



wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 10, 2018 – 11:37 PM EDT

PDB ID : 2CXG
Title : CYCLODEXTRIN GLYCOSYLTRANSFERASE COMPLEXED TO THE INHIBITOR ACARBOSE
Authors : Strokopytov, B.V.; Uitdehaag, J.C.M.; Ruiterkamp, R.; Dijkstra, B.W.
Deposited on : 1998-05-08
Resolution : 2.50 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 ⓘ symbol.

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

MolProbity	:	4.02b-467
Mogul	:	1.7.3 (157068), CSD as539be (2018)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
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-20031021

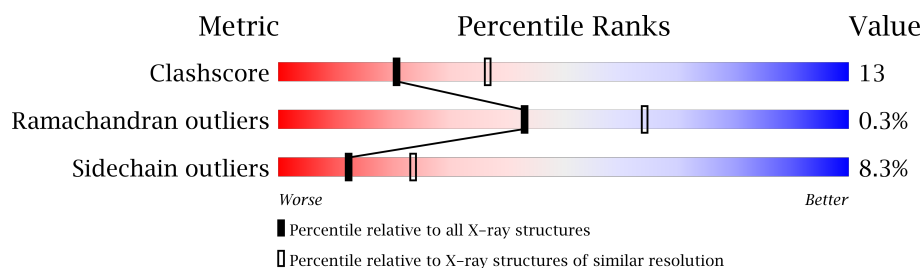
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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))
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)

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. 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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	686	

2 Entry composition [i](#)

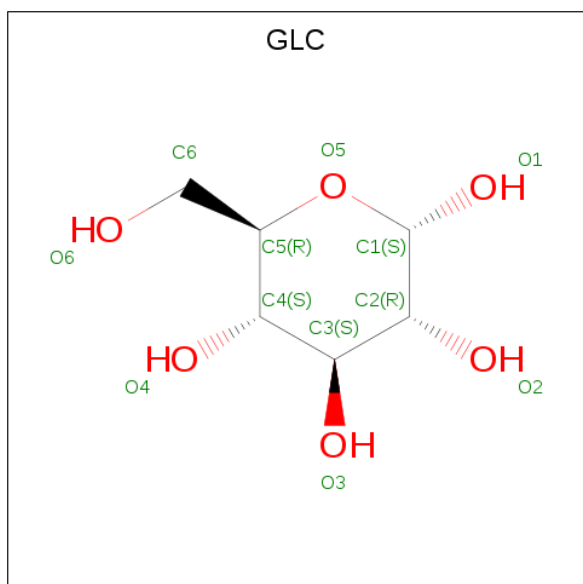
There are 6 unique types of molecules in this entry. The entry contains 5493 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 CYCLODEXTRIN GLYCOSYLTRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	686	5264	3321	900	1027	16	0	0	0

- Molecule 2 is ALPHA-D-GLUCOSE (three-letter code: GLC) (formula: C₆H₁₂O₆).



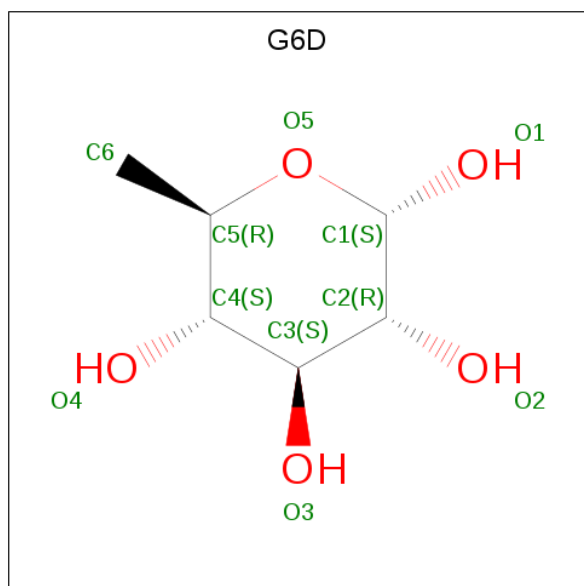
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			11	6	5		
2	A	1	Total	C	O	0	0
			12	6	6		
2	A	1	Total	C	O	0	0
			11	6	5		
2	A	1	Total	C	O	0	0
			12	6	6		
2	A	1	Total	C	O	0	0
			11	6	5		

Continued on next page...

Continued from previous page...

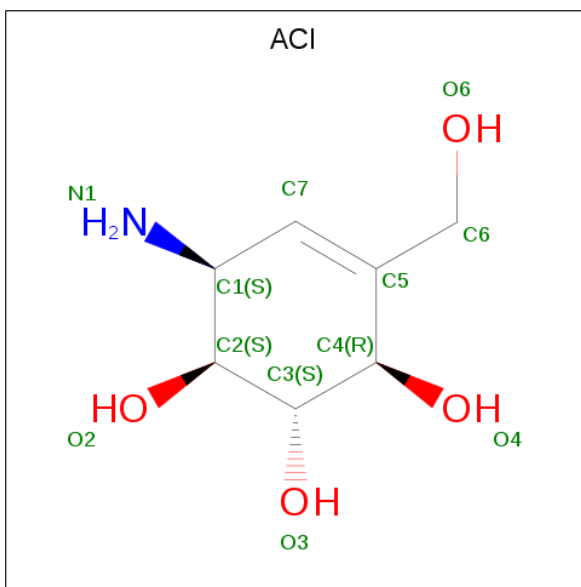
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			12	6	6		
2	A	1	Total	C	O	0	0
			11	6	5		
2	A	1	Total	C	O	0	0
			11	6	5		
2	A	1	Total	C	O	0	0
			11	6	5		
2	A	1	Total	C	O	0	0
			12	6	6		

- Molecule 3 is 6-DEOXY-ALPHA-D-GLUCOSE (three-letter code: G6D) (formula: $C_6H_{12}O_5$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is 6-AMINO-4-HYDROXYMETHYL-CYCLOHEX-4-ENE-1,2,3-TRIOL (three-letter code: ACI) (formula: $C_7H_{13}NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	N	O	0	0
			11	7	1	3		

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 6 is water.

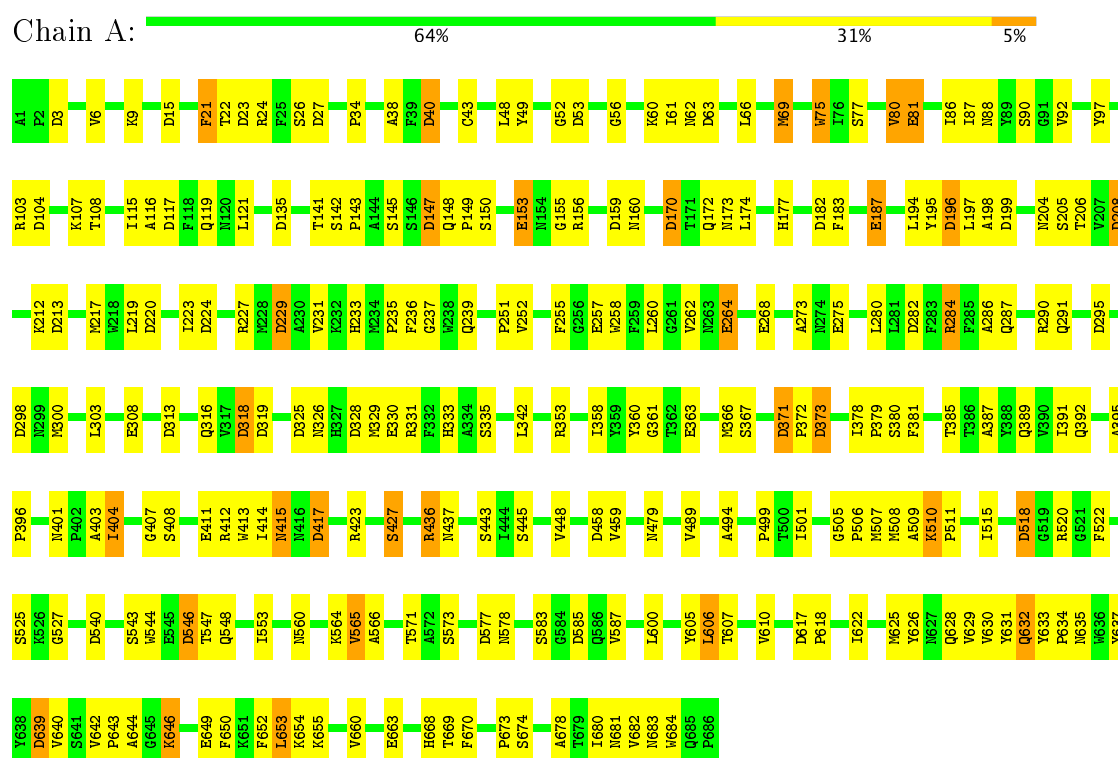
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	92	Total	O	0	0
			92	92		

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.

Note EDS was not executed.

• Molecule 1: CYCLODEXTRIN GLYCOSYLTRANSFERASE



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	120.89Å 111.90Å 65.75Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	8.00 – 2.50	Depositor
% Data completeness (in resolution range)	73.0 (8.00-2.50)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
Refinement program	TNT	Depositor
R, R_{free}	0.182 , 0.230	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	5493	wwPDB-VP
Average B, all atoms (Å ²)	14.0	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: ACI, GLC, CA, G6D

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	0.63	11/5394 (0.2%)	0.94	74/7352 (1.0%)

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	330	GLU	CD-OE2	5.40	1.31	1.25
1	A	268	GLU	CD-OE2	5.40	1.31	1.25
1	A	363	GLU	CD-OE2	5.35	1.31	1.25
1	A	411	GLU	CD-OE2	5.29	1.31	1.25
1	A	153	GLU	CD-OE2	5.29	1.31	1.25

The worst 5 of 74 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	224	ASP	CB-CG-OD2	-6.67	112.30	118.30
1	A	328	ASP	CB-CG-OD2	-6.65	112.31	118.30
1	A	27	ASP	CB-CG-OD2	-6.55	112.41	118.30
1	A	104	ASP	CB-CG-OD2	-6.39	112.55	118.30
1	A	63	ASP	CB-CG-OD2	-6.39	112.55	118.30

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	5264	0	5025	137	0
2	A	114	0	102	5	0
3	A	10	0	2	1	0
4	A	11	0	10	1	0
5	A	2	0	0	0	0
6	A	92	0	0	5	0
All	All	5493	0	5139	139	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 13.

The worst 5 of 139 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:520:ARG:HD3	1:A:547:THR:HG22	1.44	0.98
1:A:187:GLU:HG2	1:A:628:GLN:HE21	1.32	0.93
1:A:75:TRP:CZ2	1:A:227:ARG:HD3	2.20	0.76
1:A:395:ALA:HB3	1:A:396:PRO:HD3	1.69	0.74
1:A:286:ALA:O	1:A:290:ARG:HG3	1.87	0.73

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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	684/686 (100%)	637 (93%)	45 (7%)	2 (0%)	43 64

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	578	ASN
1	A	437	ASN

5.3.2 Protein sidechains [i](#)

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	556/556 (100%)	510 (92%)	46 (8%)	12	23

5 of 46 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	284	ARG
1	A	353	ARG
1	A	610	VAL
1	A	316	GLN
1	A	331	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	172	GLN
1	A	628	GLN
1	A	410	GLN
1	A	119	GLN
1	A	177	HIS

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

Of 14 ligands modelled in this entry, 2 are monoatomic - leaving 12 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	GLC	A	690	2	11,11,12	0.49	0	15,15,17	0.88	1 (6%)
2	GLC	A	691	2	12,12,12	0.31	0	17,17,17	0.75	0
2	GLC	A	693	3	11,11,12	0.35	0	15,15,17	1.35	2 (13%)
3	G6D	A	694	2,4	10,10,11	0.39	0	13,14,16	1.14	1 (7%)
4	ACI	A	695	3,2	10,11,12	1.95	3 (30%)	8,15,17	0.76	0
2	GLC	A	696	4	12,12,12	0.36	0	17,17,17	0.62	0
2	GLC	A	698	2	11,11,12	0.35	0	15,15,17	0.60	0
2	GLC	A	699	2	12,12,12	0.30	0	17,17,17	0.71	0
2	GLC	A	701	2	11,11,12	0.45	0	15,15,17	1.10	1 (6%)
2	GLC	A	702	2	11,11,12	0.45	0	15,15,17	1.85	3 (20%)
2	GLC	A	703	2	11,11,12	0.36	0	15,15,17	0.82	0
2	GLC	A	704	2	12,12,12	0.31	0	17,17,17	0.81	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	GLC	A	690	2	-	0/2/18/22	0/1/1/1
2	GLC	A	691	2	-	0/2/22/22	0/1/1/1
2	GLC	A	693	3	-	0/2/18/22	0/1/1/1
3	G6D	A	694	2,4	-	0/0/16/20	0/1/1/1
4	ACI	A	695	3,2	-	0/2/18/22	0/1/1/1
2	GLC	A	696	4	-	0/2/22/22	0/1/1/1
2	GLC	A	698	2	-	0/2/18/22	0/1/1/1
2	GLC	A	699	2	-	0/2/22/22	0/1/1/1
2	GLC	A	701	2	-	0/2/18/22	0/1/1/1
2	GLC	A	702	2	-	0/2/18/22	0/1/1/1
2	GLC	A	703	2	-	0/2/18/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GLC	A	704	2	-	0/2/22/22	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	695	ACI	C4-C5	-4.33	1.41	1.50
4	A	695	ACI	C1-C7	-3.15	1.42	1.49
4	A	695	ACI	C7-C5	2.25	1.42	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	702	GLC	O5-C5-C6	-2.28	103.11	106.88
2	A	690	GLC	C4-C3-C2	-2.02	107.95	110.61
2	A	693	GLC	C1-O5-C5	2.47	118.40	113.69
3	A	694	G6D	O5-C1-C2	2.84	115.42	110.31
2	A	701	GLC	C3-C4-C5	2.99	116.25	111.26

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

7 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	693	GLC	1	0
3	A	694	G6D	1	0
4	A	695	ACI	1	0
2	A	696	GLC	1	0
2	A	698	GLC	2	0
2	A	701	GLC	1	0
2	A	702	GLC	1	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.

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.