



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Aug 13, 2018 – 03:43 PM EDT

PDB ID : 6CVM
EMDB ID: : EMD-7770
Title : Atomic resolution cryo-EM structure of beta-galactosidase
Authors : Subramaniam, S.; Bartesaghi, A.; Banerjee, S.; Zhu, X.; Milne, J.L.S.
Deposited on : 2018-03-28
Resolution : 1.90 Å(reported)

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.7.3 (157068), CSD as539be (2018)
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

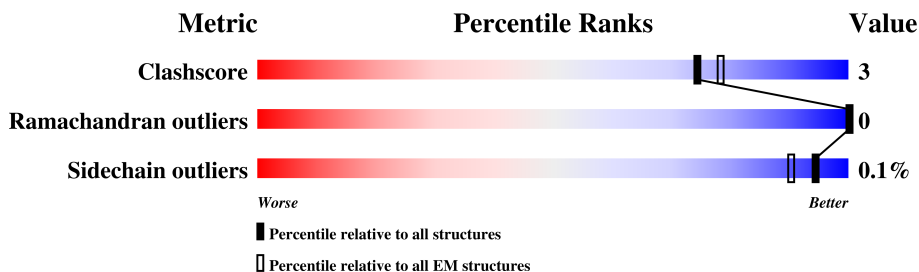
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 1.90 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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1021	91% 9%
1	B	1021	90% 9%
1	C	1021	91% 9%
1	D	1021	90% 9%

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 37062 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, 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.

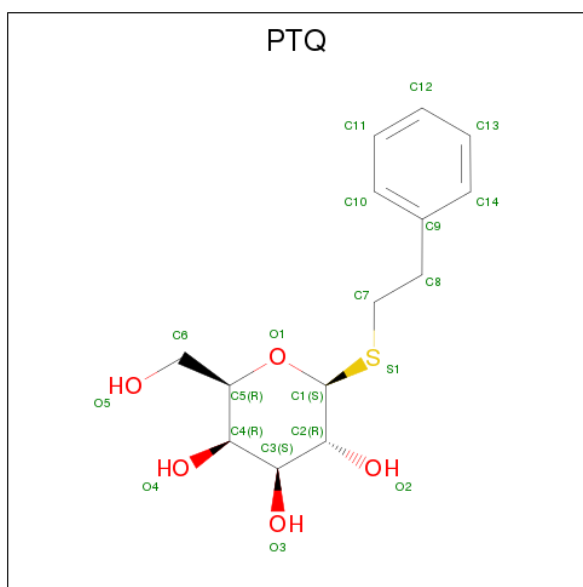
- Molecule 1 is a protein called Beta-galactosidase.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	1021	Total	C	N	O	S	0	0
			8193	5184	1450	1520	39		
1	B	1021	Total	C	N	O	S	0	0
			8193	5184	1450	1520	39		
1	C	1021	Total	C	N	O	S	0	0
			8193	5184	1450	1520	39		
1	D	1021	Total	C	N	O	S	0	0
			8193	5184	1450	1520	39		

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	334	VAL	GLU	conflict	UNP P00722
A	871	VAL	GLU	conflict	UNP P00722
B	334	VAL	GLU	conflict	UNP P00722
B	871	VAL	GLU	conflict	UNP P00722
C	334	VAL	GLU	conflict	UNP P00722
C	871	VAL	GLU	conflict	UNP P00722
D	334	VAL	GLU	conflict	UNP P00722
D	871	VAL	GLU	conflict	UNP P00722

- Molecule 2 is 2-phenylethyl 1-thio-beta-D-galactopyranoside (three-letter code: PTQ) (formula: C₁₄H₂₀O₅S).



Mol	Chain	Residues	Atoms				AltConf
2	A	1	Total	C	O	S	0
			20	14	5	1	
2	B	1	Total	C	O	S	0
			20	14	5	1	
2	C	1	Total	C	O	S	0
			20	14	5	1	
2	D	1	Total	C	O	S	0
			20	14	5	1	

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
3	B	2	Total	Mg	0
			2	2	
3	A	2	Total	Mg	0
			2	2	
3	D	2	Total	Mg	0
			2	2	
3	C	2	Total	Mg	0
			2	2	

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		AltConf
4	B	2	Total	Na	0
			2	2	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf
4	A	2	Total 2	Na 2	0
4	D	2	Total 2	Na 2	0
4	C	2	Total 2	Na 2	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		AltConf
5	A	1047	Total 1047	O 1047	0
5	B	1048	Total 1048	O 1048	0
5	C	1050	Total 1050	O 1050	0
5	D	1049	Total 1049	O 1049	0

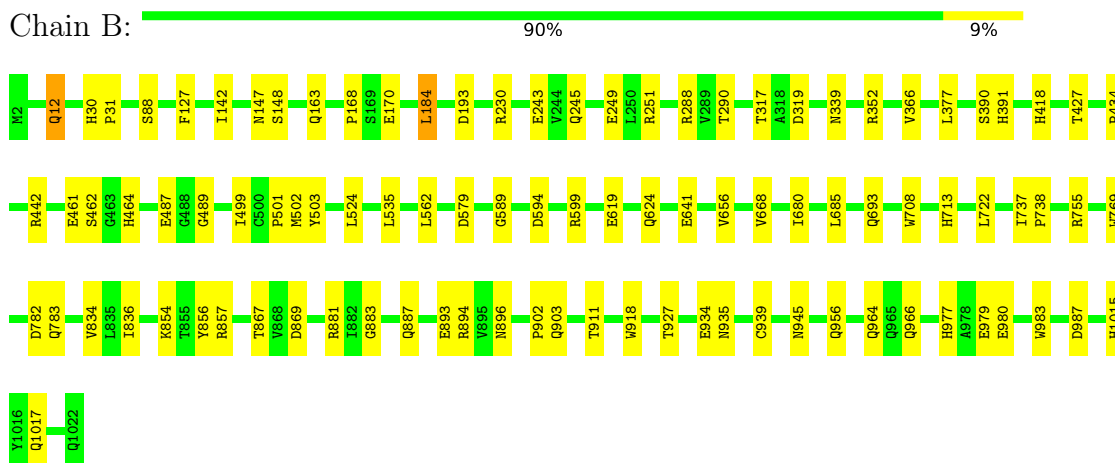
3 Residue-property plots

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.

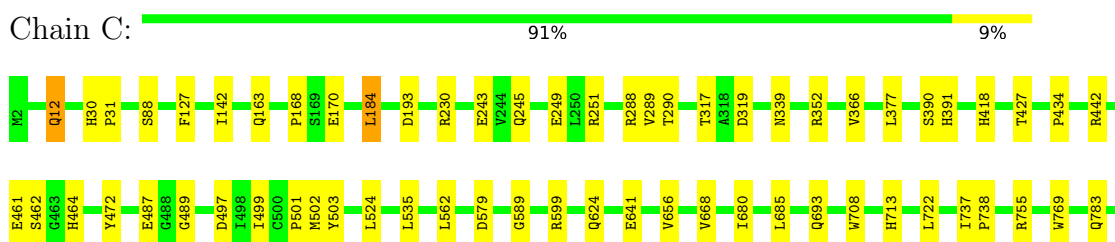
- Molecule 1: Beta-galactosidase

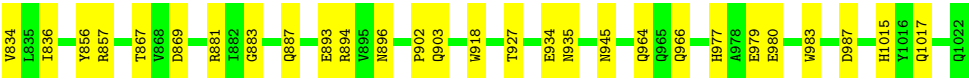


- Molecule 1: Beta-galactosidase

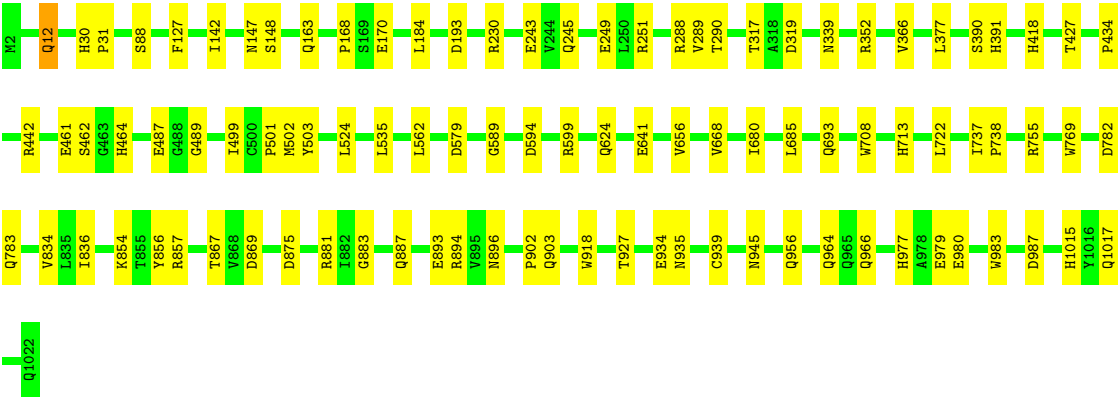
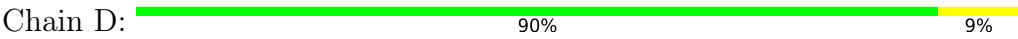


- Molecule 1: Beta-galactosidase





● Molecule 1: Beta-galactosidase



4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, D2	Depositor
Number of particles used	150321	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	45	Depositor
Minimum defocus (nm)	600	Depositor
Maximum defocus (nm)	2000	Depositor
Magnification	215000	Depositor
Image detector	GATAN K2 QUANTUM (4k x 4k)	Depositor

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: NA, MG, PTQ

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	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.54	1/8435 (0.0%)	0.57	1/11511 (0.0%)
1	B	0.55	0/8435	0.57	1/11511 (0.0%)
1	C	0.55	1/8435 (0.0%)	0.57	1/11511 (0.0%)
1	D	0.55	1/8435 (0.0%)	0.57	0/11511
All	All	0.55	3/33740 (0.0%)	0.57	3/46044 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	289	VAL	CB-CG1	-5.08	1.42	1.52
1	C	289	VAL	CB-CG1	-5.03	1.42	1.52
1	A	289	VAL	CB-CG1	-5.02	1.42	1.52

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	184	LEU	CA-CB-CG	5.01	126.82	115.30
1	B	184	LEU	CA-CB-CG	5.00	126.81	115.30
1	C	184	LEU	CA-CB-CG	5.00	126.81	115.30

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	8193	0	7798	53	0
1	B	8193	0	7798	57	0
1	C	8193	0	7798	54	0
1	D	8193	0	7798	57	0
2	A	20	0	20	1	0
2	B	20	0	20	2	0
2	C	20	0	20	1	0
2	D	20	0	20	2	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	2	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
5	A	1047	0	0	7	0
5	B	1048	0	0	8	0
5	C	1050	0	0	7	0
5	D	1049	0	0	8	0
All	All	37062	0	31272	221	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.

The worst 5 of 221 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:887:GLN:NE2	1:A:980:GLU:O	2.35	0.59
1:C:887:GLN:NE2	1:C:980:GLU:O	2.35	0.59
1:D:887:GLN:NE2	1:D:980:GLU:O	2.35	0.59
1:B:887:GLN:NE2	1:B:980:GLU:O	2.35	0.59
1:B:903:GLN:NE2	5:B:2129:HOH:O	2.36	0.59

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 PDB entries followed by that with respect to all EM entries.

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	1019/1021 (100%)	992 (97%)	27 (3%)	0	100	100
1	B	1019/1021 (100%)	992 (97%)	27 (3%)	0	100	100
1	C	1019/1021 (100%)	992 (97%)	27 (3%)	0	100	100
1	D	1019/1021 (100%)	990 (97%)	29 (3%)	0	100	100
All	All	4076/4084 (100%)	3966 (97%)	110 (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 PDB entries followed by that with respect to all EM entries.

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	873/873 (100%)	872 (100%)	1 (0%)	94	94
1	B	873/873 (100%)	872 (100%)	1 (0%)	94	94
1	C	873/873 (100%)	872 (100%)	1 (0%)	94	94
1	D	873/873 (100%)	872 (100%)	1 (0%)	94	94
All	All	3492/3492 (100%)	3488 (100%)	4 (0%)	94	94

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	GLN
1	B	12	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	12	GLN
1	D	12	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	966	GLN
1	C	12	GLN
1	D	266	GLN
1	B	783	GLN
1	B	935	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 16 are monoatomic - leaving 4 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	PTQ	A	2001	4	21,21,21	2.50	4 (19%)	27,28,28	1.76	4 (14%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	PTQ	B	2001	4	21,21,21	2.51	4 (19%)	27,28,28	1.76	4 (14%)
2	PTQ	C	2001	4	21,21,21	2.51	4 (19%)	27,28,28	1.76	4 (14%)
2	PTQ	D	2001	4	21,21,21	2.52	5 (23%)	27,28,28	1.76	4 (14%)

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	PTQ	A	2001	4	-	0/8/28/28	0/2/2/2
2	PTQ	B	2001	4	-	0/8/28/28	0/2/2/2
2	PTQ	C	2001	4	-	0/8/28/28	0/2/2/2
2	PTQ	D	2001	4	-	0/8/28/28	0/2/2/2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2001	PTQ	C7-S1	-2.66	1.78	1.81
2	C	2001	PTQ	C7-S1	-2.66	1.78	1.81
2	A	2001	PTQ	C7-S1	-2.66	1.78	1.81
2	B	2001	PTQ	C7-S1	-2.66	1.78	1.81
2	D	2001	PTQ	C11-C10	-2.02	1.34	1.38

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2001	PTQ	C3-C4-C5	-2.97	104.93	110.24
2	B	2001	PTQ	C3-C4-C5	-2.97	104.94	110.24
2	C	2001	PTQ	C3-C4-C5	-2.96	104.94	110.24
2	D	2001	PTQ	C3-C4-C5	-2.96	104.95	110.24
2	C	2001	PTQ	C7-C8-C9	-2.72	109.36	114.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	2001	PTQ	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2001	PTQ	2	0
2	C	2001	PTQ	1	0
2	D	2001	PTQ	2	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.