



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:37 AM GMT

PDB ID : 1OCJ
Title : Mutant D416A of the CELLOBIOHYDROLASE CEL6A FROM HUMICOLA
INSOLENS in complex with a THIOPENTASACCHARIDE at 1.3 angstrom
resolution
Authors : Varrot, A.; Frandsen, T.P.; Von Ossowski, I.; Boyer, V.; Driguez, H.; Schulein,
M.; Davies, G.J.
Deposited on : 2003-02-07
Resolution : 1.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

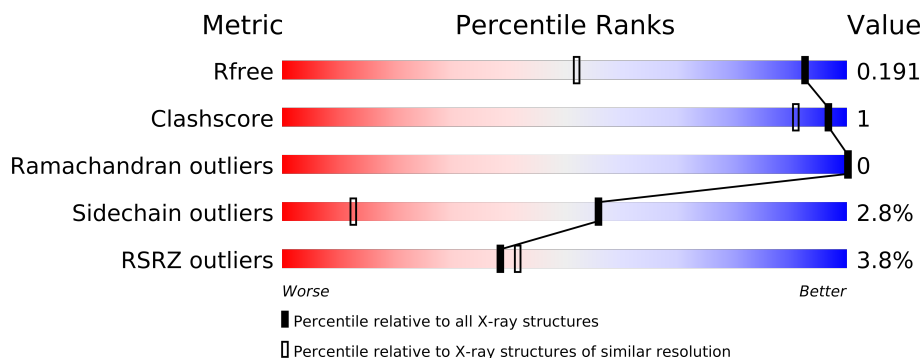
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.30 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}	66092	1025 (1.34-1.26)
Clashscore	79885	1140 (1.34-1.26)
Ramachandran outliers	78287	1093 (1.34-1.26)
Sidechain outliers	78261	1092 (1.34-1.26)
RSRZ outliers	66119	1025 (1.34-1.26)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	362	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	NAG	A	500	-	X
6	MA3	A	505	-	X
7	MG	A	506	-	X
8	ACY	A	507	-	X

2 Entry composition i

There are 9 unique types of molecules in this entry. The entry contains 3331 atoms, of which 0 are hydrogen and 0 are deuterium.

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 CELLOBIOHYDROLASE II.

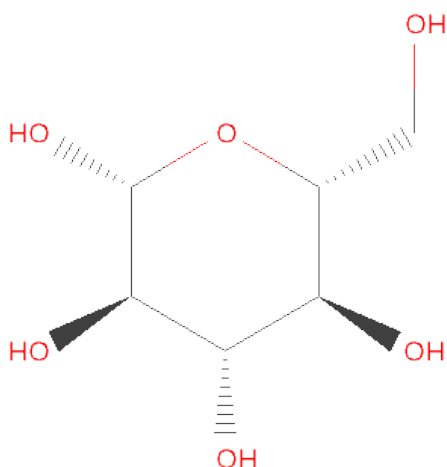
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	360	2835	1804	490	530	11	0	8	0

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



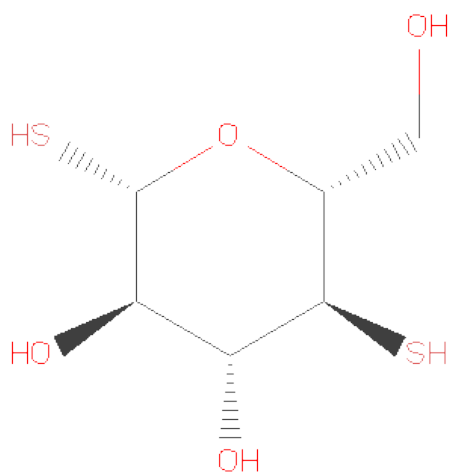
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	14	8	1	5	0	0

- Molecule 3 is SUGAR (BETA-D-GLUCOSE) (three-letter code: BGC) (formula: $C_6H_{12}O_6$).



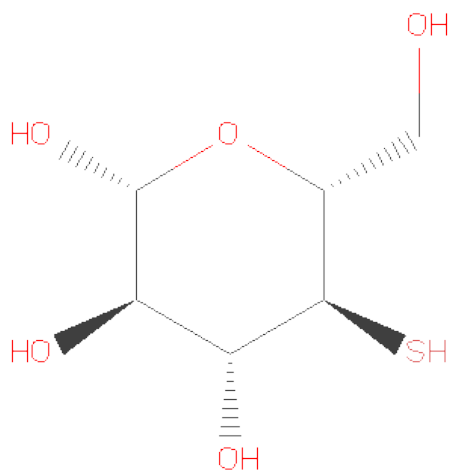
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is SUGAR (1,4-DEOXY-1,4-DITHIO-BETA-D-GLUCOPYRANOSE) (three-letter code: SSG) (formula: $C_6H_{12}O_4S_2$).



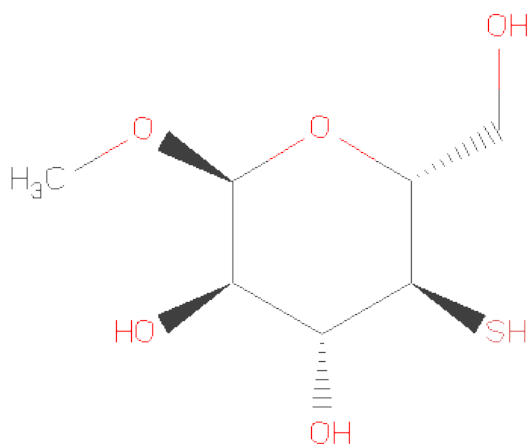
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	O	S	0	0
			11	6	4	1		
4	A	1	Total	C	O	S	0	0
			11	6	4	1		

- Molecule 5 is SUGAR (4-DEOXY-4-THIO-BETA-D-GLUCOPYRANOSE) (three-letter code: SGC) (formula: $C_6H_{12}O_5S$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
5	A	1	11	6	4	1	0	0

- Molecule 6 is SUGAR (O1-METHYL-4-DEOXY-4-THIO-ALPHA-D-GLUCOSE) (three-letter code: MA3) (formula: $C_7H_{14}O_5S$).

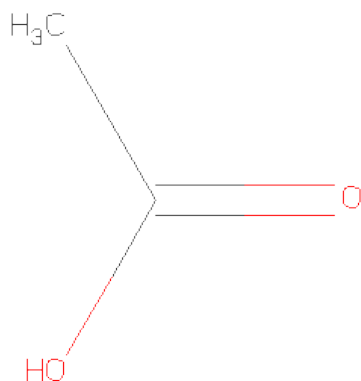


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
6	A	1	13	7	5	1	0	0

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	1	Total	Mg	0	0
			1	1		

- Molecule 8 is ACETIC ACID (three-letter code: ACY) (formula: C₂H₄O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	402	Total	O	0	18
			420	420		

i

- Molecule 1: CELLOBIOHYDROLASE II

Sequence logo showing the enrichment of amino acids in the 1000bp upstream region of the TLR4 gene. The y-axis represents information content in bits (0.00 to 0.25). The x-axis shows amino acid positions from TYR to F450. Red dots above the bars indicate positions with significant enrichment (p < 0.05).

Position	Enriched Amino Acid(s)
Y104	Tyr
Q116	Gln
T117	Thr
T118	Thr
D119	Asp
P120	Pro
A121	Ala
L122	Leu
R123	Arg
E157	Glu
D176	Asp
A184	Ala
R208	Arg
T212	Thr
M231	Met
M235	Met
D268	Asp
T281	Thr
R299	Arg
S317	Ser
S318	Ser
F338	Phe
L341	Leu
R382	Arg
Q390	Gln
D409	Asp
T410	Thr
T411	Thr
D423	Asp
K426	Lys
P427	Pro
F435	Phe
F450	Phe

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	47.44Å 67.55Å 53.66Å 90.00° 110.87° 90.00°	Depositor
Resolution (Å)	30.00 – 1.30 28.01 – 1.30	Depositor EDS
% Data completeness (in resolution range)	91.0 (30.00-1.30) 91.0 (28.01-1.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.48 (at 1.30Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.145 , 0.174 0.161 , 0.191	Depositor DCC
R_{free} test set	3548 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	11.6	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 35.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 132463 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3331	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.78% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, BGC, NAG, SSG, MA3, ACY, SGC

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	0.72	0/2944	0.84	6/4027 (0.1%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	268	ASP	CB-CG-OD2	5.95	123.65	118.30
1	A	235	MET	CG-SD-CE	-5.76	90.99	100.20
1	A	299	ARG	NE-CZ-NH2	-5.75	117.43	120.30
1	A	423	ASP	CB-CG-OD2	5.35	123.11	118.30
1	A	175	ASP	CB-CG-OD1	5.21	122.99	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2835	0	0	2	0
2	A	14	0	0	0	0
3	A	11	0	0	0	0
4	A	22	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	11	0	0	0	0
6	A	13	0	0	0	0
7	A	1	0	0	0	0
8	A	4	0	0	0	0
9	A	420	0	0	0	0
All	All	3331	0	0	2	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 1.

All (2) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:281:ILE:CG2	1:A:341[B]:LEU:CD1	2.83	0.56
1:A:208:ARG:NH1	1:A:212:ILE:CD1	2.69	0.56

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	366/362 (101%)	355 (97%)	11 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	294/288 (102%)	286 (97%)	8 (3%)	57	13

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	281	ILE
1	A	435	PHE
1	A	390	GLN
1	A	231	MET
1	A	382	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 8 ligands modelled in this entry, 1 is monoatomic - leaving 7 for Mogul analysis.

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	NAG	A	500	1	12,14,15	0.71	0	15,19,21	0.75	0
3	BGC	A	501	4	10,11,12	0.65	0	11,15,17	1.21	2 (18%)
4	SSG	A	502	3,4	10,11,12	1.87	4 (40%)	11,15,17	2.60	7 (63%)
4	SSG	A	503	5,4	10,11,12	1.79	2 (20%)	11,15,17	2.60	6 (54%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
5	SGC	A	504	4,6	10,11,12	1.30	1 (10%)	11,15,17	2.17	5 (45%)
6	MA3	A	505	5	13,13,13	1.56	2 (15%)	18,18,18	2.73	8 (44%)
8	ACY	A	507	-	3,3,3	0.37	0	3,3,3	1.00	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	500	1	-	0/6/23/26	0/1/1/1
3	BGC	A	501	4	-	0/2/19/22	0/1/1/1
4	SSG	A	502	3,4	-	0/2/19/22	0/1/1/1
4	SSG	A	503	5,4	-	0/2/19/22	0/1/1/1
5	SGC	A	504	4,6	-	0/2/19/22	0/1/1/1
6	MA3	A	505	5	-	0/4/24/24	0/1/1/1
8	ACY	A	507	-	-	0/0/0/0	0/0/0/0

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	505	MA3	O1-C1	4.45	1.47	1.40
4	A	503	SSG	C4-S4	-3.72	1.77	1.82
4	A	502	SSG	C5-C4	-3.03	1.50	1.53
4	A	503	SSG	O5-C5	-3.03	1.39	1.45
6	A	505	MA3	C4-S4	-2.70	1.78	1.82

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	505	MA3	O1-C1-C2	-6.96	99.64	108.18
4	A	503	SSG	O3-C3-C4	4.64	118.50	108.95
4	A	502	SSG	C4-C3-C2	4.38	116.92	108.41
4	A	502	SSG	O5-C5-C6	4.21	111.40	106.98
4	A	503	SSG	C3-C4-C5	3.76	117.32	110.18

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

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	360/362 (99%)	0.19	14 (3%) 37 40	7, 12, 26, 33	10 (2%)

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	450	PHE	3.8
1	A	411	THR	3.0
1	A	104	TYR	2.9
1	A	317	SER	2.6
1	A	121	ALA	2.5

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

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

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
8	ACY	A	507	4/4	0.25	6.20	17,18,18,19	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
7	MG	A	506	1/1	0.22	4.91	23,23,23,23	0
6	MA3	A	505	13/13	0.09	2.82	11,13,17,23	0
2	NAG	A	500	14/15	0.07	2.04	9,11,15,17	0
4	SSG	A	503	11/12	0.11	1.88	15,19,21,27	0
4	SSG	A	502	11/12	0.10	0.98	13,16,18,22	0
3	BGC	A	501	11/12	0.08	0.02	12,13,15,15	0
5	SGC	A	504	11/12	0.07	-0.20	11,12,13,17	0

6.5 Other polymers ⓘ

There are no such residues in this entry.