



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 14, 2023 – 05:21 pm GMT

PDB ID : 2YFN
Title : galactosidase domain of alpha-galactosidase-sucrose kinase, AgaSK
Authors : Sulzenbacher, G.; Bruel, L.; Tison-Cervera, M.; Pujol, A.; Nicoletti, C.; Perrier, J.; Galinier, A.; Ropartz, D.; Fons, M.; Pompeo, F.; Giardina, T.
Deposited on : 2011-04-07
Resolution : 1.45 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 : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : **FAILED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

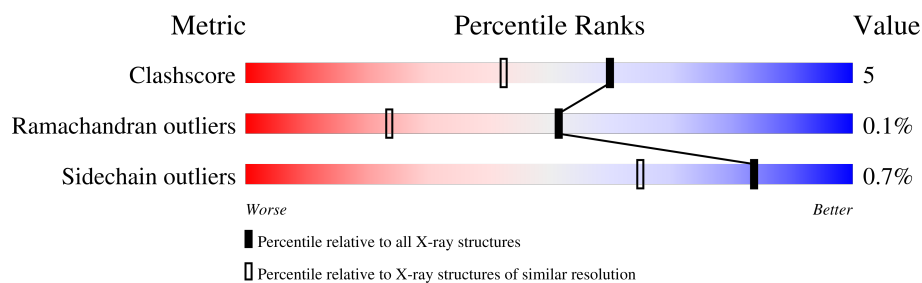
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.45 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	1202 (1.46-1.46)
Ramachandran outliers	138981	1178 (1.46-1.46)
Sidechain outliers	138945	1178 (1.46-1.46)

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 7071 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 ALPHA-GALACTOSIDASE-SUCROSE KINASE AGASK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	719	5988	3811	1004	1140	33	0	48	0

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
2	A	1	8	4	4	0	1
2	A	1	6	3	3	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Mg	0	0
			2	2		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1058	Total	O	0	0
			1058	1058		

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	105.06Å 111.65Å 154.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.19 – 1.45	Depositor
% Data completeness (in resolution range)	98.8 (36.19-1.45)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.88 (at 1.45Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.130 , 0.150	Depositor
Wilson B-factor (Å ²)	9.7	Xtriage
Anisotropy	0.192	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7071	wwPDB-VP
Average B, all atoms (Å ²)	13.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.92% 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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4, EDO, MG, GOL

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	0.65	0/6268	0.76	5/8485 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	577	ARG	NE-CZ-NH1	10.15	125.37	120.30
1	A	577	ARG	NE-CZ-NH2	-6.95	116.82	120.30
1	A	26	ASP	CB-CG-OD1	5.76	123.48	118.30
1	A	45	ASP	CB-CG-OD1	5.60	123.34	118.30
1	A	373	ARG	NE-CZ-NH2	-5.01	117.79	120.30

There are no chirality outliers.

There are no planarity outliers.

4.2 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	5988	0	5897	55	0
2	A	14	0	16	0	0
3	A	4	0	6	0	0
4	A	5	0	0	0	0
5	A	2	0	0	0	0
6	A	1058	0	0	40	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7071	0	5919	55	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 5.

The worst 5 of 55 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:321[A]:ARG:NH2	6:A:2574:HOH:O	1.61	1.28
1:A:320[B]:CYS:SG	6:A:2815:HOH:O	1.93	1.23
1:A:362[A]:MET:HE1	6:A:2241:HOH:O	1.29	1.23
1:A:663[B]:GLU:OE2	6:A:2970:HOH:O	1.60	1.19
1:A:674[B]:ARG:NH1	6:A:2992:HOH:O	1.74	1.18

There are no symmetry-related clashes.

4.3 Torsion angles [i](#)

4.3.1 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	765/720 (106%)	747 (98%)	17 (2%)	1 (0%)	51 24

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	225	SER

4.3.2 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	655/608 (108%)	650 (99%)	5 (1%)	81 62

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	156[A]	ASN
1	A	156[B]	ASN
1	A	194	ARG
1	A	251	PHE
1	A	479	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

4.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

4.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

EDS failed to run properly - this section is therefore empty.

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

EDS failed to run properly - this section is therefore empty.

5.3 Carbohydrates

EDS failed to run properly - this section is therefore empty.

5.4 Ligands

EDS failed to run properly - this section is therefore empty.

5.5 Other polymers

EDS failed to run properly - this section is therefore empty.