE	MET	modified residue	UNP P37696
A	182	MSE	MET	modified residue	UNP P37696
A	201	MSE	MET	modified residue	UNP P37696
A	258	MSE	MET	modified residue	UNP P37696

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

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

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	312	Total 312	O 312	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: Probable endoglucanase
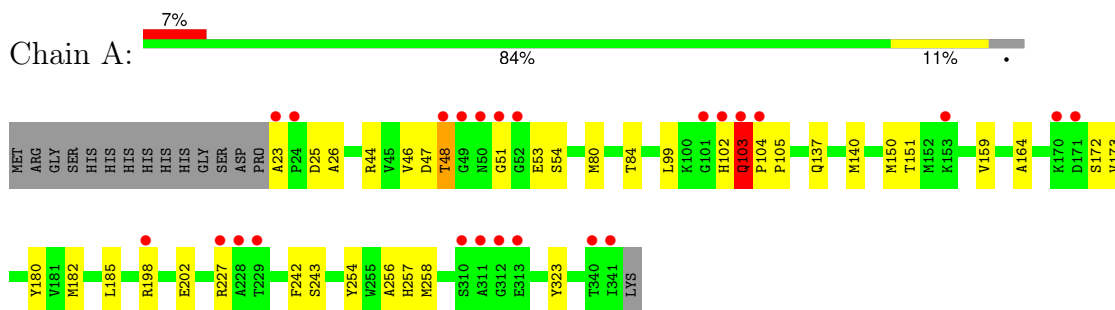
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	88.88Å 88.88Å 93.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	40.00 – 1.65 40.00 – 1.65	Depositor EDS
% Data completeness (in resolution range)	100.0 (40.00-1.65) 100.0 (40.00-1.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.07	Depositor
$< I/\sigma(I) >$ [1]	4.31 (at 1.65Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.176 , 0.195 0.177 , 0.178	Depositor DCC
R_{free} test set	2564 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	17.1	Xtriage
Anisotropy	0.250	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 38.0	EDS
L-test for twinning[2]	$< |L| > = 0.49, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.039 for h,-h-k,-l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2808	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.72% 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 (i)

5.1 Standard geometry (i)

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	0.57	0/2558	0.77	2/3477 (0.1%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	242	PHE	N-CA-C	-5.06	97.35	111.00
1	A	103	GLN	C-N-CD	5.00	138.91	128.40

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	2491	0	2419	38	0
2	A	5	0	0	0	0
3	A	312	0	0	8	0
All	All	2808	0	2419	38	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 8.

All (38) 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:46:VAL:HG13	1:A:51:GLY:HA2	1.64	0.80
1:A:180:TYR:HB3	1:A:182:MSE:HE3	1.65	0.78
1:A:180:TYR:HB3	1:A:182:MSE:CE	2.14	0.78
1:A:140:MSE:HE1	3:A:673:HOH:O	1.92	0.69
1:A:46:VAL:CG1	1:A:51:GLY:HA2	2.22	0.69
1:A:80:MSE:HG2	3:A:675:HOH:O	1.98	0.63
1:A:227:ARG:HD2	3:A:704:HOH:O	1.98	0.63
1:A:150:MSE:HG3	3:A:666:HOH:O	1.99	0.62
1:A:137:GLN:HA	1:A:140:MSE:CE	2.29	0.62
1:A:103:GLN:HB2	1:A:104:PRO:HD3	1.82	0.60
1:A:103:GLN:CB	1:A:104:PRO:HD3	2.31	0.60
1:A:172:SER:HA	1:A:227:ARG:HG3	1.83	0.60
1:A:137:GLN:HA	1:A:140:MSE:HE2	1.83	0.60
1:A:182:MSE:HE1	1:A:185:LEU:HD12	1.85	0.58
1:A:23:ALA:HB3	1:A:26:ALA:HB2	1.88	0.56
1:A:182:MSE:CE	1:A:185:LEU:HD12	2.36	0.55
1:A:150:MSE:HE1	1:A:164:ALA:HB2	1.88	0.55
1:A:53:GLU:HG2	1:A:99:LEU:HD23	1.88	0.54
1:A:182:MSE:HB3	1:A:256:ALA:HB2	1.89	0.54
1:A:44:ARG:CG	1:A:54:SER:OG	2.57	0.53
1:A:103:GLN:OE1	1:A:103:GLN:HA	2.10	0.52
1:A:140:MSE:HE3	3:A:711:HOH:O	2.11	0.51
1:A:80:MSE:HE2	3:A:547:HOH:O	2.11	0.50
1:A:182:MSE:HB3	1:A:256:ALA:CB	2.44	0.47
1:A:103:GLN:CB	1:A:104:PRO:CD	2.92	0.46
1:A:103:GLN:HB2	1:A:104:PRO:CD	2.45	0.46
1:A:25:ASP:HA	3:A:546:HOH:O	2.18	0.43
1:A:47:ASP:HB2	1:A:323:TYR:CE1	2.53	0.43
1:A:105:PRO:HD2	3:A:610:HOH:O	2.18	0.43
1:A:257:HIS:O	1:A:258:MSE:HE2	2.19	0.43
1:A:103:GLN:O	1:A:105:PRO:HD3	2.18	0.42
1:A:48:THR:HG22	1:A:51:GLY:H	1.84	0.42
1:A:172:SER:CA	1:A:227:ARG:HG3	2.48	0.42
1:A:102:HIS:O	1:A:103:GLN:C	2.57	0.41
1:A:198:ARG:O	1:A:202:GLU:HG3	2.20	0.41
1:A:173:VAL:HG13	1:A:227:ARG:HD3	2.01	0.41
1:A:151:THR:HA	1:A:159:VAL:O	2.21	0.41
1:A:80:MSE:O	1:A:84:THR:HG23	2.22	0.40

There are no symmetry-related clashes.

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	317/334 (95%)	310 (98%)	5 (2%)	2 (1%)	22　8

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103	GLN
1	A	243	SER

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	251/254 (99%)	249 (99%)	2 (1%)	79　68

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	48	THR
1	A	254	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	50	ASN
1	A	76	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	131	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 ⓘ

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	401	-	4,4,4	0.38	0	6,6,6	0.20	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.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	309/334 (92%)	0.01	24 (7%)	20	22	10, 17, 42, 57	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	312	GLY	6.3
1	A	311	ALA	6.2
1	A	341	ILE	6.1
1	A	50	ASN	5.0
1	A	103	GLN	4.5
1	A	24	PRO	4.3
1	A	23	ALA	4.3
1	A	104	PRO	4.2
1	A	102	HIS	3.8
1	A	49	GLY	3.6
1	A	101	GLY	3.4
1	A	227	ARG	3.4
1	A	313	GLU	3.4
1	A	198	ARG	3.2
1	A	48	THR	3.1
1	A	170	LYS	3.0
1	A	153	LYS	2.7
1	A	228	ALA	2.7
1	A	171	ASP	2.6
1	A	340	THR	2.5
1	A	229	THR	2.4
1	A	310	SER	2.2
1	A	52	GLY	2.1
1	A	51	GLY	2.1

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)

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(Å^2)	Q<0.9
2	SO4	A	401	5/5	0.93	0.11	42,43,45,45	0

6.5 Other polymers (i)

There are no such residues in this entry.