



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2026 – 04:49 AM UTC

PDB ID : 7VFK / pdb_00007vfk
Title : Crystal structure of SdgB (ligand-free form)
Authors : Kim, D.-G.; Baek, I.; Lee, Y.; Kim, H.S.
Deposited on : 2021-09-13
Resolution : 1.84 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
Mogul : 2022.3.0, CSD as543be (2022)
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

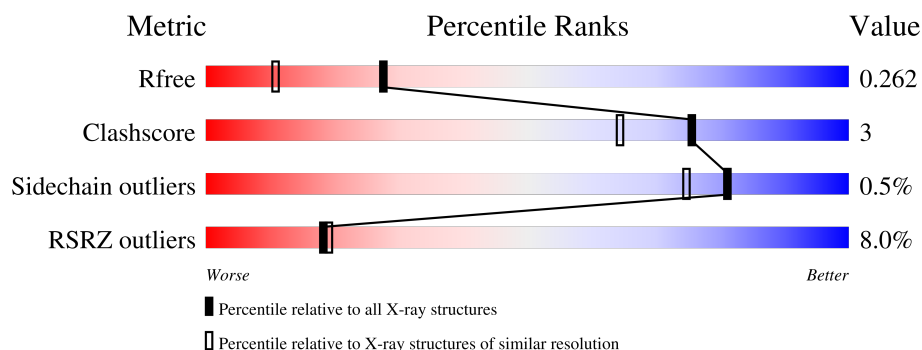
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.84 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	1296 (1.84-1.84)
Clashscore	190562	1329 (1.84-1.84)
Sidechain outliers	187428	1318 (1.84-1.84)
RSRZ outliers	180081	1296 (1.84-1.84)

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	505	 4% 92% 6% •
1	B	505	 11% 89% 9% •

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8681 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 Glycosyl transferase, group 1 family protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	499	Total	C	N	O	S	0	0	0
			4160	2686	693	770	11			
1	B	492	Total	C	N	O	S	0	0	0
			4103	2649	683	760	11			

There are 18 discrepancies between the modelled and reference sequences:

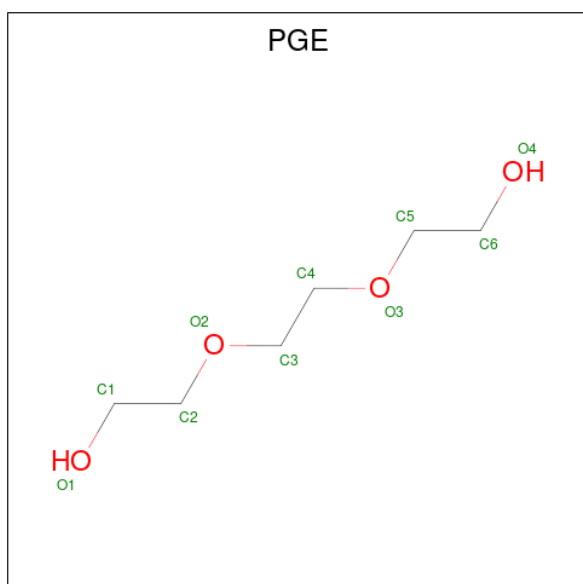
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP A0A0H2XGN0
A	497	LEU	-	expression tag	UNP A0A0H2XGN0
A	498	GLU	-	expression tag	UNP A0A0H2XGN0
A	499	HIS	-	expression tag	UNP A0A0H2XGN0
A	500	HIS	-	expression tag	UNP A0A0H2XGN0
A	501	HIS	-	expression tag	UNP A0A0H2XGN0
A	502	HIS	-	expression tag	UNP A0A0H2XGN0
A	503	HIS	-	expression tag	UNP A0A0H2XGN0
A	504	HIS	-	expression tag	UNP A0A0H2XGN0
B	0	SER	-	expression tag	UNP A0A0H2XGN0
B	497	LEU	-	expression tag	UNP A0A0H2XGN0
B	498	GLU	-	expression tag	UNP A0A0H2XGN0
B	499	HIS	-	expression tag	UNP A0A0H2XGN0
B	500	HIS	-	expression tag	UNP A0A0H2XGN0
B	501	HIS	-	expression tag	UNP A0A0H2XGN0
B	502	HIS	-	expression tag	UNP A0A0H2XGN0
B	503	HIS	-	expression tag	UNP A0A0H2XGN0
B	504	HIS	-	expression tag	UNP A0A0H2XGN0

- Molecule 2 is GLYCEROL (CCD ID: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 3 is TRIETHYLENE GLYCOL (CCD ID: PGE) (formula: $C_6H_{14}O_4$).

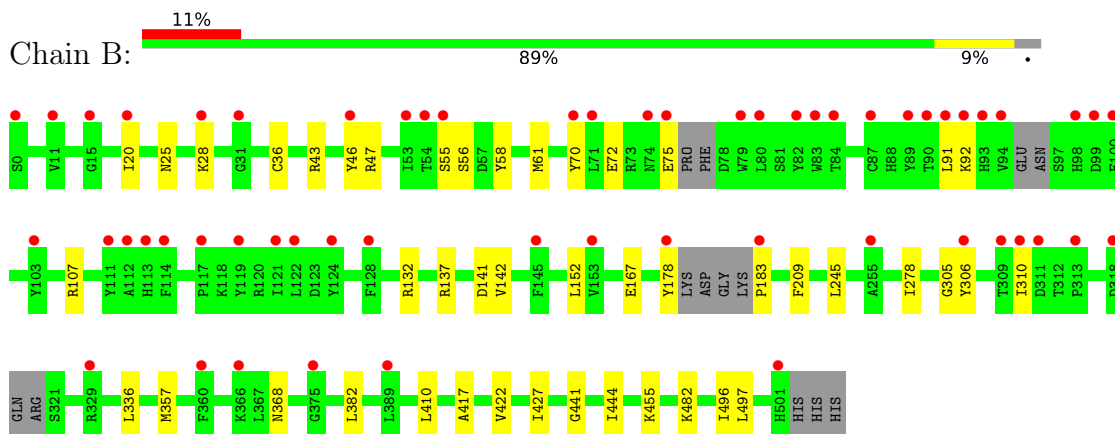


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			10	6	4		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	266	Total	O	0	0
			266	266		
4	B	112	Total	O	0	0
			112	112		

- Molecule 1: Glycosyl transferase, group 1 family protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	42.97Å 206.34Å 66.33Å 90.00° 105.06° 90.00°	Depositor
Resolution (Å)	29.26 – 1.84 29.26 – 1.84	Depositor EDS
% Data completeness (in resolution range)	98.0 (29.26-1.84) 97.9 (29.26-1.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.66 (at 1.84Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.217 , 0.259 (Not available) , 0.262	Depositor DCC
R_{free} test set	4712 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	32.3	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 38.5	EDS
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
F_o, F_c correlation	0.95	EDS
Total number of atoms	8681	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.43% 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

5.1 Standard geometry

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

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	0.39	0/4267	0.58	0/5774
1	B	0.33	0/4205	0.54	0/5686
All	All	0.36	0/8472	0.56	0/11460

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	4160	0	4065	23	0
1	B	4103	0	4006	27	0
2	A	18	0	24	1	0
2	B	12	0	16	0	0
3	B	10	0	14	0	0
4	A	266	0	0	3	0
4	B	112	0	0	1	0
All	All	8681	0	8125	49	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 3.

The worst 5 of 49 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:368:ASN:HD22	1:B:382:LEU:HD12	1.50	0.77
1:B:92:LYS:NZ	4:B:701:HOH:O	2.22	0.70
1:B:61:MET:HG3	1:B:209:PHE:CG	2.31	0.65
1:B:245:LEU:HD13	1:B:278:ILE:HG23	1.81	0.62
1:A:101:ARG:HH21	1:A:132:ARG:HH22	1.48	0.61

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

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	460/466 (99%)	457 (99%)	3 (1%)	76	68
1	B	454/466 (97%)	452 (100%)	2 (0%)	84	78
All	All	914/932 (98%)	909 (100%)	5 (0%)	81	75

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

Mol	Chain	Res	Type
1	A	86	GLU
1	A	245	LEU
1	A	427	ILE
1	B	75	GLU
1	B	427	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10

such sidechains are listed below:

Mol	Chain	Res	Type
1	B	372	GLN
1	B	379	HIS
1	B	466	GLN
1	B	35	GLN
1	B	50	GLN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

6 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
2	GOL	A	603	-	5,5,5	0.93	0	5,5,5	0.90	0
2	GOL	B	601	-	5,5,5	0.82	0	5,5,5	0.87	0
2	GOL	A	602	-	5,5,5	1.28	0	5,5,5	0.75	0
3	PGE	B	603	-	9,9,9	0.30	0	8,8,8	0.60	0
2	GOL	A	601	-	5,5,5	0.93	0	5,5,5	0.88	0
2	GOL	B	602	-	5,5,5	1.03	0	5,5,5	1.15	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	GOL	A	603	-	-	2/4/4/4	-
2	GOL	B	601	-	-	2/4/4/4	-
2	GOL	A	602	-	-	2/4/4/4	-
3	PGE	B	603	-	-	4/7/7/7	-
2	GOL	A	601	-	-	2/4/4/4	-
2	GOL	B	602	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	601	GOL	O1-C1-C2-C3
2	A	603	GOL	O1-C1-C2-C3
3	B	603	PGE	O2-C3-C4-O3
2	B	602	GOL	O2-C2-C3-O3
3	B	603	PGE	C4-C3-O2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	602	GOL	1	0

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	499/505 (98%)	0.49	22 (4%) 39 42	20, 33, 58, 81	0
1	B	492/505 (97%)	0.82	57 (11%) 9 9	25, 39, 64, 87	0
All	All	991/1010 (98%)	0.65	79 (7%) 18 19	20, 36, 62, 87	0

The worst 5 of 79 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	114	PHE	4.7
1	A	350	VAL	4.7
1	B	94	VAL	4.6
1	B	15	GLY	4.2
1	B	178	TYR	3.6

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
3	PGE	B	603	10/10	0.70	0.15	36,40,47,50	0
2	GOL	A	603	6/6	0.82	0.14	40,46,49,56	0
2	GOL	B	602	6/6	0.87	0.10	42,45,49,54	0
2	GOL	B	601	6/6	0.87	0.11	40,44,49,49	0
2	GOL	A	601	6/6	0.90	0.09	36,38,42,46	0
2	GOL	A	602	6/6	0.92	0.08	26,32,39,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.