wwPDB X-ray Structure Validation Summary Report (i)

Dec 4, 2023 – 11:08 pm GMT

PDB ID	:	2VUY
Title	:	Crystal structure of Glycogen Debranching exzyme TreX from Sulfolobus solfatarius
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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

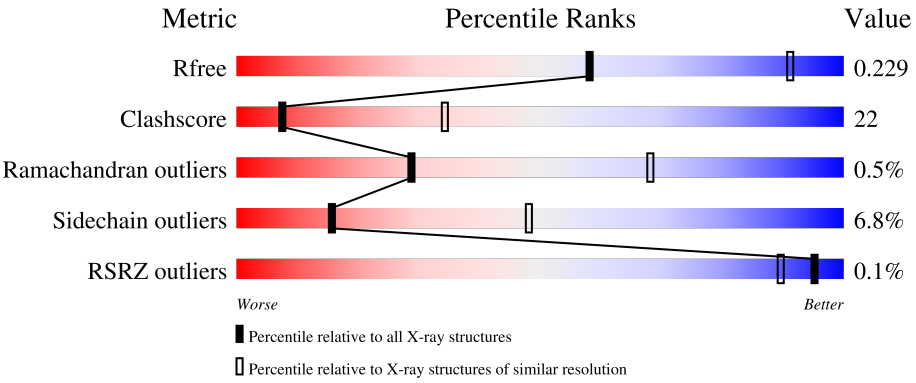
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 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 <=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 (i)

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.

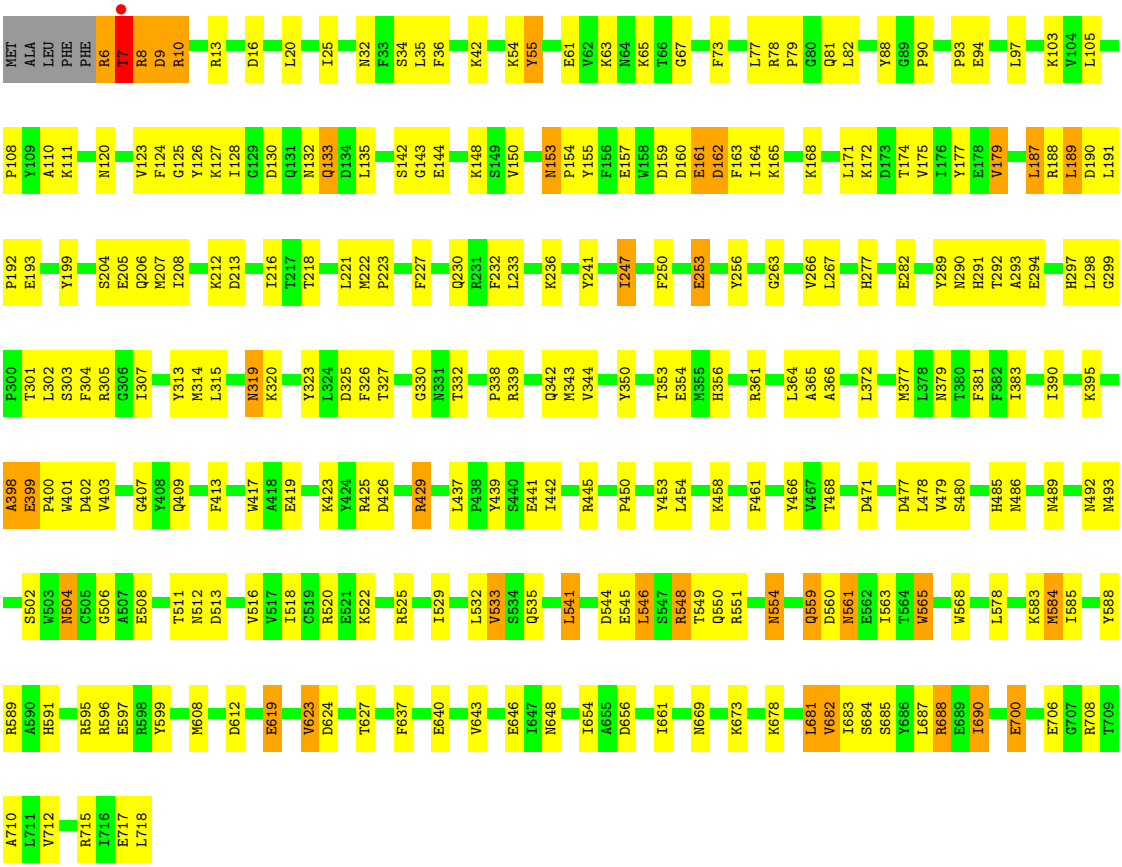
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	713	Total	C	N	O	S	0	0	0
			5826	3735	983	1089	19			
1	B	711	Total	C	N	O	S	0	0	0
			5808	3725	978	1086	19			

3 Residue-property plots (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.

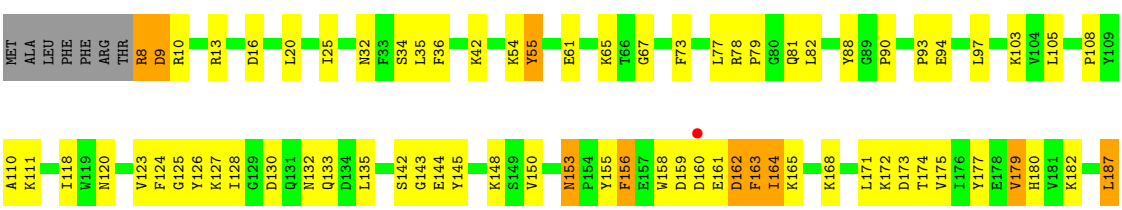
• Molecule 1: GLYCOGEN OPERON PROTEIN GLGX

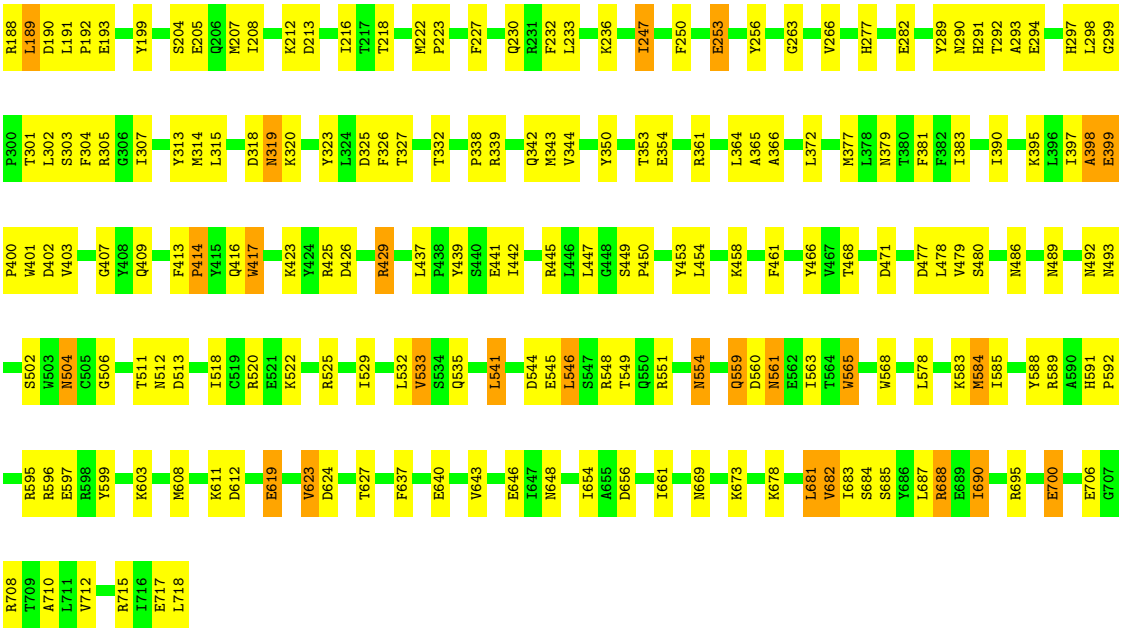
Chain A:



• Molecule 1: GLYCOGEN OPERON PROTEIN GLGX

Chain B:




R188 L189 D190 L191 P192 E193 Y199 S204 E205 Q206 M207 I208 K212 D213 I216 T217 T218 M222 P223 F227 Q230 R231 F232 L233 K236 I247 F250 E253 Y256 G263 V266 H277 E282 Y289 N290 H291 T292 A293 E294 H297 L298 G299
P300 T301 L302 S303 F304 R305 G306 I307 Y313 M314 L315 D318 N319 K320 Y323 L324 D325 F326 T327 T332 P338 R339 Q342 M343 V344 Y350 T353 E354 R361 L364 A365 A366 L372 M377 L378 N379 T380 F381 F382 I383 I390 K395 L396 I397 A398 E399
P400 W401 D402 V403 G407 Y408 Q409 F413 P414 Y415 Q416 W417 K423 Y424 R425 D426 R429 L437 P438 Y439 S440 E441 I442 R445 L446 L447 G448 S449 P450 Y453 L454 K458 F461 Y466 V467 T468 D471 D477 L478 V479 S480 N486 N489 N492 N493
S502 W503 N504 C505 G506 T511 N512 D513 I518 C519 R520 E521 K522 R525 I529 L532 V533 S534 Q535 L541 D544 E545 L546 S547 R548 T549 Q550 R551 N554 Q559 D560 N561 E562 I563 T564 W565 W568 L578 K583 M584 I585 Y588 R589 A590 H591 P592
R595 R596 E597 R598 Y599 K603 M608 K611 D612 E619 V623 D624 T627 F637 E640 V643 E646 I647 N648 I654 A655 D656 I661 N669 K673 K678 L681 V682 I683 S684 S685 Y686 L687 R688 E689 I690 R695 E700 E706 G707
R708 T709 A710 L711 V712 R715 I716 E717 L718

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	203.63Å 203.63Å 89.43Å 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
$< I/\sigma(I) >$ [1]	9.56 (at 3.01Å)	Xtriage
Refinement program		Depositor
R, R_{free}	0.206 , 0.231 0.205 , 0.229	Depositor DCC
R_{free} test set	2004 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å^2)	39.7	Xtriage
Anisotropy	0.119	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 28.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.024 for -h,-k,l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	11634	wwPDB-VP
Average B, all atoms (Å^2)	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.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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	0.33	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:O	1:B:8:ARG:HG2	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%)	34 72
1	B	709/718 (99%)	643 (91%)	62 (9%)	4 (1%)	25 64
All	All	1420/1436 (99%)	1287 (91%)	126 (9%)	7 (0%)	29 68

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 (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	629/636 (99%)	585 (93%)	44 (7%)	15	47
1	B	627/636 (99%)	586 (94%)	41 (6%)	17	50
All	All	1256/1272 (99%)	1171 (93%)	85 (7%)	16	48

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

Mol	Chain	Res	Type
1	B	247	ILE
1	B	545	GLU
1	B	318	ASP
1	B	471	ASP
1	B	559	GLN

Sometimes 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	668	ASN
1	B	601	GLN
1	B	153	ASN
1	B	561	ASN
1	B	523	GLN

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

There are no ligands in this entry.

5.7 Other polymers

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

6.1 Protein, DNA and RNA chains

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^{th} 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	713/718 (99%)	-0.57	1 (0%)	95	89	16, 32, 53, 100	0
1	B	711/718 (99%)	-0.50	1 (0%)	95	89	13, 34, 58, 83	0
All	All	1424/1436 (99%)	-0.54	2 (0%)	95	89	13, 33, 55, 100	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	7	THR	4.1
1	B	160	ASP	2.5

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.