



wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 26, 2023 – 02:25 PM EDT

PDB ID : 6D8G
Title : D341A D367A calcium binding mutant of Bacteroides uniformis beta-glucuronidase 2
Authors : Walton, W.G.; Pellock, S.J.; Redinbo, M.R.
Deposited on : 2018-04-26
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 ⓘ 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 : 2.35.1
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) : 2.35.1

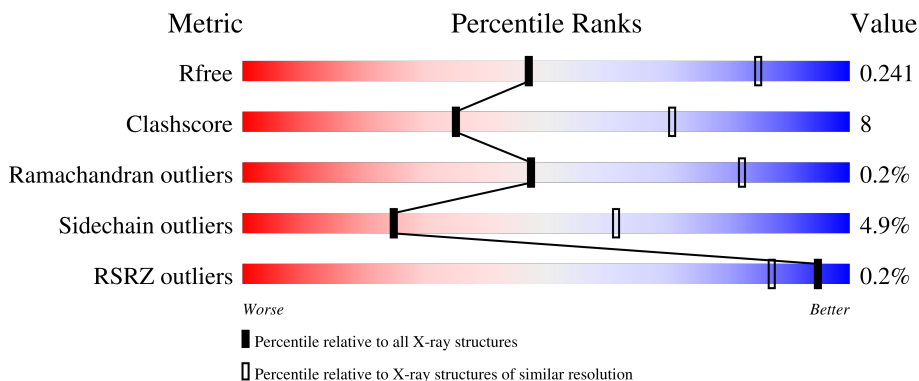
1 Overall quality at a glance

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}	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 $\leq 5\%$. 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	886	 72% 20% • 6%
1	B	886	 74% 18% • 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13437 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 hydrolases family 2, sugar binding domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	832	6723	4281	1157	1265	20	0	0	0
1	B	831	6712	4275	1153	1264	20	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	341	ALA	ASP	engineered mutation	UNP A0A078SUX9
A	367	ALA	ASP	engineered mutation	UNP A0A078SUX9
B	341	ALA	ASP	engineered mutation	UNP A0A078SUX9
B	367	ALA	ASP	engineered mutation	UNP A0A078SUX9

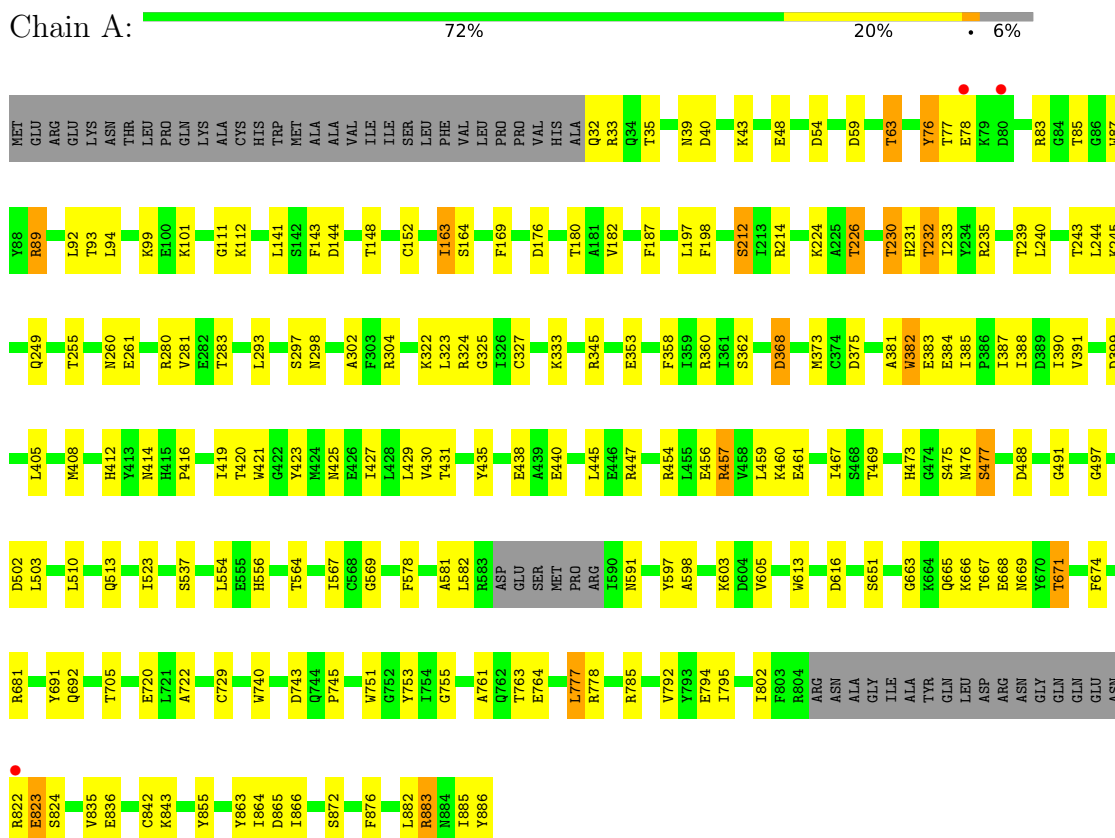
- Molecule 2 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Na	0	0
			1	1		
2	B	1	Total	Na	0	0
			1	1		

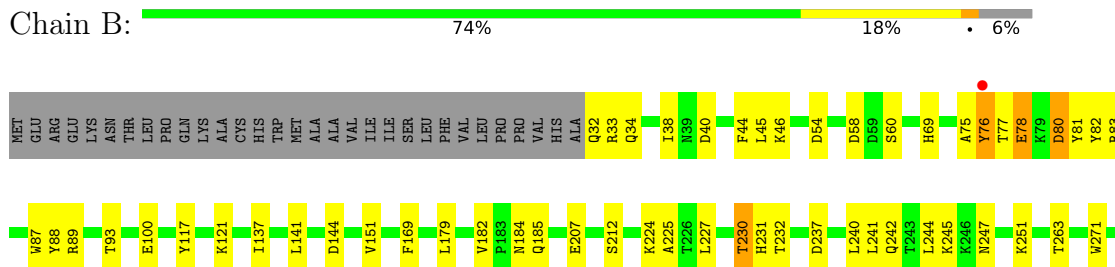
3 Residue-property plots i

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: Glycosyl hydrolases family 2, sugar binding domain protein



- Molecule 1: Glycosyl hydrolases family 2, sugar binding domain protein



P276	R277	L278	N287	R288	K289	T290	L293	L294	R304	K306	F315	P320	G325	I326	H329	E342	K343	H344	R345	M351	F388	H363	L371	E372	M373	K376	K382	P386	I387	E402	Y413	I418	I419	T420	W421	G422	Y423	E426									
Y435	K436	T437	E440	L441	L445	L449	R454	R457	V458	L459	K460	R466	I467	S468	T469	F472	H473	G474	S475	Y478	L483	T487	D488	I489	G497	L503	P521	G530	S631	I547	T564	I567	C568	G569	G570	T571	H572	W573	M574	F575	A581						
L582	ARG	ASP	GLU	SER	MET	PRO	ARG	I590	K593	R600	Q610	K615	D616	V619	T628	T628	G636	Y641	N669	Y670	T671	P677	F678	S679	R701	I702	E720	L721	A722	W740	L741	P742	G749	S750	W751	T760	A761	Q762	T763	E764	I765	Q766	F774	Q775			
T776	L777	R785	F786	D787	E794	T800	D801	I802	F803	R804	ARG	ASN	ALA	GLY	ILE	ALA	TYR	GLN	LEU	ASP	ARG	ASN	GLY	GLN	GLN	GLU	ASN	R822	E823	S824	I828	R849	A850	K853	K854	Y855	Y856	I857	I864	S870	T871	S872	G873	T874	K881	I885	Y886

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	74.99Å 142.02Å 181.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.55 – 3.00 29.55 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.55-3.00) 100.0 (29.55-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.82 (at 3.00Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575)	Depositor
R, R_{free}	0.166 , 0.241 0.166 , 0.241	Depositor DCC
R_{free} test set	2000 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	67.7	Xtrriage
Anisotropy	0.459	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 34.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	13437	wwPDB-VP
Average B, all atoms (Å ²)	57.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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 [i](#)

5.1 Standard geometry [i](#)

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

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	0.45	0/6900	0.61	1/9366 (0.0%)
1	B	0.45	1/6889 (0.0%)	0.60	1/9352 (0.0%)
All	All	0.45	1/13789 (0.0%)	0.61	2/18718 (0.0%)

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	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	144	ASP	C-N	6.65	1.49	1.34

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	382	TRP	CA-CB-CG	8.12	129.13	113.70
1	B	382	TRP	CA-CB-CG	6.50	126.05	113.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	381	ALA	Peptide

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	6723	0	6499	116	0
1	B	6712	0	6486	104	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
All	All	13437	0	12985	217	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 8.

The worst 5 of 217 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:382:TRP:HB3	1:B:420:THR:HG23	1.42	0.99
1:B:54:ASP:O	1:B:89:ARG:NH2	1.98	0.95
1:A:382:TRP:HB3	1:A:420:THR:HG23	1.60	0.83
1:B:247:ASN:HB2	1:B:288:ARG:HH21	1.45	0.81
1:B:413:TYR:O	1:B:466:ARG:NH2	2.13	0.80

There are no symmetry-related clashes.

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	826/886 (93%)	765 (93%)	61 (7%)	0	100 100
1	B	825/886 (93%)	769 (93%)	53 (6%)	3 (0%)	34 72

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1651/1772 (93%)	1534 (93%)	114 (7%)	3 (0%)	47 82

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	573	TRP
1	B	636	GLY
1	B	386	PRO

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	719/765 (94%)	688 (96%)	31 (4%)	29 66
1	B	718/765 (94%)	679 (95%)	39 (5%)	22 57
All	All	1437/1530 (94%)	1367 (95%)	70 (5%)	25 61

5 of 70 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	616	ASP
1	B	701	ARG
1	B	857	ILE
1	A	777	LEU
1	A	743	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	432	GLN
1	B	479	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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 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.

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	832/886 (93%)	-0.54	3 (0%) 92 79	34, 54, 78, 117	0
1	B	831/886 (93%)	-0.47	1 (0%) 95 89	41, 58, 83, 108	0
All	All	1663/1772 (93%)	-0.51	4 (0%) 95 87	34, 56, 81, 117	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	822	ARG	2.5
1	A	78	GLU	2.4
1	B	76	TYR	2.3
1	A	80	ASP	2.1

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 monosaccharides 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(Å ²)	Q<0.9
2	NA	A	901	1/1	0.94	0.28	39,39,39,39	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	NA	B	901	1/1	0.99	0.35	39,39,39,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.