Full wwPDB X-ray Structure Validation Report (i)

Sep 13, 2023 – 03:27 PM EDT

PDB ID	:	4QB1
Title	:	Structure of CBM35 from Paenibacillus barcinonensis
Authors	:	Sainz-Polo, M.A.; Sanz-Aparicio, J.
Deposited on	:	2014-05-06
Resolution	:	2.19 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

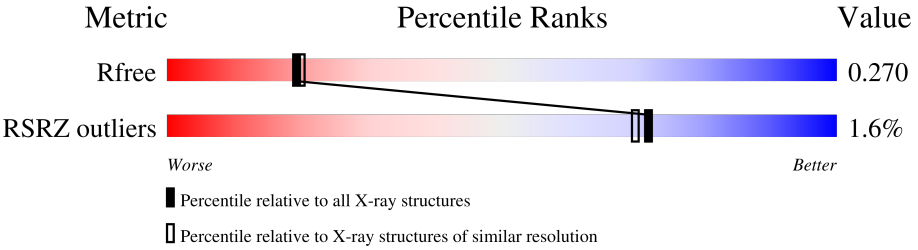
MolProbity	:	FAILED
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.19 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}	130704	4898 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1036 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 Xyn30D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	128	974	605	159	207	3	0	0	0

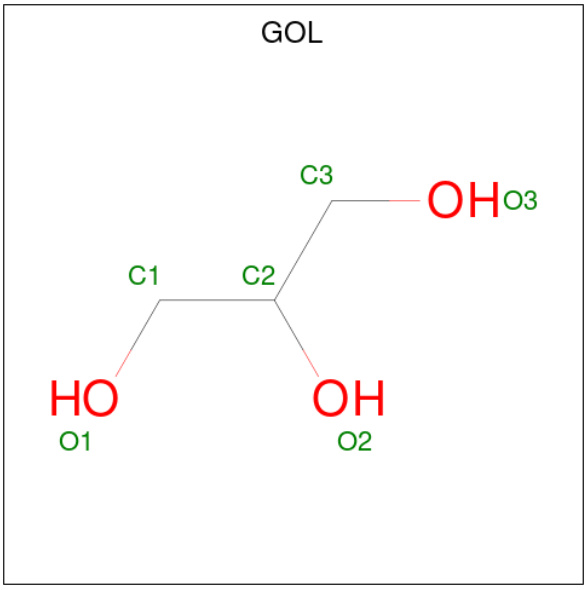
There are 23 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-22	MET	-	initiating methionine	UNP H6WCZ0
A	-21	GLY	-	expression tag	UNP H6WCZ0
A	-20	SER	-	expression tag	UNP H6WCZ0
A	-19	SER	-	expression tag	UNP H6WCZ0
A	-18	HIS	-	expression tag	UNP H6WCZ0
A	-17	HIS	-	expression tag	UNP H6WCZ0
A	-16	HIS	-	expression tag	UNP H6WCZ0
A	-15	HIS	-	expression tag	UNP H6WCZ0
A	-14	HIS	-	expression tag	UNP H6WCZ0
A	-13	HIS	-	expression tag	UNP H6WCZ0
A	-12	SER	-	expression tag	UNP H6WCZ0
A	-11	SER	-	expression tag	UNP H6WCZ0
A	-10	GLY	-	expression tag	UNP H6WCZ0
A	-9	LEU	-	expression tag	UNP H6WCZ0
A	-8	VAL	-	expression tag	UNP H6WCZ0
A	-7	PRO	-	expression tag	UNP H6WCZ0
A	-6	ARG	-	expression tag	UNP H6WCZ0
A	-5	GLY	-	expression tag	UNP H6WCZ0
A	-4	HIS	-	expression tag	UNP H6WCZ0
A	-3	SER	-	expression tag	UNP H6WCZ0
A	-2	MET	-	expression tag	UNP H6WCZ0
A	-1	ALA	-	expression tag	UNP H6WCZ0
A	0	SER	-	expression tag	UNP H6WCZ0

- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	54	54	54	0	0

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	39.40Å 47.45Å 103.02Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.51 – 2.19 43.10 – 2.19	Depositor EDS
% Data completeness (in resolution range)	98.6 (51.51-2.19) 98.6 (43.10-2.19)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.02	Depositor
$< I/\sigma(I) >$ [1]	8.82 (at 2.20Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.209 , 0.269 0.218 , 0.270	Depositor DCC
R_{free} test set	494 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å^2)	17.8	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 49.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	1036	wwPDB-VP
Average B, all atoms (Å^2)	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.14% 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.

4 Model quality (i)

4.1 Standard geometry (i)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts (i)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles (i)

4.3.1 Protein backbone (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains (i)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA (i)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains (i)

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

4.5 Carbohydrates (i)

There are no monosaccharides in this entry.

4.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
3	GOL	A	203	2	5,5,5	0.68	0	5,5,5	0.42	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
3	GOL	A	203	2	-	3/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	203	GOL	O2-C2-C3-O3
3	A	203	GOL	C1-C2-C3-O3
3	A	203	GOL	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

4.7 Other polymers (i)

There are no such residues in this entry.

4.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

5 Fit of model and data (i)

5.1 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	128/164 (78%)	0.66	2 (1%) 72 70	13, 16, 23, 33	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	49	THR	2.4
1	A	141	SER	2.0

5.2 Non-standard residues in protein, DNA, RNA chains (i)

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

5.3 Carbohydrates (i)

There are no monosaccharides in this entry.

5.4 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
3	GOL	A	203	6/6	0.86	0.19	16,17,18,18	0
2	CA	A	202	1/1	0.98	0.09	16,16,16,16	0
2	CA	A	201	1/1	0.99	0.13	14,14,14,14	0

5.5 Other polymers (i)

There are no such residues in this entry.