



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 14, 2020 – 12:18 AM BST

PDB ID : 6D7J  
Title : The Crystal Structure of Parabacteroides merdae Beta-Glucuronidase (GUS)  
with Glycerol in Active-Site  
Authors : Little, M.S.; Redinbo, M.R.  
Deposited on : 2018-04-24  
Resolution : 2.24 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](mailto: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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.4.dev1  
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) : 2.14.4.dev1



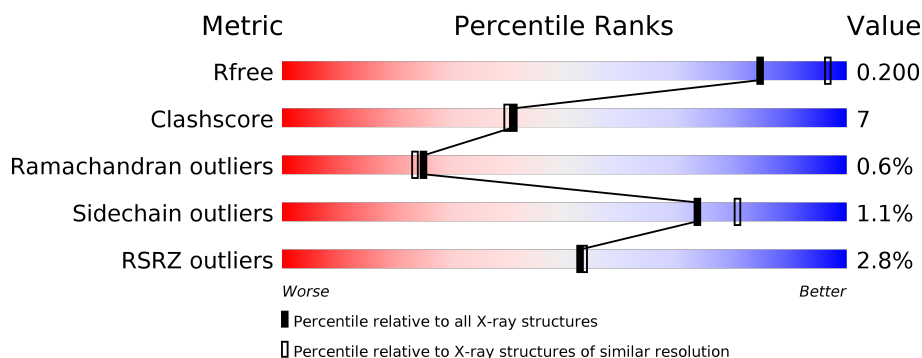
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.24 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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 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	830	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 14%, green 82%, grey 14%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>82%</span> <span>14%</span> <span>• •</span> </div> </div>
1	B	830	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 2%, orange 2%, yellow 12%, green 83%, grey 12%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>2%</span> <span>83%</span> <span>12%</span> <span>5%</span> </div> </div>
1	C	830	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, orange 1%, yellow 12%, green 82%, grey 12%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>%</span> <span>82%</span> <span>12%</span> <span>• 5%</span> </div> </div>
1	D	830	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 6%, orange 6%, yellow 14%, green 80%, grey 14%);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> <span>6%</span> <span>80%</span> <span>14%</span> <span>• 5%</span> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit crite-



ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	B	917	-	-	X	-
3	GOL	C	902	-	-	X	-
3	GOL	C	904	-	-	X	-
3	GOL	C	909	-	X	-	-
3	GOL	C	913	-	-	X	-
3	GOL	C	918	-	-	X	-
4	PRO	A	914	-	-	X	-



## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 27845 atoms, of which 442 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-Glucuronidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	799	Total	C	N	O	S	0	1	0
			6385	4054	1106	1195	30			
1	B	788	Total	C	N	O	S	0	0	0
			6287	3993	1083	1181	30			
1	C	789	Total	C	N	O	S	0	0	0
			6288	3993	1084	1180	31			
1	D	787	Total	C	N	O	S	0	0	0
			6237	3961	1076	1169	31			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	HIS	-	expression tag	UNP K5Z WV5
A	2	HIS	-	expression tag	UNP K5Z WV5
A	3	HIS	-	expression tag	UNP K5Z WV5
A	4	HIS	-	expression tag	UNP K5Z WV5
A	5	HIS	-	expression tag	UNP K5Z WV5
A	6	HIS	-	expression tag	UNP K5Z WV5
A	7	SER	-	expression tag	UNP K5Z WV5
A	8	SER	-	expression tag	UNP K5Z WV5
A	9	GLY	-	expression tag	UNP K5Z WV5
A	10	VAL	-	expression tag	UNP K5Z WV5
A	11	ASP	-	expression tag	UNP K5Z WV5
A	12	LEU	-	expression tag	UNP K5Z WV5
A	13	GLY	-	expression tag	UNP K5Z WV5
A	14	THR	-	expression tag	UNP K5Z WV5
A	15	GLU	-	expression tag	UNP K5Z WV5
A	16	ASN	-	expression tag	UNP K5Z WV5
A	17	LEU	-	expression tag	UNP K5Z WV5
A	18	TYR	-	expression tag	UNP K5Z WV5
A	19	PHE	-	expression tag	UNP K5Z WV5
A	20	GLN	-	expression tag	UNP K5Z WV5
A	21	SER	-	expression tag	UNP K5Z WV5

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	22	ASN	-	expression tag	UNP K5Z WV5
B	1	HIS	-	expression tag	UNP K5Z WV5
B	2	HIS	-	expression tag	UNP K5Z WV5
B	3	HIS	-	expression tag	UNP K5Z WV5
B	4	HIS	-	expression tag	UNP K5Z WV5
B	5	HIS	-	expression tag	UNP K5Z WV5
B	6	HIS	-	expression tag	UNP K5Z WV5
B	7	SER	-	expression tag	UNP K5Z WV5
B	8	SER	-	expression tag	UNP K5Z WV5
B	9	GLY	-	expression tag	UNP K5Z WV5
B	10	VAL	-	expression tag	UNP K5Z WV5
B	11	ASP	-	expression tag	UNP K5Z WV5
B	12	LEU	-	expression tag	UNP K5Z WV5
B	13	GLY	-	expression tag	UNP K5Z WV5
B	14	THR	-	expression tag	UNP K5Z WV5
B	15	GLU	-	expression tag	UNP K5Z WV5
B	16	ASN	-	expression tag	UNP K5Z WV5
B	17	LEU	-	expression tag	UNP K5Z WV5
B	18	TYR	-	expression tag	UNP K5Z WV5
B	19	PHE	-	expression tag	UNP K5Z WV5
B	20	GLN	-	expression tag	UNP K5Z WV5
B	21	SER	-	expression tag	UNP K5Z WV5
B	22	ASN	-	expression tag	UNP K5Z WV5
C	1	HIS	-	expression tag	UNP K5Z WV5
C	2	HIS	-	expression tag	UNP K5Z WV5
C	3	HIS	-	expression tag	UNP K5Z WV5
C	4	HIS	-	expression tag	UNP K5Z WV5
C	5	HIS	-	expression tag	UNP K5Z WV5
C	6	HIS	-	expression tag	UNP K5Z WV5
C	7	SER	-	expression tag	UNP K5Z WV5
C	8	SER	-	expression tag	UNP K5Z WV5
C	9	GLY	-	expression tag	UNP K5Z WV5
C	10	VAL	-	expression tag	UNP K5Z WV5
C	11	ASP	-	expression tag	UNP K5Z WV5
C	12	LEU	-	expression tag	UNP K5Z WV5
C	13	GLY	-	expression tag	UNP K5Z WV5
C	14	THR	-	expression tag	UNP K5Z WV5
C	15	GLU	-	expression tag	UNP K5Z WV5
C	16	ASN	-	expression tag	UNP K5Z WV5
C	17	LEU	-	expression tag	UNP K5Z WV5
C	18	TYR	-	expression tag	UNP K5Z WV5
C	19	PHE	-	expression tag	UNP K5Z WV5

*Continued on next page...*



*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	20	GLN	-	expression tag	UNP K5Z WV5
C	21	SER	-	expression tag	UNP K5Z WV5
C	22	ASN	-	expression tag	UNP K5Z WV5
D	1	HIS	-	expression tag	UNP K5Z WV5
D	2	HIS	-	expression tag	UNP K5Z WV5
D	3	HIS	-	expression tag	UNP K5Z WV5
D	4	HIS	-	expression tag	UNP K5Z WV5
D	5	HIS	-	expression tag	UNP K5Z WV5
D	6	HIS	-	expression tag	UNP K5Z WV5
D	7	SER	-	expression tag	UNP K5Z WV5
D	8	SER	-	expression tag	UNP K5Z WV5
D	9	GLY	-	expression tag	UNP K5Z WV5
D	10	VAL	-	expression tag	UNP K5Z WV5
D	11	ASP	-	expression tag	UNP K5Z WV5
D	12	LEU	-	expression tag	UNP K5Z WV5
D	13	GLY	-	expression tag	UNP K5Z WV5
D	14	THR	-	expression tag	UNP K5Z WV5
D	15	GLU	-	expression tag	UNP K5Z WV5
D	16	ASN	-	expression tag	UNP K5Z WV5
D	17	LEU	-	expression tag	UNP K5Z WV5
D	18	TYR	-	expression tag	UNP K5Z WV5
D	19	PHE	-	expression tag	UNP K5Z WV5
D	20	GLN	-	expression tag	UNP K5Z WV5
D	21	SER	-	expression tag	UNP K5Z WV5
D	22	ASN	-	expression tag	UNP K5Z WV5

- Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total K 1 1	0	0
2	A	1	Total K 1 1	0	0
2	D	1	Total K 1 1	0	0
2	C	1	Total K 1 1	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	A	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	B	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		

*Continued on next page...*

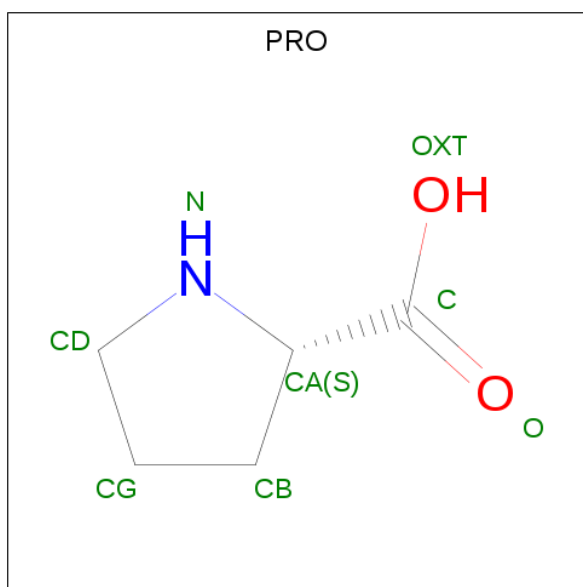


*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	C	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		
3	D	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 4 is PROLINE (three-letter code: PRO) (formula: C<sub>5</sub>H<sub>9</sub>NO<sub>2</sub>).





Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	0	0
			17	5	9	1	2		
4	D	1	Total	C	H	N	O	0	0
			17	5	9	1	2		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Na	0	0
			1	1		
5	A	1	Total	Na	0	0
			1	1		
5	D	1	Total	Na	0	0
			1	1		
5	C	1	Total	Na	0	0
			1	1		

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	560	Total	O	0	0
			560	560		
6	B	493	Total	O	0	0
			493	493		
6	C	445	Total	O	0	0
			445	445		

*Continued on next page...*



*Continued from previous page...*

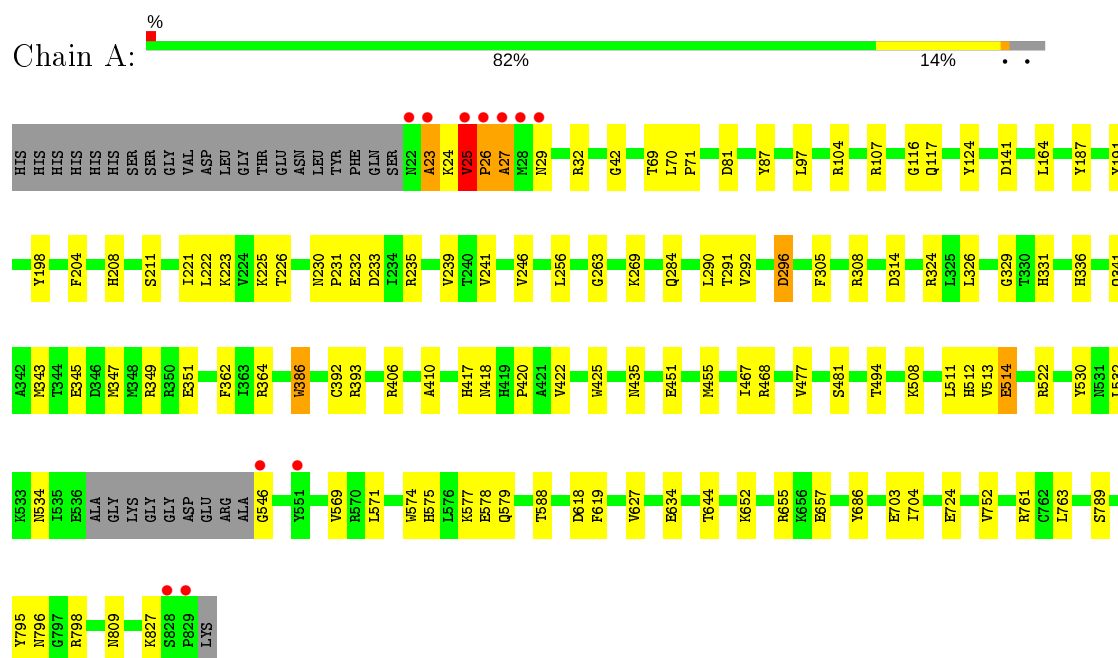
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	D	366	Total 366	O 366	0	0



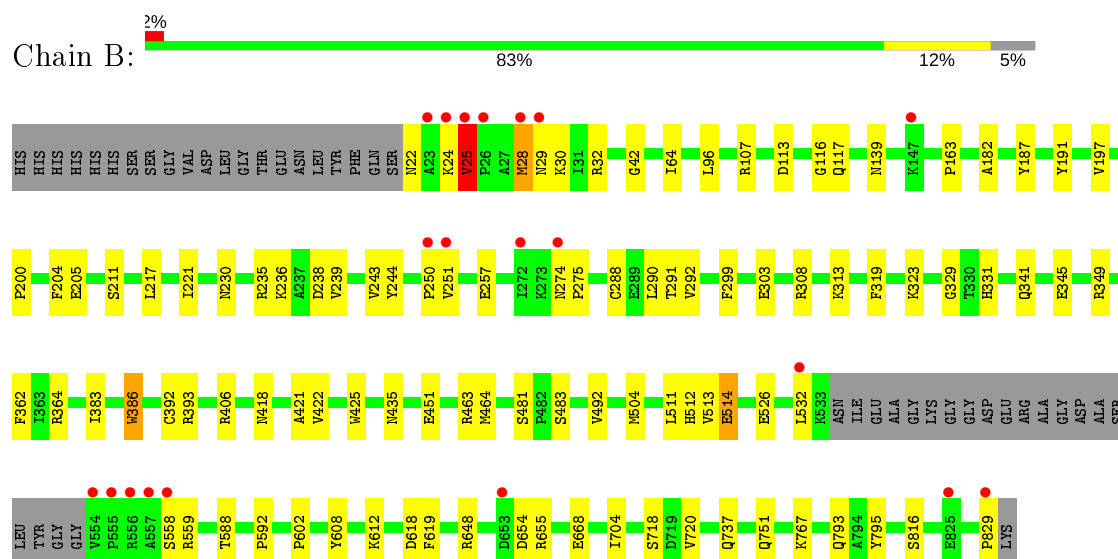
### 3 Residue-property plots

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: Beta-Glucuronidase

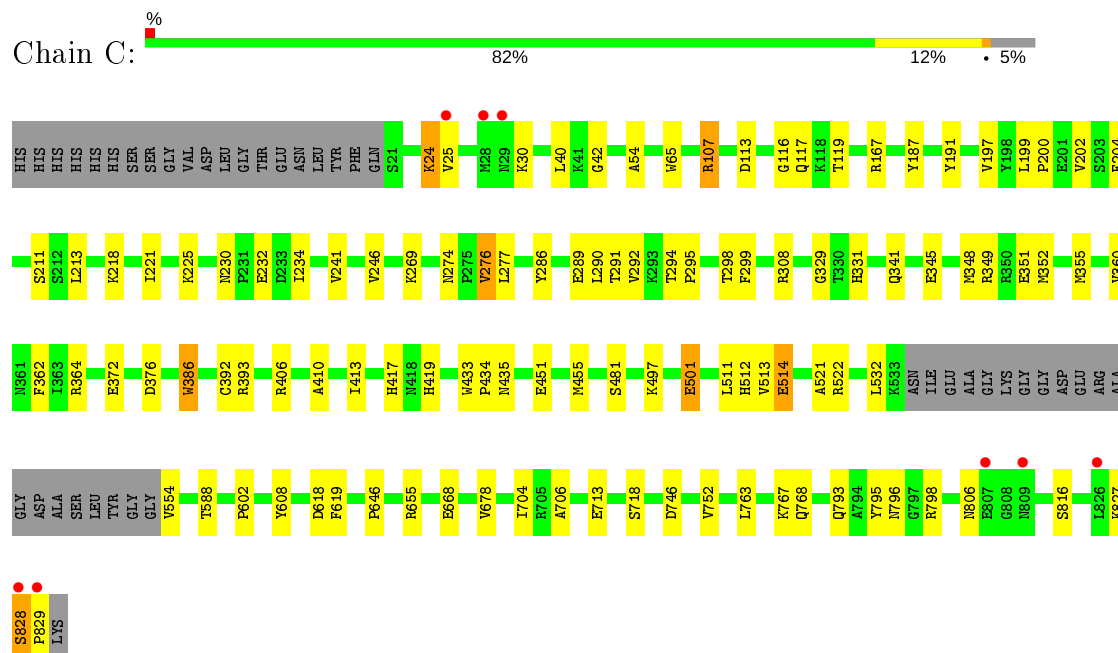


#### • Molecule 1: Beta-Glucuronidase

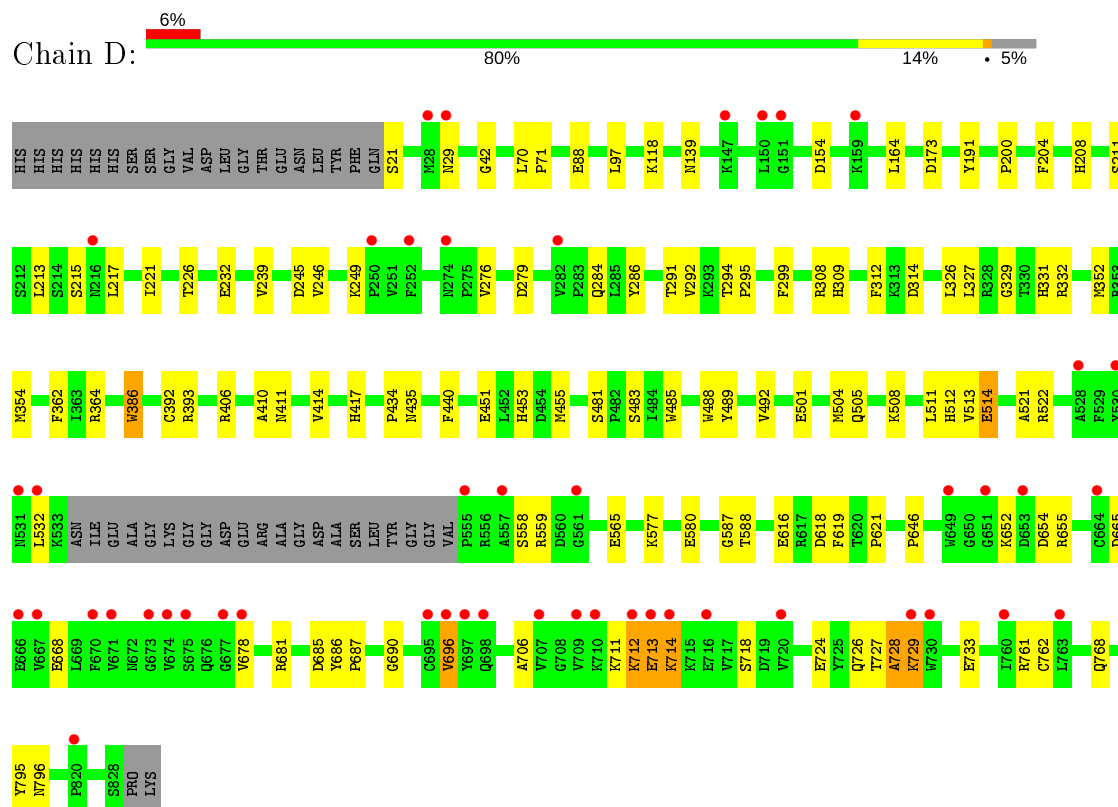




● Molecule 1: Beta-Glucuronidase



● Molecule 1: Beta-Glucuronidase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.92Å 171.97Å 125.32Å 90.00° 107.84° 90.00°	Depositor
Resolution (Å)	48.34 – 2.24 48.34 – 2.24	Depositor EDS
% Data completeness (in resolution range)	95.0 (48.34-2.24) 88.7 (48.34-2.24)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.13 _2998	Depositor
R, $R_{free}$	0.165 , 0.200 0.165 , 0.200	Depositor DCC
$R_{free}$ test set	1953 reflections (1.16%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.5	Xtriage
Anisotropy	0.289	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 55.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	27845	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.35% 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, 0.333 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: GOL, K, 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	0.41	0/6548	0.59	0/8888
1	B	0.41	0/6448	0.58	0/8756
1	C	0.38	0/6448	0.59	1/8757 (0.0%)
1	D	0.36	0/6397	0.56	0/8695
All	All	0.39	0/25841	0.58	1/35096 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	107	ARG	NE-CZ-NH2	-9.33	115.63	120.30

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	6385	0	6136	94	0
1	B	6287	0	6026	78	0
1	C	6288	0	6029	90	1
1	D	6237	0	5939	90	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	72	96	94	13	0
3	B	96	128	127	22	0
3	C	102	136	134	25	0
3	D	48	64	64	2	0
4	A	8	9	7	5	0
4	D	8	9	7	2	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	C	1	0	0	0	0
5	D	1	0	0	0	0
6	A	560	0	0	14	2
6	B	493	0	0	14	0
6	C	445	0	0	12	1
6	D	366	0	0	23	0
All	All	27403	442	24563	355	2

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 7.

The worst 5 of 355 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:C:107:ARG:HH22	3:C:902:GOL:H32	1.13	1.08
1:B:235:ARG:NH2	6:B:1001:HOH:O	1.83	1.08
1:D:762:CYS:SG	6:D:1033:HOH:O	2.12	1.06
1:A:305:PHE:HB2	4:A:914:PRO:HD3	1.42	1.01
1:D:655:ARG:HA	6:D:1012:HOH:O	1.60	1.00

All (2) 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)
6:A:1182:HOH:O	6:C:1363:HOH:O[1_656]	1.96	0.24
1:C:269:LYS:NZ	6:A:1007:HOH:O[1_454]	2.09	0.11



## 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	796/830 (96%)	762 (96%)	28 (4%)	6 (1%)	19	16
1	B	784/830 (94%)	754 (96%)	27 (3%)	3 (0%)	34	35
1	C	785/830 (95%)	759 (97%)	23 (3%)	3 (0%)	34	35
1	D	783/830 (94%)	752 (96%)	25 (3%)	6 (1%)	19	16
All	All	3148/3320 (95%)	3027 (96%)	103 (3%)	18 (1%)	25	23

5 of 18 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	PRO
1	A	27	ALA
1	A	435	ASN
1	B	435	ASN
1	D	713	GLU

### 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	679/719 (94%)	673 (99%)	6 (1%)	78	84
1	B	669/719 (93%)	661 (99%)	8 (1%)	71	78
1	C	669/719 (93%)	662 (99%)	7 (1%)	76	82
1	D	658/719 (92%)	650 (99%)	8 (1%)	71	78
All	All	2675/2876 (93%)	2646 (99%)	29 (1%)	73	80



5 of 29 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	795	TYR
1	C	276	VAL
1	D	696	VAL
1	C	24	LYS
1	C	308	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	208	HIS

### 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 63 ligands modelled in this entry, 8 are monoatomic - leaving 55 for Mogul analysis.

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	$\# Z  > 2$
3	GOL	A	904	-	5,5,5	0.87	0	5,5,5	1.16	0
3	GOL	B	916	-	5,5,5	1.34	1 (20%)	5,5,5	1.36	1 (20%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	C	916	-	5,5,5	0.82	0	5,5,5	0.97	0
3	GOL	A	911	-	5,5,5	0.83	0	5,5,5	0.94	0
3	GOL	A	908	-	5,5,5	1.16	0	5,5,5	1.36	1 (20%)
3	GOL	C	908	-	5,5,5	1.05	0	5,5,5	1.20	1 (20%)
3	GOL	A	902	-	5,5,5	0.69	0	5,5,5	1.23	0
3	GOL	C	913	-	5,5,5	0.81	0	5,5,5	1.02	0
3	GOL	A	913	-	5,5,5	0.97	0	5,5,5	1.57	1 (20%)
3	GOL	C	917	-	5,5,5	0.81	0	5,5,5	1.06	1 (20%)
3	GOL	B	911	-	5,5,5	0.62	0	5,5,5	0.61	0
3	GOL	C	904	-	5,5,5	1.09	0	5,5,5	1.39	1 (20%)
3	GOL	C	903	-	5,5,5	0.72	0	5,5,5	1.07	1 (20%)
3	GOL	B	904	-	5,5,5	0.88	0	5,5,5	1.02	0
3	GOL	D	909	-	5,5,5	0.99	0	5,5,5	1.39	1 (20%)
3	GOL	D	908	-	5,5,5	1.01	0	5,5,5	0.90	0
3	GOL	D	903	-	5,5,5	0.85	0	5,5,5	1.10	0
3	GOL	D	904	-	5,5,5	0.64	0	5,5,5	1.01	0
3	GOL	C	911	-	5,5,5	0.67	0	5,5,5	0.97	0
3	GOL	C	918	-	5,5,5	0.74	0	5,5,5	1.02	0
3	GOL	C	915	-	5,5,5	1.01	0	5,5,5	1.03	0
3	GOL	B	908	-	5,5,5	0.79	0	5,5,5	1.05	0
3	GOL	C	906	-	5,5,5	0.67	0	5,5,5	0.90	0
3	GOL	B	906	-	5,5,5	0.79	0	5,5,5	0.95	0
3	GOL	C	902	-	5,5,5	0.93	0	5,5,5	1.15	0
3	GOL	C	909	-	5,5,5	1.31	1 (20%)	5,5,5	1.56	1 (20%)
3	GOL	A	909	-	5,5,5	1.41	1 (20%)	5,5,5	1.62	1 (20%)
4	PRO	D	910	-	5,8,8	0.36	0	6,10,10	1.07	0
4	PRO	A	914	-	5,8,8	0.61	0	6,10,10	0.96	0
3	GOL	C	905	-	5,5,5	0.87	0	5,5,5	1.22	1 (20%)
3	GOL	D	907	-	5,5,5	0.69	0	5,5,5	1.22	1 (20%)
3	GOL	B	902	-	5,5,5	0.78	0	5,5,5	0.76	0
3	GOL	B	917	-	5,5,5	0.98	0	5,5,5	0.63	0
3	GOL	A	910	-	5,5,5	0.77	0	5,5,5	0.87	0
3	GOL	B	905	-	5,5,5	1.05	1 (20%)	5,5,5	0.82	0
3	GOL	A	907	-	5,5,5	0.87	0	5,5,5	0.79	0
3	GOL	C	907	-	5,5,5	0.88	0	5,5,5	1.19	1 (20%)
3	GOL	D	905	-	5,5,5	1.15	0	5,5,5	1.44	1 (20%)
3	GOL	A	903	-	5,5,5	1.15	0	5,5,5	0.76	0
3	GOL	B	912	-	5,5,5	0.77	0	5,5,5	0.93	0
3	GOL	B	915	-	5,5,5	0.90	0	5,5,5	1.05	0
3	GOL	C	912	-	5,5,5	0.85	0	5,5,5	0.99	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	A	912	-	5,5,5	1.20	1 (20%)	5,5,5	1.31	1 (20%)
3	GOL	D	902	-	5,5,5	1.05	0	5,5,5	1.27	1 (20%)
3	GOL	B	907	-	5,5,5	1.11	1 (20%)	5,5,5	1.34	1 (20%)
3	GOL	A	905	-	5,5,5	0.89	0	5,5,5	1.02	0
3	GOL	B	903	-	5,5,5	0.84	0	5,5,5	0.91	0
3	GOL	B	909	-	5,5,5	0.37	0	5,5,5	0.92	0
3	GOL	A	906	-	5,5,5	0.72	0	5,5,5	1.02	0
3	GOL	B	910	-	5,5,5	0.89	0	5,5,5	1.24	1 (20%)
3	GOL	C	910	-	5,5,5	0.81	0	5,5,5	0.99	0
3	GOL	B	914	-	5,5,5	0.85	0	5,5,5	0.71	0
3	GOL	D	906	-	5,5,5	0.87	0	5,5,5	1.15	1 (20%)
3	GOL	C	914	-	5,5,5	1.19	1 (20%)	5,5,5	1.38	1 (20%)
3	GOL	B	913	-	5,5,5	0.77	0	5,5,5	0.88	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
3	GOL	A	904	-	-	2/4/4/4	-
3	GOL	B	916	-	-	1/4/4/4	-
3	GOL	C	916	-	-	2/4/4/4	-
3	GOL	A	911	-	-	2/4/4/4	-
3	GOL	A	908	-	-	0/4/4/4	-
3	GOL	C	908	-	-	4/4/4/4	-
3	GOL	A	902	-	-	2/4/4/4	-
3	GOL	C	913	-	-	0/4/4/4	-
3	GOL	A	913	-	-	2/4/4/4	-
3	GOL	C	917	-	-	4/4/4/4	-
3	GOL	B	911	-	-	2/4/4/4	-
3	GOL	C	904	-	-	2/4/4/4	-
3	GOL	C	903	-	-	0/4/4/4	-
3	GOL	B	904	-	-	2/4/4/4	-
3	GOL	D	909	-	-	0/4/4/4	-
3	GOL	D	908	-	-	2/4/4/4	-
3	GOL	D	903	-	-	0/4/4/4	-
3	GOL	D	904	-	-	2/4/4/4	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	C	911	-	-	0/4/4/4	-
3	GOL	C	918	-	-	2/4/4/4	-
3	GOL	C	915	-	-	4/4/4/4	-
3	GOL	B	908	-	-	2/4/4/4	-
3	GOL	C	906	-	-	4/4/4/4	-
3	GOL	B	906	-	-	0/4/4/4	-
3	GOL	C	902	-	-	2/4/4/4	-
3	GOL	C	909	-	-	4/4/4/4	-
3	GOL	A	909	-	-	2/4/4/4	-
4	PRO	D	910	-	-	0/0/11/11	0/1/1/1
4	PRO	A	914	-	-	0/0/11/11	0/1/1/1
3	GOL	C	905	-	-	2/4/4/4	-
3	GOL	D	907	-	-	2/4/4/4	-
3	GOL	B	902	-	-	0/4/4/4	-
3	GOL	B	917	-	-	4/4/4/4	-
3	GOL	A	910	-	-	2/4/4/4	-
3	GOL	B	905	-	-	0/4/4/4	-
3	GOL	A	907	-	-	1/4/4/4	-
3	GOL	C	907	-	-	4/4/4/4	-
3	GOL	D	905	-	-	0/4/4/4	-
3	GOL	A	903	-	-	3/4/4/4	-
3	GOL	B	912	-	-	2/4/4/4	-
3	GOL	B	915	-	-	2/4/4/4	-
3	GOL	C	912	-	-	2/4/4/4	-
3	GOL	A	912	-	-	2/4/4/4	-
3	GOL	D	902	-	-	2/4/4/4	-
3	GOL	B	907	-	-	1/4/4/4	-
3	GOL	A	905	-	-	3/4/4/4	-
3	GOL	B	903	-	-	0/4/4/4	-
3	GOL	B	909	-	-	2/4/4/4	-
3	GOL	A	906	-	-	4/4/4/4	-
3	GOL	B	910	-	-	2/4/4/4	-
3	GOL	C	910	-	-	4/4/4/4	-
3	GOL	B	914	-	-	2/4/4/4	-
3	GOL	D	906	-	-	2/4/4/4	-
3	GOL	C	914	-	-	2/4/4/4	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GOL	B	913	-	-	0/4/4/4	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	909	GOL	O2-C2	-2.56	1.35	1.43
3	C	909	GOL	O2-C2	-2.45	1.36	1.43
3	B	907	GOL	O2-C2	-2.22	1.36	1.43
3	A	912	GOL	O2-C2	-2.09	1.37	1.43
3	B	916	GOL	C3-C2	2.04	1.60	1.51

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	909	GOL	C3-C2-C1	-3.28	98.96	111.70
3	A	913	GOL	C3-C2-C1	-3.13	99.55	111.70
3	D	905	GOL	C3-C2-C1	-2.91	100.38	111.70
3	C	909	GOL	C3-C2-C1	-2.79	100.88	111.70
3	A	912	GOL	C3-C2-C1	-2.69	101.26	111.70

There are no chirality outliers.

5 of 99 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	904	GOL	C1-C2-C3-O3
3	A	911	GOL	C1-C2-C3-O3
3	C	908	GOL	C1-C2-C3-O3
3	A	902	GOL	O1-C1-C2-C3
3	C	906	GOL	C1-C2-C3-O3

There are no ring outliers.

34 monomers are involved in 69 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	904	GOL	3	0
3	B	916	GOL	2	0
3	C	916	GOL	2	0
3	A	908	GOL	1	0
3	A	902	GOL	1	0
3	C	913	GOL	4	0
3	A	913	GOL	1	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	904	GOL	4	0
3	C	903	GOL	1	0
3	B	904	GOL	2	0
3	D	908	GOL	1	0
3	D	904	GOL	1	0
3	C	911	GOL	2	0
3	C	918	GOL	4	0
3	B	908	GOL	2	0
3	C	906	GOL	1	0
3	B	906	GOL	2	0
3	C	902	GOL	4	0
4	D	910	PRO	2	0
4	A	914	PRO	5	0
3	B	902	GOL	1	0
3	B	917	GOL	6	0
3	B	905	GOL	3	0
3	C	907	GOL	1	0
3	A	903	GOL	3	0
3	B	912	GOL	1	0
3	B	915	GOL	1	0
3	C	912	GOL	1	0
3	A	912	GOL	2	0
3	A	905	GOL	1	0
3	B	909	GOL	1	0
3	A	906	GOL	1	0
3	B	910	GOL	1	0
3	C	914	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, 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	799/830 (96%)	-0.17	11 (1%) 75 76	16, 25, 40, 68	0
1	B	788/830 (94%)	0.01	20 (2%) 57 58	16, 25, 43, 70	0
1	C	789/830 (95%)	0.00	8 (1%) 82 83	18, 27, 44, 79	0
1	D	787/830 (94%)	0.27	48 (6%) 21 20	19, 35, 56, 83	0
All	All	3163/3320 (95%)	0.03	87 (2%) 53 53	16, 28, 49, 83	0

The worst 5 of 87 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	829	PRO	7.9
1	A	28	MET	6.9
1	A	23	ALA	5.5
1	B	28	MET	5.0
1	D	532	LEU	5.0

### 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 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, 95<sup>th</sup> 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	GOL	C	918	6/6	0.55	0.33	44,52,57,59	0
3	GOL	C	910	6/6	0.62	0.29	36,50,62,65	0
3	GOL	D	906	6/6	0.62	0.19	54,65,76,76	0
3	GOL	C	915	6/6	0.69	0.36	46,59,71,72	0
3	GOL	B	917	6/6	0.74	0.34	41,53,63,63	0
3	GOL	C	902	6/6	0.76	0.28	45,55,67,67	0
3	GOL	D	907	6/6	0.77	0.28	51,61,72,79	0
3	GOL	C	916	6/6	0.78	0.28	43,58,70,70	0
3	GOL	A	913	6/6	0.80	0.27	29,47,52,58	0
3	GOL	D	908	6/6	0.81	0.42	44,57,68,69	0
3	GOL	B	916	6/6	0.81	0.32	39,49,60,60	0
3	GOL	A	910	6/6	0.81	0.36	45,56,71,75	0
3	GOL	D	905	6/6	0.81	0.23	37,45,51,52	0
3	GOL	A	909	6/6	0.81	0.28	29,42,58,69	0
3	GOL	B	914	6/6	0.81	0.29	34,51,73,73	0
3	GOL	C	905	6/6	0.81	0.20	46,55,60,63	0
3	GOL	C	913	6/6	0.82	0.28	43,52,63,63	0
3	GOL	B	912	6/6	0.82	0.25	37,45,57,58	0
3	GOL	A	905	6/6	0.82	0.42	38,49,59,62	0
3	GOL	B	906	6/6	0.83	0.26	33,45,61,61	0
3	GOL	B	911	6/6	0.84	0.23	24,41,48,49	0
3	GOL	D	909	6/6	0.84	0.26	35,44,60,60	0
3	GOL	B	913	6/6	0.84	0.33	39,49,60,72	0
4	PRO	D	910	8/8	0.85	0.23	39,52,62,62	0
3	GOL	A	911	6/6	0.86	0.20	35,53,71,86	0
3	GOL	A	903	6/6	0.86	0.29	34,47,57,61	0
4	PRO	A	914	8/8	0.86	0.20	22,37,42,46	0
3	GOL	B	905	6/6	0.86	0.28	31,38,47,56	0
3	GOL	B	902	6/6	0.87	0.32	34,54,65,69	0
3	GOL	D	903	6/6	0.87	0.30	41,50,60,64	0
3	GOL	A	904	6/6	0.88	0.31	31,39,57,60	0
3	GOL	B	904	6/6	0.88	0.46	39,60,76,76	0
3	GOL	D	904	6/6	0.88	0.19	39,47,55,59	0
3	GOL	C	908	6/6	0.88	0.26	35,44,61,61	0
3	GOL	A	912	6/6	0.89	0.33	40,52,67,67	0
3	GOL	A	908	6/6	0.89	0.18	44,53,63,70	0
3	GOL	C	907	6/6	0.90	0.34	32,40,51,61	0
3	GOL	C	917	6/6	0.90	0.33	30,50,60,68	0
3	GOL	A	906	6/6	0.91	0.34	27,45,56,62	0
3	GOL	B	908	6/6	0.91	0.28	30,39,62,70	0
3	GOL	C	906	6/6	0.91	0.25	46,56,61,67	0
3	GOL	C	911	6/6	0.91	0.22	40,52,63,70	0
3	GOL	B	903	6/6	0.91	0.34	40,54,65,76	0

Continued on next page...



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	GOL	B	915	6/6	0.92	0.18	32,39,45,48	0
3	GOL	B	910	6/6	0.92	0.22	33,42,54,54	0
3	GOL	B	909	6/6	0.93	0.12	21,27,32,33	0
3	GOL	C	909	6/6	0.93	0.16	21,28,36,43	0
3	GOL	C	912	6/6	0.93	0.14	28,34,41,41	0
3	GOL	C	903	6/6	0.93	0.22	36,45,60,62	0
3	GOL	D	902	6/6	0.94	0.12	26,36,42,48	0
3	GOL	C	914	6/6	0.94	0.14	27,32,36,40	0
3	GOL	B	907	6/6	0.94	0.17	23,29,41,41	0
3	GOL	C	904	6/6	0.96	0.12	28,34,40,48	0
3	GOL	A	902	6/6	0.96	0.12	25,30,35,36	0
3	GOL	A	907	6/6	0.96	0.14	18,28,37,41	0
5	NA	C	919	1/1	0.97	0.10	14,14,14,14	0
5	NA	D	911	1/1	0.97	0.09	18,18,18,18	0
5	NA	A	915	1/1	0.98	0.10	16,16,16,16	0
2	K	A	901	1/1	0.99	0.07	20,20,20,20	0
5	NA	B	918	1/1	0.99	0.11	16,16,16,16	0
2	K	D	901	1/1	0.99	0.11	24,24,24,24	0
2	K	C	901	1/1	0.99	0.13	26,26,26,26	0
2	K	B	901	1/1	1.00	0.08	17,17,17,17	0

## 6.5 Other polymers

There are no such residues in this entry.