



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 10:00 PM UTC

PDB ID : 6TS8 / pdb_00006ts8
Title : Chaetomium thermophilum UDP-Glucose Glucosyl Transferase (UGGT) double cysteine mutant G177C/A786C.
Authors : Roversi, P.; Zitzmann, N.; Ibba, R.; Hensen, M.; Chandran, A.
Deposited on : 2019-12-20
Resolution : 4.60 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

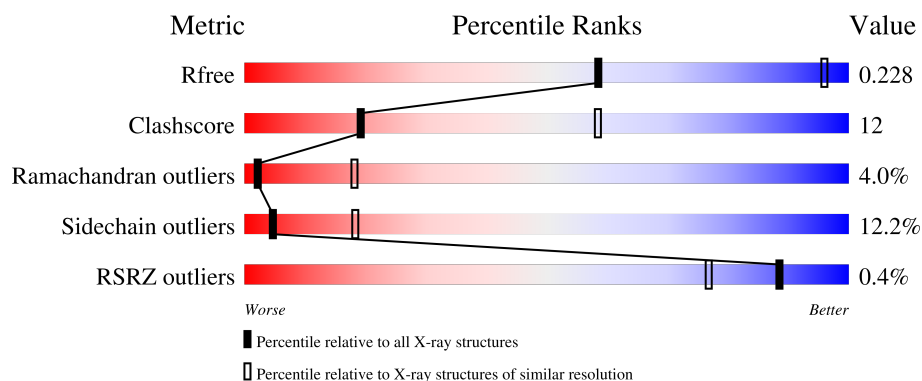
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.60 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}	180053	1007 (5.28-3.92)
Clashscore	190562	1022 (5.26-3.94)
Ramachandran outliers	187476	1069 (5.30-3.90)
Sidechain outliers	187428	1051 (5.30-3.90)
RSRZ outliers	180081	1002 (5.28-3.92)

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 of 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	1382	
1	B	1382	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 20112 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 UDP-glucose-glycoprotein glucosyltransferase-like protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1254	Total	C	N	O	S	0	0	0
			10056	6439	1711	1879	27			
1	B	1254	Total	C	N	O	S	0	0	0
			10056	6439	1711	1879	27			

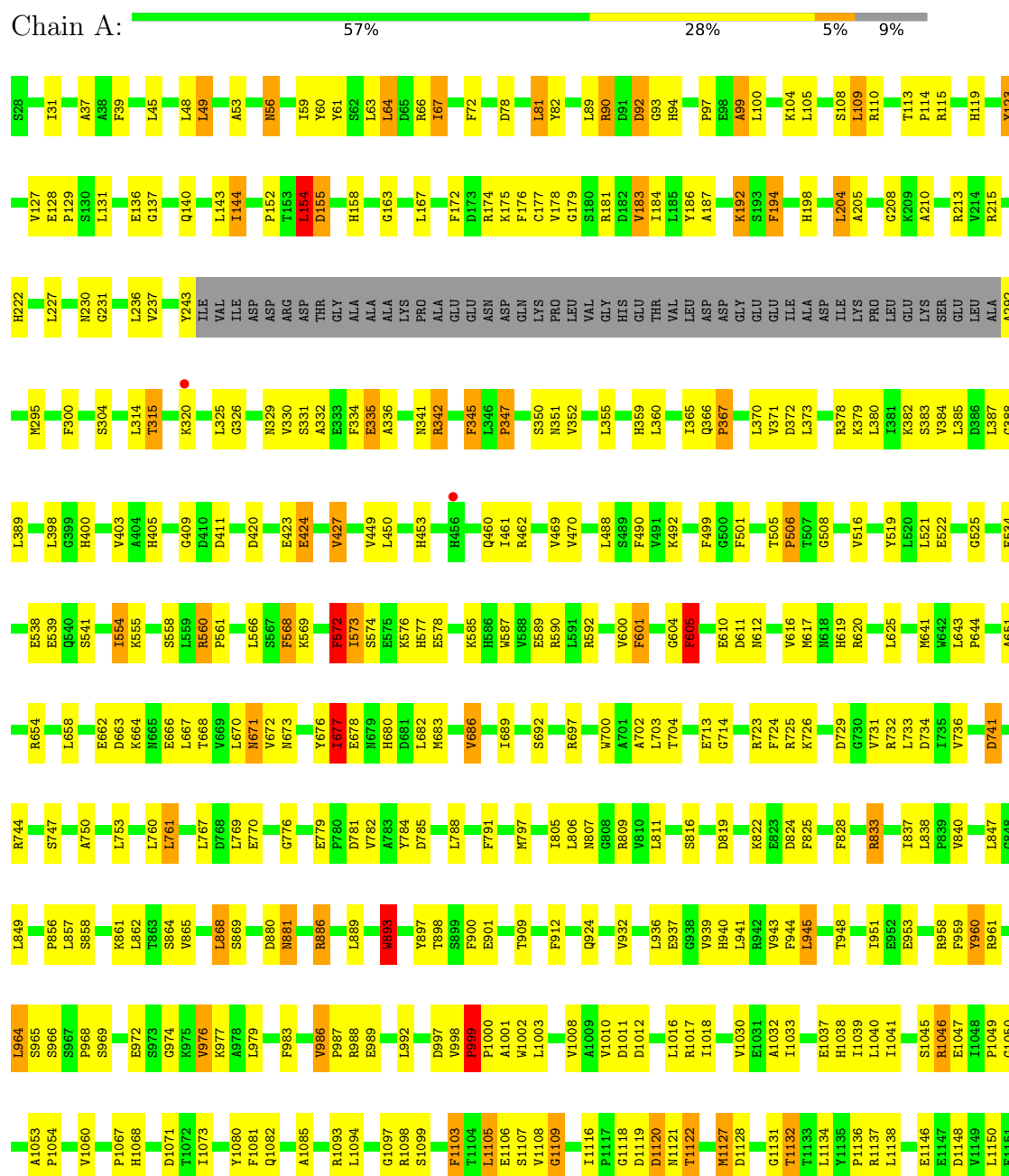
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	177	CYS	GLY	engineered mutation	UNP G0SB58
A	786	CYS	ALA	engineered mutation	UNP G0SB58
B	177	CYS	GLY	engineered mutation	UNP G0SB58
B	786	CYS	ALA	engineered mutation	UNP G0SB58

3 Residue-property plots

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: UDP-glucose-glycoprotein glucosyltransferase-like protein





[illegible]

4 Data and refinement statistics

Property	Value	Source
Space group	P 43	Depositor
Cell constants a, b, c, α , β , γ	139.05Å 139.05Å 176.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	139.04 – 4.60 139.04 – 4.77	Depositor EDS
% Data completeness (in resolution range)	32.1 (139.04-4.60) 36.0 (139.04-4.77)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.76 (at 4.88Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.203 , 0.234 0.202 , 0.228	Depositor DCC
R_{free} test set	339 reflections (2.03%)	wwPDB-VP
Wilson B-factor (Å ²)	206.4	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.47 , 718.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Estimated twinning fraction	0.369 for h,-k,-l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	20112	wwPDB-VP
Average B, all atoms (Å ²)	173.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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](#)

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	0.97	5/10293 (0.0%)	1.27	13/13959 (0.1%)
1	B	0.97	5/10293 (0.0%)	1.27	12/13959 (0.1%)
All	All	0.97	10/20586 (0.0%)	1.27	25/27918 (0.1%)

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	750	ALA	C-O	-6.34	1.16	1.24
1	B	750	ALA	C-O	-6.10	1.16	1.24
1	A	519	TYR	C-O	-6.05	1.17	1.24
1	B	519	TYR	C-O	-5.82	1.17	1.24
1	A	152	PRO	C-O	-5.53	1.17	1.24

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	304	SER	CA-C-N	5.85	130.67	122.08
1	B	304	SER	C-N-CA	5.85	130.67	122.08
1	A	304	SER	CA-C-N	5.81	130.62	122.08
1	A	304	SER	C-N-CA	5.81	130.62	122.08
1	B	345	PHE	CA-CB-CG	5.52	119.32	113.80

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	10056	0	9955	233	0
1	B	10056	0	9955	236	0
All	All	20112	0	19910	469	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 12.

The worst 5 of 469 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:1109:GLY:H	1:A:1136:PRO:HA	1.12	1.11
1:B:1109:GLY:H	1:B:1136:PRO:HA	1.12	1.06
1:A:382:LYS:HA	1:A:385:LEU:HG	1.51	0.93
1:B:382:LYS:HA	1:B:385:LEU:HG	1.51	0.92
1:B:833:ARG:HA	1:B:837:ILE:HB	1.55	0.87

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	1246/1382 (90%)	988 (79%)	208 (17%)	50 (4%)	2	18
1	B	1246/1382 (90%)	989 (79%)	207 (17%)	50 (4%)	2	18
All	All	2492/2764 (90%)	1977 (79%)	415 (17%)	100 (4%)	2	18

5 of 100 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	155	ASP
1	A	424	GLU
1	A	427	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	506	PRO
1	A	541	SER

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	1097/1197 (92%)	964 (88%)	133 (12%)	5	18
1	B	1097/1197 (92%)	963 (88%)	134 (12%)	5	18
All	All	2194/2394 (92%)	1927 (88%)	267 (12%)	5	18

5 of 267 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1046	ARG
1	B	1120	ASP
1	B	1312	TYR
1	A	1046	ARG
1	A	1016	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	158	HIS
1	B	366	GLN
1	B	1397	GLN
1	B	166	GLN
1	B	341	ASN

5.3.3 RNA ⓘ

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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 [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	1254/1382 (90%)	-0.90	2 (0%) 91 82	73, 141, 251, 333	0
1	B	1254/1382 (90%)	-0.81	7 (0%) 85 73	106, 188, 295, 398	0
All	All	2508/2764 (90%)	-0.85	9 (0%) 88 77	73, 168, 272, 398	0

The worst 5 of 9 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	181	ARG	3.6
1	B	851	ASP	3.2
1	B	308	PHE	3.0
1	A	456	HIS	3.0
1	B	667	LEU	2.7

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.