W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Dec 19, 2023 – 06:55 AM EST

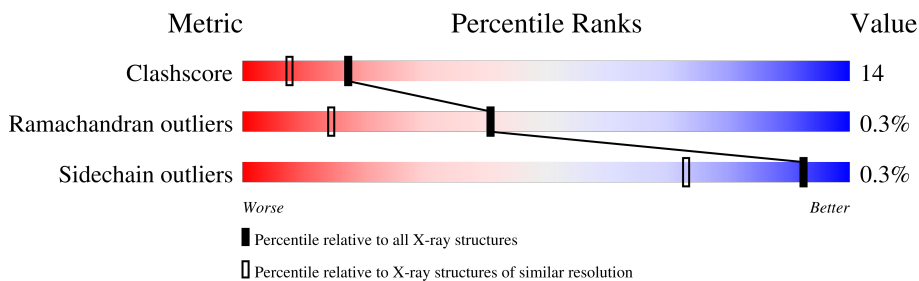|		
PDB ID	:	1IS9
Title	:	Endoglucanase A from Clostridium thermocellum at atomic resolution
Authors	:	Schmidt, A.; Gonzalez, A.; Morris, R.J.; Costabel, M.; Alzari, P.M.; Lamzin, V.S.
Deposited on	:	2001-11-26
Resolution	:	1.03 Å(reported)

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 1.03 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
Clashscore		14
Ramachandran outliers		0.3%
Sidechain outliers		0.3%

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))
Clashscore	141614	1677 (1.10-0.98)
Ramachandran outliers	138981	1591 (1.10-0.98)
Sidechain outliers	138945	1589 (1.10-0.98)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	363	85% 14% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	CL	A	1010	-	-	X	-
3	CL	A	1014	-	-	X	-

2 Entry composition ⓘ

There are 4 unique types of molecules in this entry. The entry contains 3600 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 A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	358	Total 2827	C 1790	N 476	O 550	S 11	0	7	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	135	GLN	LEU	conflict	UNP P04955

- Molecule 2 is MERCURY (II) ION (three-letter code: HG) (formula: Hg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	6	Total 6	Hg 6	0	0

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	7	Total 7	Cl 7	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	760	Total 760	O 760	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. 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.

Note EDS was not executed.

• Molecule 1: endoglucanase A
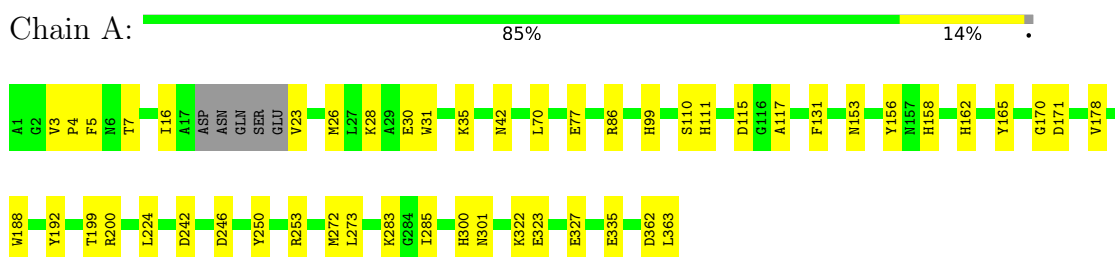
4 Data and refinement statistics ⓘ

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.44Å 62.83Å 103.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 1.03	Depositor
% Data completeness (in resolution range)	98.0 (25.00-1.03)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
Refinement program	REFMAC	Depositor
R, R_{free}	0.131 , 0.149	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3600	wwPDB-VP
Average B, all atoms (Å^2)	10.0	wwPDB-VP

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: CL, HG

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	0.38	0/2940	0.92	8/3994 (0.2%)

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	2

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	86	ARG	NE-CZ-NH1	11.51	126.05	120.30
1	A	192	TYR	CB-CG-CD1	7.20	125.32	121.00
1	A	86	ARG	NE-CZ-NH2	-6.44	117.08	120.30
1	A	192	TYR	CB-CG-CD2	-5.98	117.41	121.00
1	A	200	ARG	NE-CZ-NH2	-5.21	117.69	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	3	VAL	Peptide,Mainchain

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	2827	0	2638	72	1
2	A	6	0	0	0	0
3	A	7	0	0	7	0
4	A	760	0	0	55	9
All	All	3600	0	2638	72	9

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 14.

The worst 5 of 72 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:131:PHE:HD1	4:A:1722:HOH:O	1.02	1.33
1:A:110:SER:CB	4:A:1725:HOH:O	1.74	1.30
1:A:171:ASP:HA	4:A:1738:HOH:O	1.32	1.27
1:A:283:LYS:HG2	4:A:1717:HOH:O	1.28	1.21
1:A:131:PHE:CD1	4:A:1722:HOH:O	1.79	1.16

The worst 5 of 9 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)
1:A:283:LYS:CD	4:A:1698:HOH:O[1_655]	1.77	0.43
4:A:1417:HOH:O	4:A:1569:HOH:O[4_445]	1.90	0.30
4:A:1643:HOH:O	4:A:1736:HOH:O[4_545]	1.95	0.25
4:A:1672:HOH:O	4:A:1754:HOH:O[3_645]	1.95	0.25
4:A:1094:HOH:O	4:A:1492:HOH:O[4_555]	2.08	0.12

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	361/363 (99%)	352 (98%)	8 (2%)	1 (0%)	41	14

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	178	VAL

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	293/292 (100%)	292 (100%)	1 (0%)	92	73

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	188	TRP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	162	HIS
1	A	203	GLN
1	A	300	HIS
1	A	153	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	154	ASN

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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 13 ligands modelled in this entry, 13 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 ⓘ

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 ⓘ

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.