



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 10:16 am GMT

PDB ID : 1GNX
Title : B-GLUCOSIDASE FROM STREPTOMYCES SP
Authors : Guasch, A.; Perez-Pons, J.A.; Vallmitjana, M.; Querol, E.; Coll, M.
Deposited on : 2001-10-10
Resolution : 1.68 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at 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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

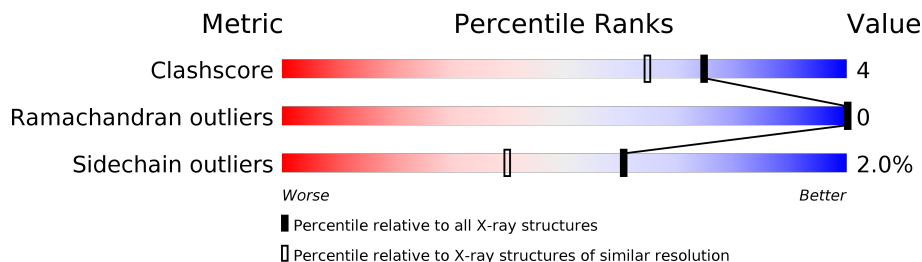
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.68 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	112137	5803 (1.70-1.66)
Ramachandran outliers	110173	5704 (1.70-1.66)
Sidechain outliers	110143	5703 (1.70-1.66)

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 ≥ 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	479	
1	B	479	

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 7617 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.

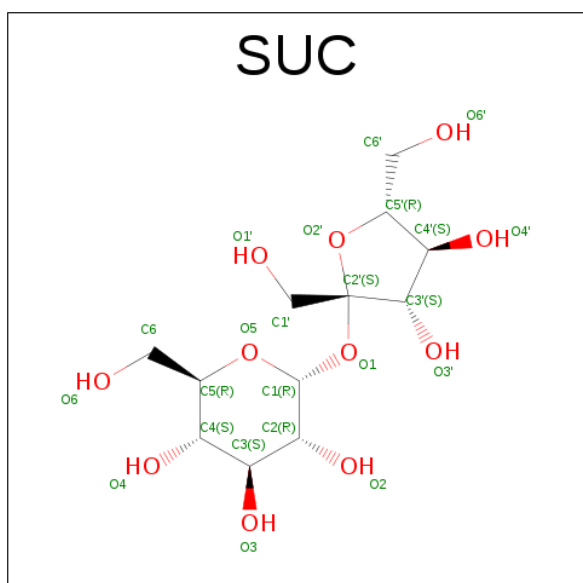
- Molecule 1 is a protein called BETA-GLUCOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	448	Total	C	N	O	S	75	0	1
			3491	2214	628	644	5			
1	B	448	Total	C	N	O	S	58	0	1
			3493	2215	628	645	5			

There are 6 discrepancies between the modelled and reference sequences:

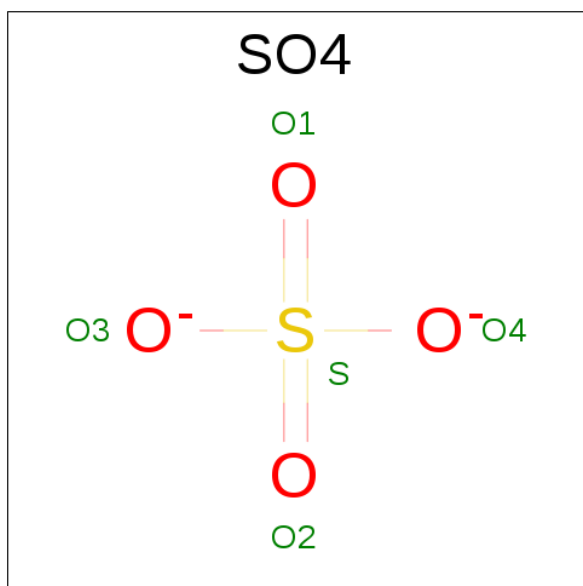
Chain	Residue	Modelled	Actual	Comment	Reference
A	143	ALA	PRO	CONFLICT	UNP Q59976
A	150	ALA	PRO	CONFLICT	UNP Q59976
A	246	ASP	GLU	CONFLICT	UNP Q59976
B	143	ALA	PRO	CONFLICT	UNP Q59976
B	150	ALA	PRO	CONFLICT	UNP Q59976
B	246	ASP	GLU	CONFLICT	UNP Q59976

- Molecule 2 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: C₁₂H₂₂O₁₁).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	1	0
			23	12	11		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		
3	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 4 is water.


Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	302	Total	O	0	0
			302	302		
4	B	293	Total	O	0	0
			293	293		

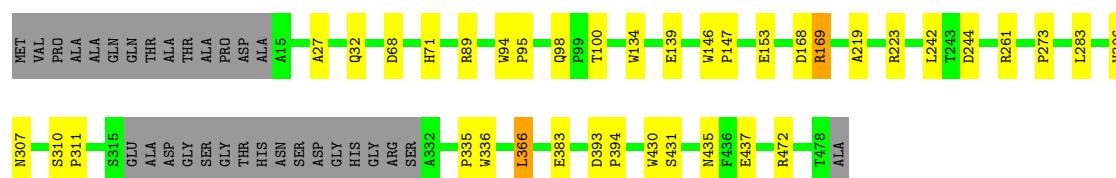
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.


Note EDS was not executed.

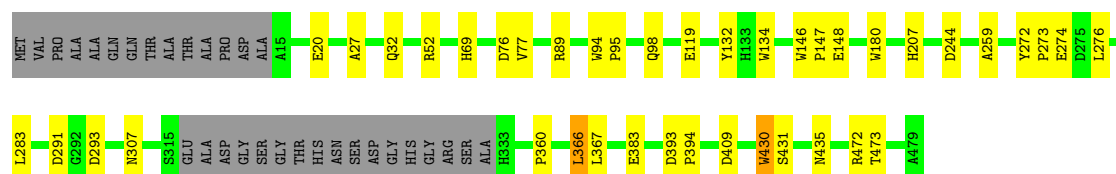
• Molecule 1: BETA-GLUCOSIDASE

Chain A: 



• Molecule 1: BETA-GLUCOSIDASE

Chain B: 



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	96.70Å 111.28Å 185.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 1.68	Depositor
% Data completeness (in resolution range)	82.0 (40.00-1.68)	Depositor
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.190 , 0.210	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	7617	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: SUC, SO4

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	0/3596	0.61	1/4916 (0.0%)
1	B	0.31	0/3598	0.61	1/4919 (0.0%)
All	All	0.32	0/7194	0.61	2/9835 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	430	TRP	N-CA-C	-5.72	95.56	111.00
1	B	430	TRP	N-CA-C	-5.59	95.89	111.00

There are no chirality outliers.

There are no planarity outliers.

5.2 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	3491	0	3328	24	0
1	B	3493	0	3330	26	1
2	A	23	0	22	0	0
3	A	5	0	0	0	0
3	B	10	0	0	0	0
4	A	302	0	0	2	0
4	B	293	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	7617	0	6680	50	1

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 4.

All (50) 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:94:TRP:HE1	1:A:98:GLN:HE21	1.35	0.75
1:B:69:HIS:HD2	1:B:76:ASP:OD2	1.78	0.67
1:B:307:ASN:ND2	1:B:383:GLU:HB2	2.10	0.67
1:B:94:TRP:HE1	1:B:98:GLN:HE21	1.43	0.65
1:A:273:PRO:HG3	4:A:2172:HOH:O	1.96	0.64
1:B:307:ASN:CG	1:B:383:GLU:HB2	2.19	0.62
1:B:148:GLU:HG3	4:B:2128:HOH:O	2.00	0.61
1:A:68:ASP:OD2	1:A:71:HIS:HD2	1.85	0.59
1:A:169:ARG:HD2	4:A:2128:HOH:O	2.03	0.58
1:A:307:ASN:CG	1:A:383:GLU:HB2	2.24	0.58
1:A:94:TRP:HE1	1:A:98:GLN:NE2	2.01	0.57
1:A:366:LEU:HD23	1:A:366:LEU:C	2.28	0.54
1:A:307:ASN:ND2	1:A:383:GLU:HB2	2.24	0.52
1:B:27:ALA:O	1:B:431:SER:HA	2.08	0.52
1:A:27:ALA:O	1:A:431:SER:HA	2.10	0.51
1:B:94:TRP:N	1:B:95:PRO:CD	2.75	0.50
1:B:366:LEU:C	1:B:366:LEU:HD23	2.31	0.50
1:B:360:PRO:HG2	1:B:409:ASP:HB3	1.94	0.50
1:B:273:PRO:HG2	1:B:276:LEU:HB3	1.94	0.49
1:B:32:GLN:O	1:B:435:ASN:HB2	2.12	0.49
1:B:207:HIS:HD2	1:B:272:TYR:OH	1.95	0.49
1:A:393:ASP:HB2	1:A:394:PRO:CD	2.43	0.49
1:A:94:TRP:HB3	1:A:95:PRO:HD3	1.94	0.49
1:A:89:ARG:NH1	1:A:383:GLU:HG3	2.29	0.48
1:B:180:TRP:CD1	1:B:259:ALA:HA	2.49	0.47
1:A:261:ARG:HB3	1:A:273:PRO:HG2	1.97	0.47
1:B:89:ARG:NH1	1:B:383:GLU:HG3	2.30	0.46
1:A:134:TRP:CD1	1:A:134:TRP:N	2.83	0.46
1:B:134:TRP:CD1	1:B:134:TRP:N	2.82	0.46
1:B:207:HIS:HE1	1:B:293:ASP:OD2	1.99	0.46
1:B:146:TRP:N	1:B:147:PRO:CD	2.79	0.45
1:B:472:ARG:NH1	1:B:473:THR:CG2	2.80	0.45
1:B:94:TRP:HB3	1:B:95:PRO:HD3	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:32:GLN:O	1:A:435:ASN:HB2	2.16	0.45
1:A:94:TRP:N	1:A:95:PRO:CD	2.80	0.44
1:A:335:PRO:HG2	1:A:336:TRP:CD1	2.52	0.44
1:B:393:ASP:HB2	1:B:394:PRO:CD	2.48	0.44
1:B:430:TRP:CE2	1:B:431:SER:HB2	2.53	0.44
1:A:219:ALA:O	1:A:223:ARG:HG3	2.18	0.44
1:A:437:GLU:O	1:A:437:GLU:HG3	2.18	0.43
1:A:306:VAL:HG11	1:A:366:LEU:HD21	2.01	0.42
1:A:310:SER:OG	1:A:311:PRO:HD2	2.20	0.42
1:B:366:LEU:HD23	1:B:367:LEU:N	2.34	0.42
1:B:94:TRP:HE1	1:B:98:GLN:NE2	2.15	0.41
1:A:100:THR:HA	1:A:139:GLU:OE2	2.21	0.41
1:B:77:VAL:HG11	1:B:119:GLU:HG3	2.03	0.41
1:A:146:TRP:N	1:A:147:PRO:CD	2.85	0.40
1:B:132:TYR:CZ	1:B:134:TRP:HA	2.57	0.40
1:B:273:PRO:HG2	1:B:276:LEU:CB	2.50	0.40
1:A:393:ASP:HB2	1:A:394:PRO:HD2	2.02	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:244:ASP:OD1	1:B:244:ASP:OD1[3_655]	2.04	0.16

5.3 Torsion angles [i](#)

5.3.1 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	444/479 (93%)	433 (98%)	11 (2%)	0	100	100
1	B	444/479 (93%)	434 (98%)	10 (2%)	0	100	100
All	All	888/958 (93%)	867 (98%)	21 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	346/367 (94%)	338 (98%)	8 (2%)	56	33
1	B	347/367 (95%)	341 (98%)	6 (2%)	66	47
All	All	693/734 (94%)	679 (98%)	14 (2%)	60	40

All (14) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	153	GLU
1	A	168	ASP
1	A	169	ARG
1	A	242	LEU
1	A	244	ASP
1	A	283	LEU
1	A	366	LEU
1	A	472	ARG
1	B	20	GLU
1	B	52	ARG
1	B	274	GLU
1	B	283	LEU
1	B	291	ASP
1	B	366	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	59	ASN
1	A	71	HIS
1	A	98	GLN
1	A	138	GLN
1	A	142	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	237	HIS
1	A	307	ASN
1	A	346	GLN
1	B	69	HIS
1	B	98	GLN
1	B	207	HIS
1	B	210	ASN
1	B	218	GLN
1	B	299	GLN
1	B	307	ASN
1	B	346	GLN
1	B	384	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](#)

4 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	SUC	A	600	-	24,24,24	1.09	1 (4%)	36,36,36	2.08	13 (36%)
3	SO4	A	610	-	4,4,4	0.45	0	6,6,6	0.07	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
3	SO4	B	620	-	4,4,4	0.49	0	6,6,6	0.08	0
3	SO4	B	630	-	4,4,4	0.46	0	6,6,6	0.05	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	SUC	A	600	-	-	0/12/51/51	0/2/2/2
3	SO4	A	610	-	-	0/0/0/0	0/0/0/0
3	SO4	B	620	-	-	0/0/0/0	0/0/0/0
3	SO4	B	630	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	600	SUC	O5-C1	-2.47	1.35	1.41

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	600	SUC	O2-C2-C1	-3.25	103.23	110.03
2	A	600	SUC	O2'-C2'-C3'	-2.23	100.73	105.57
2	A	600	SUC	O2'-C2'-C1'	-2.13	102.53	108.03
2	A	600	SUC	C3-C4-C5	-2.10	106.51	110.22
2	A	600	SUC	C6'-C5'-C4'	-2.06	110.06	115.05
2	A	600	SUC	O1-C2'-C3'	2.31	115.80	108.11
2	A	600	SUC	C2'-C3'-C4'	2.36	107.75	102.06
2	A	600	SUC	O2'-C5'-C6'	2.45	115.79	108.71
2	A	600	SUC	O1-C2'-O2'	3.26	120.58	110.55
2	A	600	SUC	O5-C5-C6	3.95	115.87	106.41
2	A	600	SUC	C2'-O1-C1	4.30	129.09	117.62
2	A	600	SUC	C1-O5-C5	4.39	121.99	113.72
2	A	600	SUC	O2-C2-C3	5.54	122.42	110.36

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.