



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 06:04 AM GMT

PDB ID : 2VUY  
Title : Crystal structure of Glycogen Debranching exzyme TreX from Sulfolobus sol-  
fatarius  
Authors : Song, H.-N.; Yoon, S.-M.; Cha, H.-J.; Park, K.-H.; Woo, E.-J.  
Deposited on : 2008-06-02  
Resolution : 3.00 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  
<http://wwpdb.org/validation/2016/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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

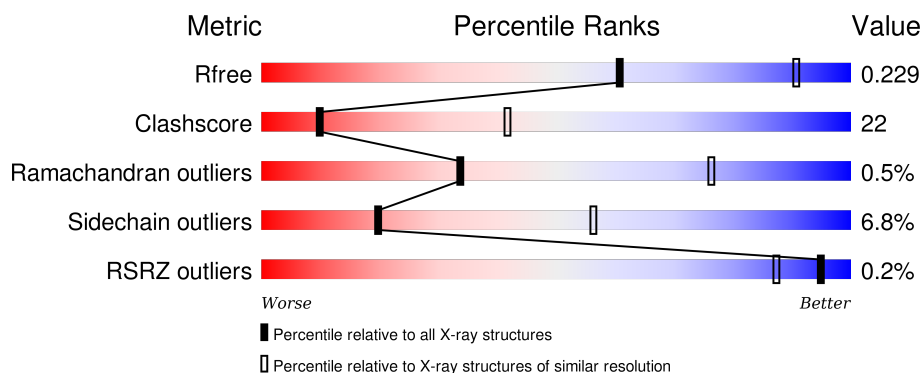
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	718	 63% 32% 5% •
1	B	718	 62% 32% 5% •

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 11634 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.

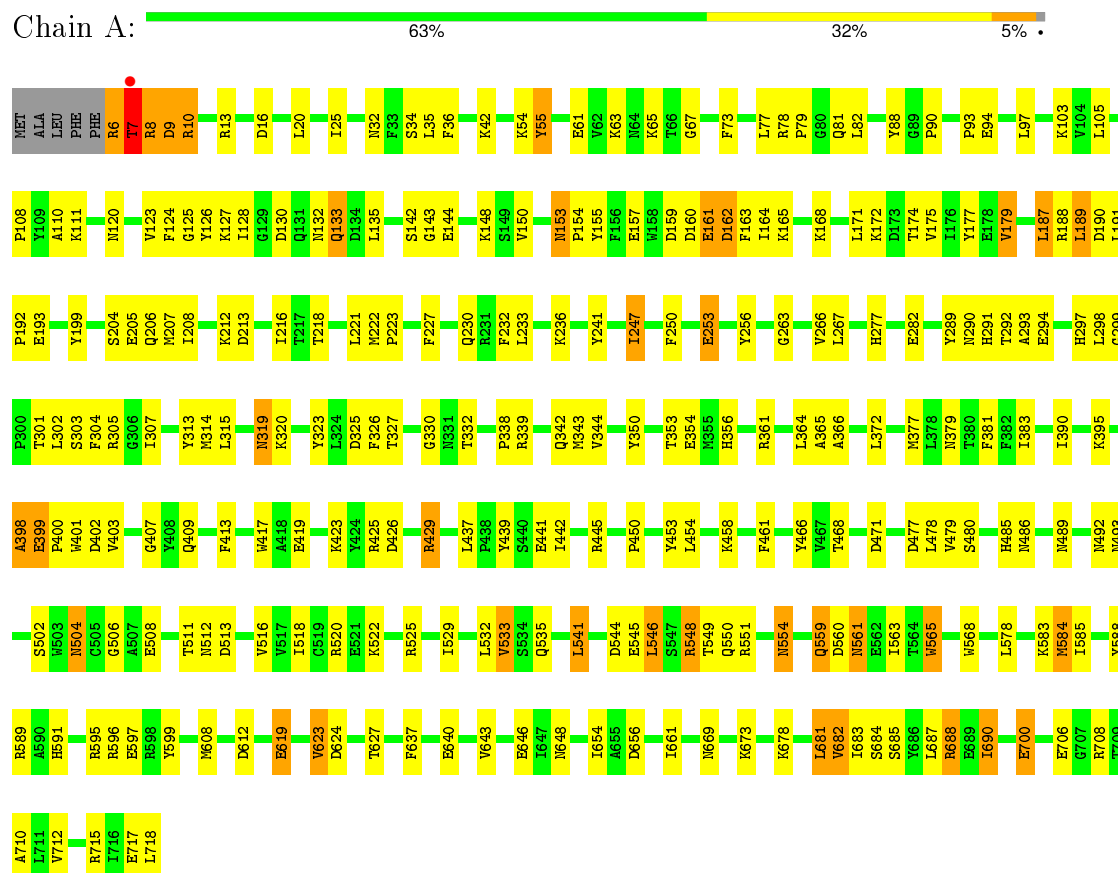
- Molecule 1 is a protein called GLYCOGEN OPERON PROTEIN GLGX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	713	Total	C	N	O	S	0	0	0
			5826	3735	983	1089	19			
1	B	711	Total	C	N	O	S	0	0	0
			5808	3725	978	1086	19			

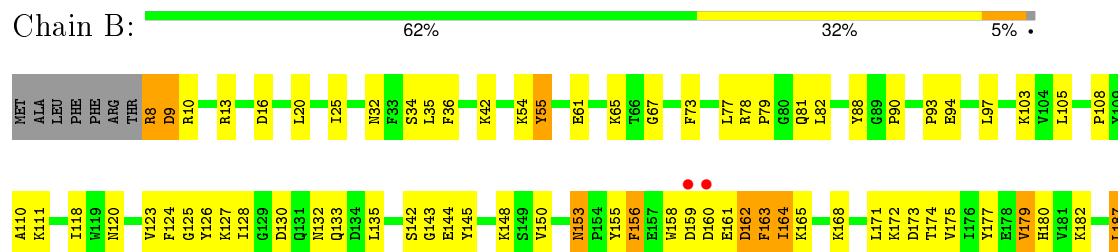
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: GLYCOGEN OPERON PROTEIN GLGX



#### • Molecule 1: GLYCOGEN OPERON PROTEIN GLGX



R708	R595	P400	P300	R188
T709	R596	H401	T301	L189
A710	R597	D402	L302	D190
L711	R598	V403	S303	L191
V712	R599	G407	F304	P192
		Y408	R305	E193
R715	R603	Q409	G306	Y199
I716			I307	
E717			Y313	S204
L718	R608	F413	M314	E205
		P414	L315	D206
	R611	Y415		W207
	D612	Q416	D318	I208
	R619	W417	M319	
			K320	K212
	R623	K423	Y323	D213
	D624	Y424	I324	I216
		R425	D325	T217
		D426	F326	T218
	R627	R429	T327	W222
			T332	P223
	R637	L437		F227
	B640	P438	P338	
		Y439	R339	
	R643	S440	Q342	Q230
		E441	M343	R231
	B646	I442	V344	L233
	I647		Y350	K236
	R648	R445	T353	I247
		L446	E354	F250
	R654	L447	R361	E253
	A655	G448	L364	Y256
	D656	S449	A365	
		P450	A366	Q263
	I661		L372	Y266
	R669	Y453	M377	H277
		L454	N379	
	R673	K458	T380	E282
		D560	F381	Y289
	R678	M561	I383	M290
		E562		H291
	L681	I563	I390	T292
	V682	T564		A293
	I683	W565		E294
	S684			
	S685	W568		
	V686	D471		
	L687	D477		
	R688	L478		
	E689	V479		
	I690	S480		
		I585		
	R695	N486		
		Y588		
	E700	R589		
		A590		
	E706	H591		
	G707	P592		

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	203.63 Å   203.63 Å   89.43 Å 90.00°   90.00°   120.00°	Depositor
Resolution (Å)	29.86 – 3.00 49.12 – 3.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (29.86-3.00) 92.8 (49.12-3.00)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	9.56 (at 3.01 Å)	Xtriage
Refinement program	?	Depositor
R, $R_{free}$	0.206 , 0.231 0.205 , 0.229	Depositor DCC
$R_{free}$ test set	2004 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	39.7	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 28.3	EDS
Estimated twinning fraction	0.024 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 42358 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11634	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.33	0/5978	0.57	0/8104
1	B	0.33	0/5960	0.57	0/8080
All	All	0.33	0/11938	0.57	0/16184

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	1	2
1	B	0	3
All	All	1	5

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	7	THR	CB

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	398	ALA	Peptide
1	A	7	THR	Peptide
1	B	398	ALA	Peptide
1	B	417	TRP	Peptide
1	B	9	ASP	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	5826	0	5642	262	0
1	B	5808	0	5622	249	0
All	All	11634	0	11264	507	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 22.

The worst 5 of 507 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:6:ARG:NH1	1:A:7:THR:HG22	1.29	1.43
1:A:6:ARG:HG2	1:A:7:THR:N	1.51	1.18
1:A:6:ARG:CG	1:A:7:THR:H	1.59	1.15
1:A:511:THR:HG22	1:A:513:ASP:H	1.13	1.10
1:B:8:ARG:HG2	1:B:8:ARG:O	1.51	1.10

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	711/718 (99%)	644 (91%)	64 (9%)	3 (0%)	39 80
1	B	709/718 (99%)	643 (91%)	62 (9%)	4 (1%)	30 72
All	All	1420/1436 (99%)	1287 (91%)	126 (9%)	7 (0%)	34 76



5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	9	ASP
1	A	700	GLU
1	B	156	PHE
1	B	414	PRO
1	B	700	GLU

### 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	629/636 (99%)	585 (93%)	44 (7%)	19	55
1	B	627/636 (99%)	586 (94%)	41 (6%)	21	58
All	All	1256/1272 (99%)	1171 (93%)	85 (7%)	20	56

5 of 85 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	619	GLU
1	B	55	TYR
1	B	584	MET
1	A	623	VAL
1	A	690	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 58 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	561	ASN
1	B	102	ASN
1	B	559	GLN
1	A	601	GLN
1	A	669	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](#)

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	713/718 (99%)	-0.56	1 (0%) 95 90	16, 32, 53, 100	0
1	B	711/718 (99%)	-0.49	2 (0%) 94 84	13, 34, 58, 83	0
All	All	1424/1436 (99%)	-0.53	3 (0%) 95 87	13, 33, 55, 100	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	7	THR	4.2
1	B	160	ASP	2.6
1	B	159	ASP	2.0

### 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](#)

There are no ligands in this entry.

### 6.5 Other polymers [i](#)

There are no such residues in this entry.