



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 18, 2024 – 01:15 PM EDT

PDB ID : 3WZT  
Title : Crystal structure of Trx3 domain of UGGT (detergent-unbound form)  
Authors : Zhu, T.; Satoh, T.; Kato, K.  
Deposited on : 2014-10-03  
Resolution : 3.40 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](mailto: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 : 4.02b-467  
Mogul : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

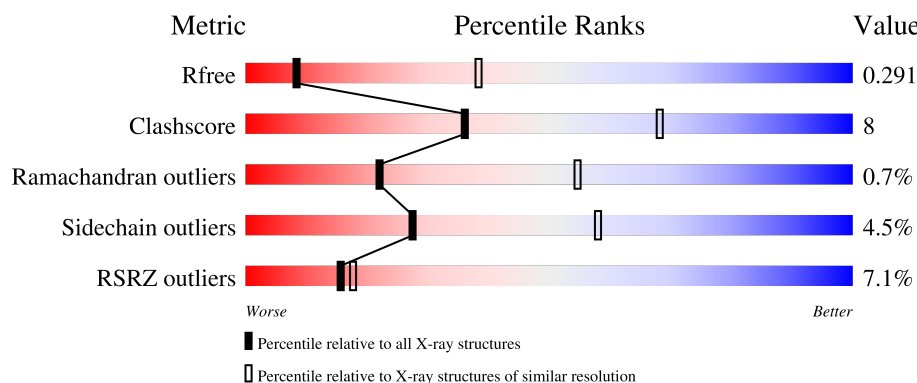
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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  $\geq 3$ , 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	163	<div> <div>2%</div> <div>83%</div> <div>10%</div> <div>5%</div> </div>
1	B	163	<div> <div>2%</div> <div>85%</div> <div>10%</div> <div>5%</div> </div>
1	C	163	<div> <div>2%</div> <div>75%</div> <div>9%</div> <div>13%</div> </div>
1	D	163	<div> <div>3%</div> <div>80%</div> <div>13%</div> <div>6%</div> </div>
1	E	163	<div> <div>9%</div> <div>39%</div> <div>16%</div> <div>42%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	163	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: red (17%), green (62%), yellow (6%), and grey (32%). The percentages are labeled below the bar. A small black dot is located at the end of the yellow segment.

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 6452 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	155	Total	C	N	O	Se	0	0	0
			1239	783	208	246	2			
1	B	156	Total	C	N	O	Se	0	0	0
			1246	788	209	247	2			
1	C	142	Total	C	N	O	Se	0	0	0
			1127	715	189	221	2			
1	D	154	Total	C	N	O	Se	0	0	0
			1231	779	206	244	2			
1	E	94	Total	C	N	O	Se	0	0	0
			738	469	122	145	2			
1	F	111	Total	C	N	O	Se	0	0	0
			871	552	150	168	1			

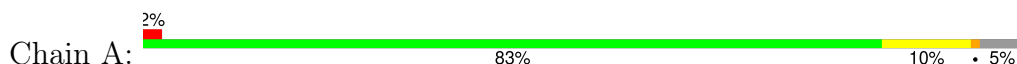
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	669	GLY	-	expression tag	UNP G0SB58
A	670	SER	-	expression tag	UNP G0SB58
B	669	GLY	-	expression tag	UNP G0SB58
B	670	SER	-	expression tag	UNP G0SB58
C	669	GLY	-	expression tag	UNP G0SB58
C	670	SER	-	expression tag	UNP G0SB58
D	669	GLY	-	expression tag	UNP G0SB58
D	670	SER	-	expression tag	UNP G0SB58
E	669	GLY	-	expression tag	UNP G0SB58
E	670	SER	-	expression tag	UNP G0SB58
F	669	GLY	-	expression tag	UNP G0SB58
F	670	SER	-	expression tag	UNP G0SB58

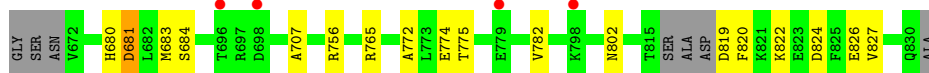
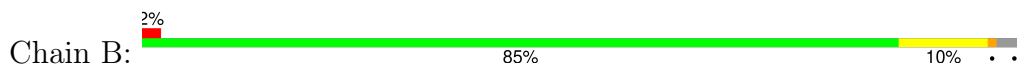
### 3 Residue-property plots [i](#)

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.

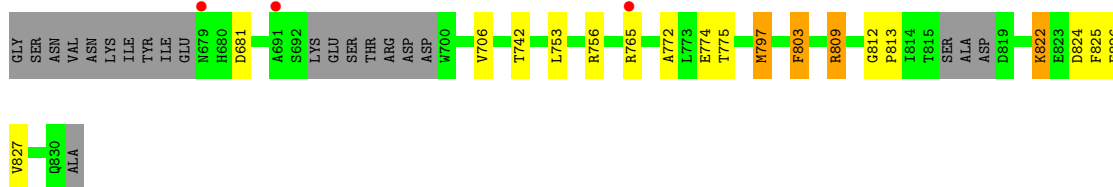
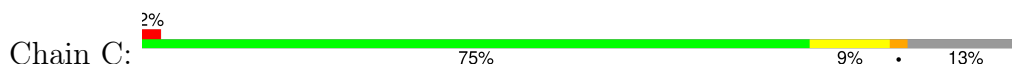
- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein



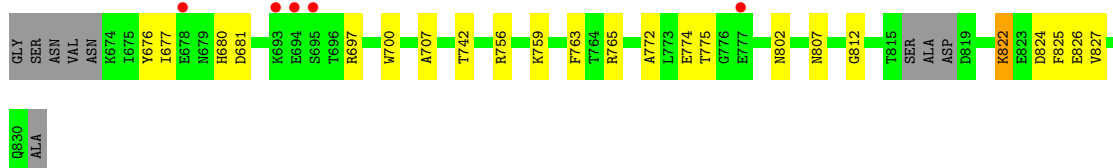
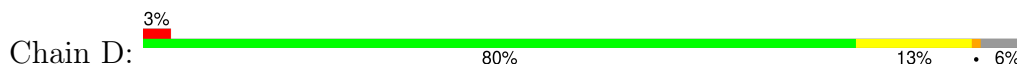
- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein



- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein

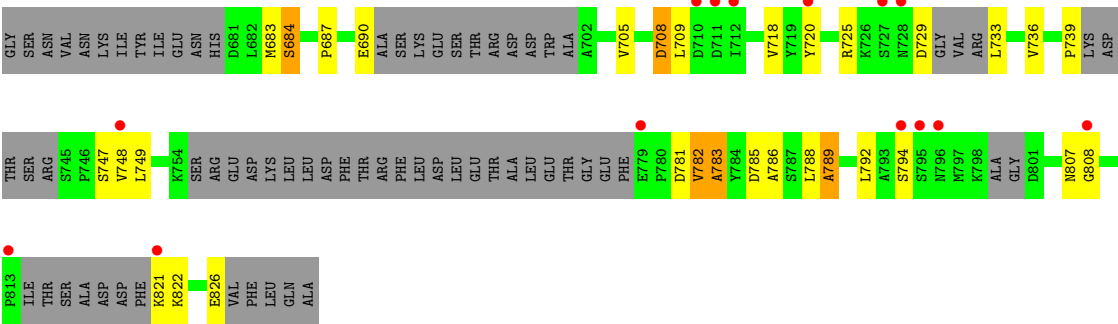


- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein

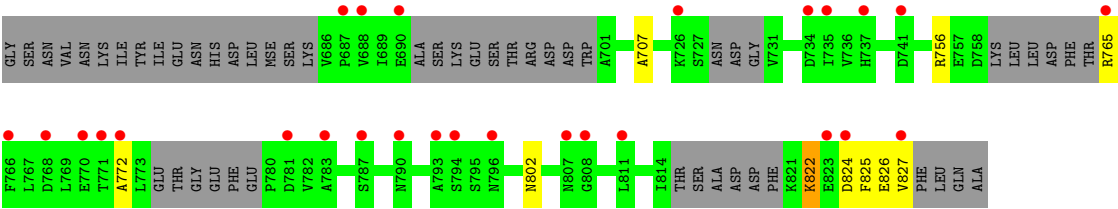


- Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein





● Molecule 1: UDP-glucose-glycoprotein glucosyltransferase-like protein



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	196.40Å 196.40Å 196.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.40 19.84 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.4 (20.00-3.40) 100.0 (19.84-3.40)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.39 (at 3.44Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, $R_{free}$	0.235 , 0.292 0.234 , 0.291	Depositor DCC
$R_{free}$ test set	865 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	89.3	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 61.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.022 for -l,-k,-h	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	6452	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	99.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.97% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

### 5.1 Standard geometry

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	0.54	0/1257	0.69	2/1695 (0.1%)
1	B	0.55	0/1264	0.69	1/1705 (0.1%)
1	C	0.53	0/1143	0.74	4/1541 (0.3%)
1	D	0.53	0/1249	0.67	0/1684
1	E	0.54	0/742	0.69	0/993
1	F	0.43	0/878	0.63	0/1179
All	All	0.53	0/6533	0.69	7/8797 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	E	0	6
All	All	0	7

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	797	MSE	CA-CB-CG	8.07	127.01	113.30
1	A	765	ARG	NE-CZ-NH1	7.90	124.25	120.30
1	C	803	PHE	CB-CA-C	-5.77	98.85	110.40
1	A	765	ARG	NE-CZ-NH2	-5.54	117.53	120.30
1	C	803	PHE	CB-CG-CD1	5.17	124.42	120.80

There are no chirality outliers.

5 of 7 planarity outliers are listed below:



Mol	Chain	Res	Type	Group
1	A	673	ASN	Peptide
1	E	705	VAL	Peptide
1	E	733	LEU	Peptide
1	E	782	VAL	Peptide
1	E	783	ALA	Peptide

## 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	1239	0	1216	12	0
1	B	1246	0	1225	16	0
1	C	1127	0	1107	25	0
1	D	1231	0	1210	42	0
1	E	738	0	731	33	0
1	F	871	0	877	4	0
All	All	6452	0	6366	102	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 8.

The worst 5 of 102 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:B:782:VAL:CG2	1:D:812:GLY:HA2	1.80	1.10
1:B:782:VAL:HG21	1:D:812:GLY:HA2	1.33	1.09
1:D:700:TRP:HH2	1:E:783:ALA:H	1.11	0.96
1:D:700:TRP:HH2	1:E:783:ALA:N	1.62	0.96
1:B:782:VAL:CG2	1:D:812:GLY:CA	2.56	0.82

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 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	151/163 (93%)	142 (94%)	9 (6%)	0	100	100
1	B	152/163 (93%)	146 (96%)	6 (4%)	0	100	100
1	C	136/163 (83%)	129 (95%)	7 (5%)	0	100	100
1	D	150/163 (92%)	143 (95%)	7 (5%)	0	100	100
1	E	80/163 (49%)	66 (82%)	9 (11%)	5 (6%)	1	9
1	F	99/163 (61%)	96 (97%)	3 (3%)	0	100	100
All	All	768/978 (78%)	722 (94%)	41 (5%)	5 (1%)	22	55

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	708	ASP
1	E	807	ASN
1	E	789	ALA
1	E	748	VAL
1	E	687	PRO

### 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	137/140 (98%)	131 (96%)	6 (4%)	28	58
1	B	138/140 (99%)	133 (96%)	5 (4%)	35	63
1	C	124/140 (89%)	117 (94%)	7 (6%)	21	51

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	136/140 (97%)	130 (96%)	6 (4%)	28	58
1	E	83/140 (59%)	78 (94%)	5 (6%)	19	49
1	F	96/140 (69%)	93 (97%)	3 (3%)	40	68
All	All	714/840 (85%)	682 (96%)	32 (4%)	27	58

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	826	GLU
1	F	765	ARG
1	C	742	THR
1	C	681	ASP
1	F	822	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	728	ASN
1	D	680	HIS
1	D	728	ASN
1	E	796	ASN
1	E	807	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	153/163 (93%)	-0.09	4 (2%) 56 54	50, 73, 115, 176	0
1	B	154/163 (94%)	-0.13	4 (2%) 56 54	52, 74, 112, 167	0
1	C	140/163 (85%)	-0.04	3 (2%) 63 62	59, 87, 123, 155	0
1	D	152/163 (93%)	0.08	5 (3%) 46 45	55, 89, 146, 179	0
1	E	92/163 (56%)	0.81	14 (15%) 2 2	99, 132, 171, 190	0
1	F	110/163 (67%)	1.13	27 (24%) 0 0	100, 139, 163, 183	0
All	All	801/978 (81%)	0.21	57 (7%) 16 18	50, 90, 155, 190	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	690	GLU	6.0
1	F	807	ASN	5.6
1	B	696	THR	5.4
1	D	693	LYS	4.9
1	F	796	ASN	4.9

### 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 monosaccharides 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.