



Full wwPDB X-ray Structure Validation Report ⓘ

May 28, 2020 – 07:16 pm BST

PDB ID : 1CGT
Title : STRUCTURE OF CYCLODEXTRIN GLYCOSYLTRANSFERASE REFINED AT 2.0 ANGSTROMS RESOLUTION
Authors : Klein, C.; Schulz, G.E.
Deposited on : 1992-06-10
Resolution : 2.00 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 ⓘ symbol.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

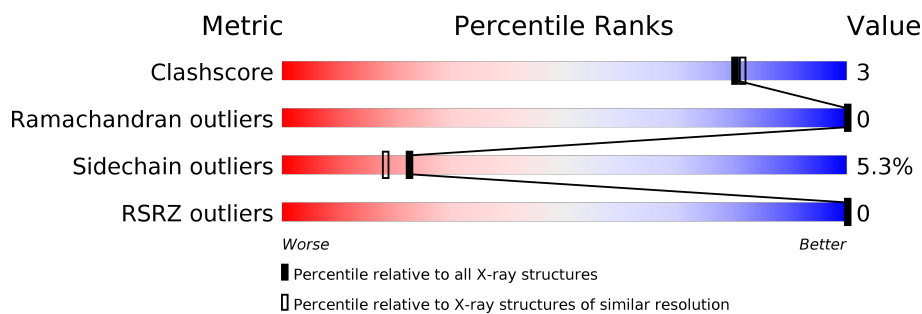
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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 $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	684	 86% 11% •

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 5857 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 GLYCOSYL-TRANSFERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	684	5267	3323	891	1040	13	0	0	0

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

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

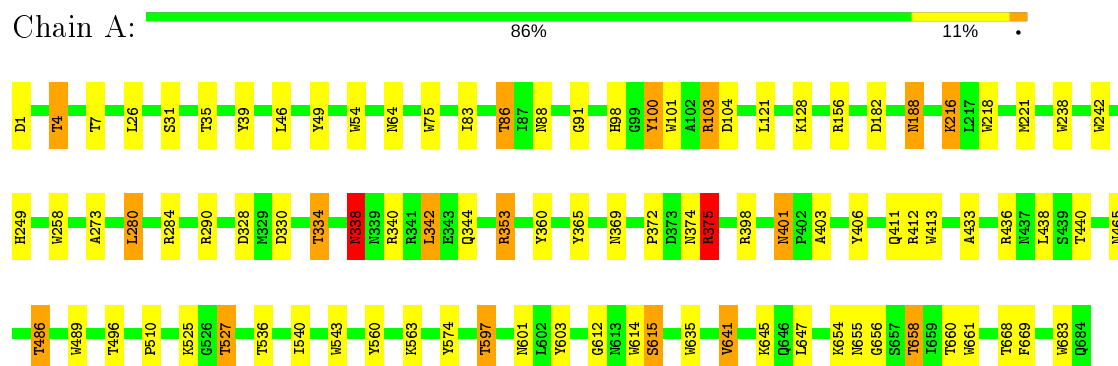
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	588	Total	O	0	0
			588	588		

3 Residue-property plots [i](#)

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). 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: CYCLODEXTRIN GLYCOSYL-TRANSFERASE



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	94.80Å 104.70Å 114.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.00 – 2.00 72.89 – 1.99	Depositor EDS
% Data completeness (in resolution range)	(Not available) (7.00-2.00) 92.1 (72.89-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$	-	Xtrriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.166 , (Not available) 0.158 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	22.9	Xtrriage
Anisotropy	0.171	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 76.8	EDS
L-test for twinning ¹	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	5857	wwPDB-VP
Average B, all atoms (Å ²)	24.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.20% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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:
CA

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	0.81	0/5398	1.48	69/7366 (0.9%)

There are no bond length outliers.

All (69) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	284	ARG	NE-CZ-NH2	-22.90	108.85	120.30
1	A	284	ARG	NE-CZ-NH1	17.66	129.13	120.30
1	A	375	ARG	NE-CZ-NH1	14.42	127.51	120.30
1	A	375	ARG	NE-CZ-NH2	-13.65	113.48	120.30
1	A	661	TRP	CD1-CG-CD2	8.68	113.25	106.30
1	A	412	ARG	NE-CZ-NH1	8.46	124.53	120.30
1	A	413	TRP	CD1-CG-CD2	8.45	113.06	106.30
1	A	614	TRP	CD1-CG-CD2	8.32	112.95	106.30
1	A	342	LEU	CA-CB-CG	7.91	133.48	115.30
1	A	258	TRP	CD1-CG-CD2	7.83	112.57	106.30
1	A	436	ARG	NE-CZ-NH2	-7.73	116.43	120.30
1	A	563	LYS	CB-CG-CD	-7.71	91.56	111.60
1	A	683	TRP	CD1-CG-CD2	7.51	112.31	106.30
1	A	413	TRP	CE2-CD2-CG	-7.41	101.37	107.30
1	A	4	THR	N-CA-CB	-7.34	96.35	110.30
1	A	489	TRP	CD1-CG-CD2	7.30	112.14	106.30
1	A	353	ARG	NE-CZ-NH2	-7.22	116.69	120.30
1	A	641	VAL	N-CA-CB	-7.12	95.84	111.50
1	A	683	TRP	CE2-CD2-CG	-7.10	101.62	107.30
1	A	258	TRP	CE2-CD2-CG	-7.05	101.66	107.30
1	A	661	TRP	CE2-CD2-CG	-7.05	101.66	107.30
1	A	489	TRP	CE2-CD2-CG	-7.03	101.68	107.30
1	A	49	TYR	CB-CG-CD1	-6.88	116.87	121.00
1	A	86	THR	N-CA-CB	-6.81	97.36	110.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	101	TRP	CD1-CG-CD2	6.75	111.70	106.30
1	A	614	TRP	CE2-CD2-CG	-6.71	101.93	107.30
1	A	365	TYR	CB-CG-CD2	-6.70	116.98	121.00
1	A	635	TRP	CD1-CG-CD2	6.58	111.56	106.30
1	A	543	TRP	CD1-CG-CD2	6.55	111.54	106.30
1	A	75	TRP	CE2-CD2-CG	-6.53	102.08	107.30
1	A	635	TRP	CE2-CD2-CG	-6.52	102.08	107.30
1	A	242	TRP	CD1-CG-CD2	6.46	111.47	106.30
1	A	543	TRP	CE2-CD2-CG	-6.42	102.16	107.30
1	A	75	TRP	CD1-CG-CD2	6.39	111.41	106.30
1	A	238	TRP	CD1-CG-CD2	6.36	111.39	106.30
1	A	641	VAL	CB-CA-C	6.34	123.45	111.40
1	A	101	TRP	CE2-CD2-CG	-6.32	102.25	107.30
1	A	54	TRP	CE2-CD2-CG	-6.29	102.27	107.30
1	A	54	TRP	CD1-CG-CD2	6.25	111.30	106.30
1	A	218	TRP	CD1-CG-CD2	6.23	111.28	106.30
1	A	360	TYR	CB-CG-CD2	-6.18	117.29	121.00
1	A	238	TRP	CE2-CD2-CG	-6.09	102.43	107.30
1	A	218	TRP	CE2-CD2-CG	-6.09	102.43	107.30
1	A	413	TRP	CB-CG-CD1	-6.08	119.09	127.00
1	A	290	ARG	NE-CZ-NH1	6.05	123.33	120.30
1	A	597	THR	N-CA-CB	-5.90	99.10	110.30
1	A	284	ARG	CB-CG-CD	-5.89	96.28	111.60
1	A	104	ASP	CB-CG-OD1	5.84	123.56	118.30
1	A	103	ARG	NE-CZ-NH2	-5.80	117.40	120.30
1	A	100	TYR	CB-CG-CD2	-5.78	117.53	121.00
1	A	156	ARG	NE-CZ-NH1	5.76	123.18	120.30
1	A	654	LYS	CB-CG-CD	-5.67	96.86	111.60
1	A	242	TRP	CE2-CD2-CG	-5.63	102.80	107.30
1	A	398	ARG	NE-CZ-NH2	-5.61	117.49	120.30
1	A	603	TYR	CB-CG-CD2	-5.58	117.65	121.00
1	A	661	TRP	CG-CD1-NE1	-5.57	104.53	110.10
1	A	413	TRP	CG-CD1-NE1	-5.47	104.62	110.10
1	A	221	MET	CA-CB-CG	-5.46	104.02	113.30
1	A	398	ARG	NE-CZ-NH1	5.42	123.01	120.30
1	A	406	TYR	CB-CG-CD1	-5.33	117.80	121.00
1	A	614	TRP	CG-CD1-NE1	-5.30	104.80	110.10
1	A	39	TYR	CB-CG-CD2	-5.25	117.85	121.00
1	A	328	ASP	CB-CG-OD2	-5.24	113.58	118.30
1	A	574	TYR	CB-CG-CD1	-5.22	117.87	121.00
1	A	338	ASN	CB-CG-ND2	-5.21	104.20	116.70
1	A	656	GLY	CA-C-N	-5.09	106.00	117.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	560	TYR	CB-CG-CD2	-5.08	117.95	121.00
1	A	101	TRP	CG-CD1-NE1	-5.06	105.04	110.10
1	A	635	TRP	CG-CD2-CE3	5.05	138.45	133.90

There are no chirality outliers.

There are no planarity outliers.

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	5267	0	5008	28	0
2	A	2	0	0	0	0
3	A	588	0	0	4	0
All	All	5857	0	5008	28	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.

All (28) 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:98:HIS:HD2	1:A:100:TYR:H	1.30	0.80
1:A:340:ARG:HH12	1:A:465:ASN:HD22	1.40	0.68
1:A:334:THR:HG21	3:A:1111:HOH:O	1.93	0.67
1:A:340:ARG:HH12	1:A:465:ASN:ND2	1.99	0.60
1:A:88:ASN:HD21	1:A:91:GLY:H	1.52	0.57
1:A:647:LEU:HD12	3:A:1238:HOH:O	2.04	0.56
1:A:273:ALA:HB2	1:A:280:LEU:HD22	1.87	0.56
1:A:655:ASN:OD1	1:A:658:THR:HG23	2.07	0.55
1:A:338:ASN:HD21	1:A:438:LEU:HD11	1.72	0.54
1:A:527:THR:HB	1:A:536:THR:HG22	1.90	0.54
1:A:401:ASN:HD22	1:A:403:ALA:H	1.57	0.53
1:A:64:ASN:OD1	1:A:128:LYS:NZ	2.41	0.52
1:A:330:ASP:HB3	1:A:369:ASN:HD22	1.78	0.48
1:A:525:LYS:HD2	1:A:540:ILE:HB	1.95	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:401:ASN:ND2	1:A:403:ALA:H	2.11	0.48
1:A:612:GLY:O	1:A:615:SER:HB3	2.13	0.47
1:A:344:GLN:HG2	1:A:486:THR:CG2	2.45	0.46
1:A:88:ASN:ND2	1:A:91:GLY:H	2.11	0.46
1:A:1:ASP:OD2	1:A:7:THR:HG21	2.16	0.46
1:A:330:ASP:HB3	1:A:369:ASN:ND2	2.32	0.44
1:A:645:LYS:NZ	3:A:1253:HOH:O	2.51	0.44
1:A:669:PHE:HD2	3:A:1238:HOH:O	2.00	0.44
1:A:182:ASP:H	1:A:188:ASN:HD21	1.65	0.43
1:A:344:GLN:HG2	1:A:486:THR:HG21	2.01	0.43
1:A:433:ALA:O	1:A:486:THR:HA	2.19	0.42
1:A:83:ILE:HA	1:A:103:ARG:HD3	2.00	0.42
1:A:216:LYS:HG3	1:A:249:HIS:CE1	2.54	0.41
1:A:374:ASN:OD1	1:A:375:ARG:HD3	2.22	0.40

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	682/684 (100%)	665 (98%)	17 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	566/566 (100%)	536 (95%)	30 (5%)	22 18

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	4	THR
1	A	26	LEU
1	A	31	SER
1	A	35	THR
1	A	46	LEU
1	A	86	THR
1	A	121	LEU
1	A	188	ASN
1	A	216	LYS
1	A	280	LEU
1	A	334	THR
1	A	338	ASN
1	A	342	LEU
1	A	353	ARG
1	A	372	PRO
1	A	375	ARG
1	A	401	ASN
1	A	411	GLN
1	A	440	THR
1	A	486	THR
1	A	496	THR
1	A	510	PRO
1	A	527	THR
1	A	597	THR
1	A	601	ASN
1	A	615	SER
1	A	641	VAL
1	A	658	THR
1	A	660	THR
1	A	668	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	29	ASN
1	A	88	ASN
1	A	98	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	188	ASN
1	A	204	ASN
1	A	249	HIS
1	A	318	ASN
1	A	338	ASN
1	A	369	ASN
1	A	401	ASN
1	A	465	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 2 ligands modelled in this entry, 2 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 [i](#)

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 [i](#)

6.1 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	684/684 (100%)	-0.56	0 100 100	11, 20, 39, 67	0

There are no RSRZ outliers to report.

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 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
2	CA	A	686	1/1	0.98	0.06	18,18,18,18	0
2	CA	A	685	1/1	1.00	0.10	19,19,19,19	0

6.5 Other polymers [i](#)

There are no such residues in this entry.