



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:23 AM GMT

PDB ID : 3ANJ  
Title : Crystal structure of unsaturated glucuronyl hydrolase from *Streptococcus agalactiae*  
Authors : Nakamichi, Y.; Maruyama, Y.; Mikami, B.; Hashimoto, W.; Murata, K.  
Deposited on : 2010-09-02  
Resolution : 1.95 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](mailto: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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

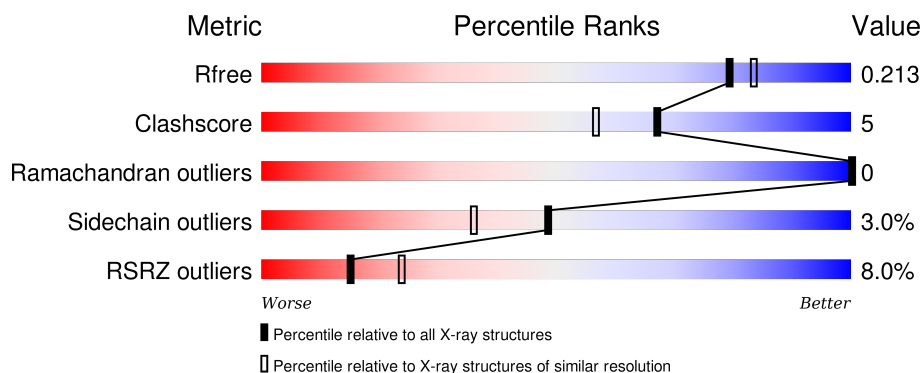
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

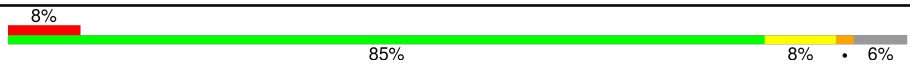
The reported resolution of this entry is 1.95 Å.

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))
$R_{free}$	91344	1833 (1.96-1.96)
Clashscore	102246	1953 (1.96-1.96)
Ramachandran outliers	100387	1936 (1.96-1.96)
Sidechain outliers	100360	1936 (1.96-1.96)
RSRZ outliers	91569	1835 (1.96-1.96)

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  $\geq 3$ , 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.

Mol	Chain	Length	Quality of chain
1	A	398	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3359 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 Putative uncharacterized protein gbs1889.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	376	Total	C	N	O	S	0	17	0
			3250	2089	540	609	12			

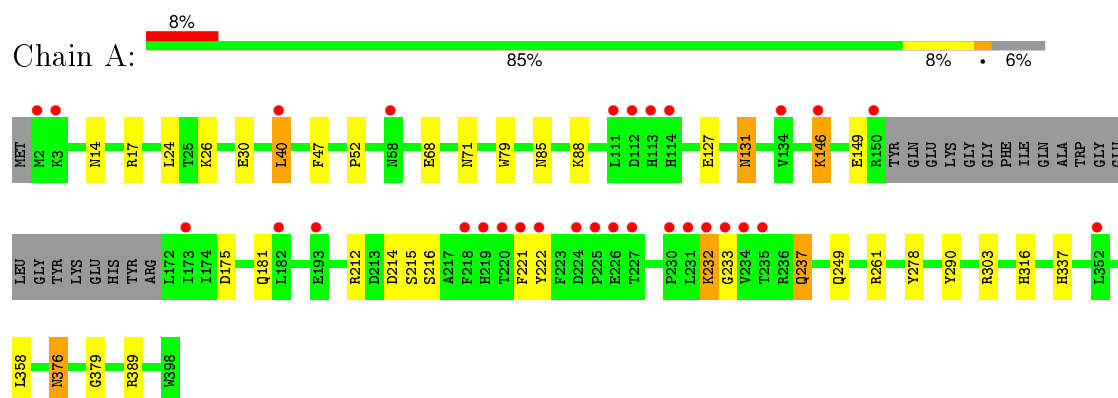
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	109	Total	O	0	0
			109	109		

### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: Putative uncharacterized protein gbs1889



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.84Å 53.19Å 70.08Å 90.00° 96.62° 90.00°	Depositor
Resolution (Å)	50.00 – 1.95 23.69 – 1.95	Depositor EDS
% Data completeness (in resolution range)	98.8 (50.00-1.95) 98.9 (23.69-1.95)	Depositor EDS
$R_{merge}$	0.01	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.68 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.189 , 0.213 0.191 , 0.213	Depositor DCC
$R_{free}$ test set	1394 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.7	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 52.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 27822 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3359	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.95% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.38	0/3334	0.48	0/4514

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 [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	3250	0	3135	31	0
2	A	109	0	0	0	0
All	All	3359	0	3135	31	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.

All (31) 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:24:LEU:H	1:A:337:HIS:HD2	1.21	0.88
1:A:316:HIS:HD2	1:A:389:ARG:HH11	1.38	0.70
1:A:68:GLU:H	1:A:71:ASN:HD21	1.40	0.68
1:A:175[B]:ASP:OD1	1:A:175[B]:ASP:N	2.27	0.67

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:232:LYS:HB3	1:A:233:GLY:HA2	1.76	0.65
1:A:14[A]:ASN:O	1:A:17[A]:ARG:HG2	1.97	0.65
1:A:376:ASN:ND2	1:A:379:GLY:H	1.96	0.64
1:A:376:ASN:HD22	1:A:379:GLY:H	1.46	0.62
1:A:24:LEU:H	1:A:337:HIS:CD2	2.11	0.58
1:A:237:GLN:HE21	1:A:237:GLN:HA	1.69	0.57
1:A:316:HIS:CD2	1:A:389:ARG:HH11	2.22	0.57
1:A:232:LYS:CB	1:A:233:GLY:HA2	2.33	0.57
1:A:175[A]:ASP:C	1:A:175[A]:ASP:OD2	2.47	0.52
1:A:127:GLU:HG2	1:A:131:ASN:HD21	1.75	0.52
1:A:47:PHE:CZ	1:A:52:PRO:HB3	2.48	0.48
1:A:181:GLN:NE2	1:A:261:ARG:HH21	2.12	0.48
1:A:40[A]:LEU:HD13	1:A:358:LEU:HD21	1.96	0.47
1:A:212:ARG:CZ	1:A:216[A]:SER:OG	2.63	0.46
1:A:216[A]:SER:HA	1:A:249:GLN:HE22	1.81	0.45
1:A:131:ASN:HD22	1:A:131:ASN:C	2.19	0.45
1:A:146:LYS:O	1:A:149:GLU:HG2	2.17	0.44
1:A:79:TRP:CZ2	1:A:127:GLU:HB2	2.54	0.43
1:A:216[A]:SER:HB3	1:A:278:TYR:CZ	2.53	0.43
1:A:214:ASP:O	1:A:215:SER:HB2	2.19	0.43
1:A:14[B]:ASN:O	1:A:17[B]:ARG:HG2	2.19	0.42
1:A:290:TYR:HA	1:A:303:ARG:HA	2.02	0.42
1:A:316:HIS:HD2	1:A:389:ARG:NH1	2.12	0.41
1:A:216[B]:SER:HA	1:A:249:GLN:HE22	1.84	0.41
1:A:216[B]:SER:HB2	1:A:278:TYR:CZ	2.56	0.41
1:A:26:LYS:O	1:A:30:GLU:HB2	2.22	0.40

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	389/398 (98%)	382 (98%)	7 (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	347/347 (100%)	336 (97%)	11 (3%)	46	33

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40[A]	LEU
1	A	40[B]	LEU
1	A	85	ASN
1	A	88	LYS
1	A	131	ASN
1	A	146	LYS
1	A	221	PHE
1	A	222	TYR
1	A	232	LYS
1	A	237	GLN
1	A	376	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	71	ASN
1	A	85	ASN
1	A	131	ASN
1	A	181	GLN
1	A	189	GLN
1	A	202	HIS
1	A	237	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	249	GLN
1	A	316	HIS
1	A	337	HIS
1	A	376	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](#)

There are no ligands in this entry.

### 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	376/398 (94%)	0.56	30 (7%) 15 24	15, 25, 45, 60	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	233	GLY	13.9
1	A	231	LEU	12.6
1	A	234	VAL	12.3
1	A	221	PHE	11.6
1	A	222	TYR	8.5
1	A	220	THR	8.4
1	A	235	THR	7.0
1	A	232	LYS	6.7
1	A	219	HIS	4.6
1	A	2	MET	3.9
1	A	226	GLU	3.8
1	A	224	ASP	3.7
1	A	230	PRO	3.4
1	A	218	PHE	3.1
1	A	352	LEU	3.1
1	A	112	ASP	3.0
1	A	113	HIS	3.0
1	A	150	ARG	2.8
1	A	134	VAL	2.4
1	A	111	LEU	2.3
1	A	40[A]	LEU	2.3
1	A	146	LYS	2.3
1	A	227	THR	2.3
1	A	182	LEU	2.2
1	A	193	GLU	2.2
1	A	114	HIS	2.2
1	A	225	PRO	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	173	ILE	2.1
1	A	58	ASN	2.1
1	A	3	LYS	2.0

## 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 carbohydrates in this entry.

## 6.4 Ligands [i](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.