



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 01:39 pm BST

PDB ID : 6DXU  
Title : Crystal Structure of Parabacteroides merdae Beta-Glucuronidase (GUS)  
Authors : Little, M.S.; Redinbo, M.R.  
Deposited on : 2018-06-30  
Resolution : 1.90 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

<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 : 2.11  
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) : 2.11

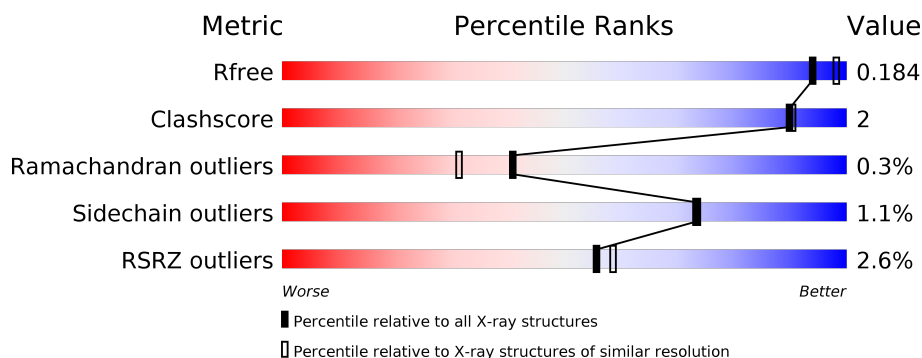
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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>%</div> <div> <div></div> <div>89%</div> <div>7%</div> </div> </div>
1	B	830	<div> <div>3%</div> <div> <div></div> <div>92%</div> <div>5%</div> </div> </div>
1	C	830	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>6%</div> </div> </div>
1	D	830	<div> <div>3%</div> <div> <div></div> <div>91%</div> <div>5%</div> </div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 28061 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 hydrolase family 2, TIM barrel domain protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	800	Total	C	N	O	S	0	0	0
			6359	4044	1098	1187	30			
1	A	774	Total	C	N	O	S	0	0	0
			6208	3952	1065	1161	30			
1	B	803	Total	C	N	O	S	0	1	0
			6353	4040	1091	1192	30			
1	C	806	Total	C	N	O	S	0	0	0
			6383	4053	1106	1194	30			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	1	HIS	-	expression tag	UNP A7AG62
D	2	HIS	-	expression tag	UNP A7AG62
D	3	HIS	-	expression tag	UNP A7AG62
D	4	HIS	-	expression tag	UNP A7AG62
D	5	HIS	-	expression tag	UNP A7AG62
D	6	HIS	-	expression tag	UNP A7AG62
D	7	SER	-	expression tag	UNP A7AG62
D	8	SER	-	expression tag	UNP A7AG62
D	9	GLY	-	expression tag	UNP A7AG62
D	10	VAL	-	expression tag	UNP A7AG62
D	11	ASP	-	expression tag	UNP A7AG62
D	12	LEU	-	expression tag	UNP A7AG62
D	13	GLY	-	expression tag	UNP A7AG62
D	14	THR	-	expression tag	UNP A7AG62
D	15	GLU	-	expression tag	UNP A7AG62
D	16	ASN	-	expression tag	UNP A7AG62
D	17	LEU	-	expression tag	UNP A7AG62
D	18	TYR	-	expression tag	UNP A7AG62
D	19	PHE	-	expression tag	UNP A7AG62
D	20	GLN	-	expression tag	UNP A7AG62
D	21	SER	-	expression tag	UNP A7AG62

*Continued on next page...*

*Continued from previous page...*

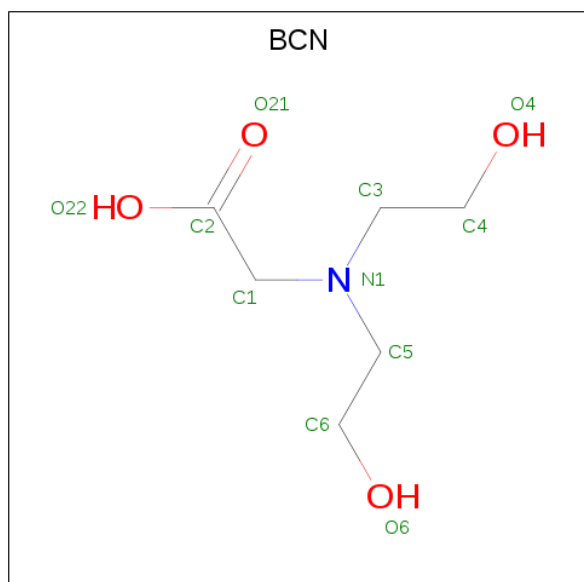
Chain	Residue	Modelled	Actual	Comment	Reference
D	22	ASN	-	expression tag	UNP A7AG62
A	1	HIS	-	expression tag	UNP A7AG62
A	2	HIS	-	expression tag	UNP A7AG62
A	3	HIS	-	expression tag	UNP A7AG62
A	4	HIS	-	expression tag	UNP A7AG62
A	5	HIS	-	expression tag	UNP A7AG62
A	6	HIS	-	expression tag	UNP A7AG62
A	7	SER	-	expression tag	UNP A7AG62
A	8	SER	-	expression tag	UNP A7AG62
A	9	GLY	-	expression tag	UNP A7AG62
A	10	VAL	-	expression tag	UNP A7AG62
A	11	ASP	-	expression tag	UNP A7AG62
A	12	LEU	-	expression tag	UNP A7AG62
A	13	GLY	-	expression tag	UNP A7AG62
A	14	THR	-	expression tag	UNP A7AG62
A	15	GLU	-	expression tag	UNP A7AG62
A	16	ASN	-	expression tag	UNP A7AG62
A	17	LEU	-	expression tag	UNP A7AG62
A	18	TYR	-	expression tag	UNP A7AG62
A	19	PHE	-	expression tag	UNP A7AG62
A	20	GLN	-	expression tag	UNP A7AG62
A	21	SER	-	expression tag	UNP A7AG62
A	22	ASN	-	expression tag	UNP A7AG62
B	1	HIS	-	expression tag	UNP A7AG62
B	2	HIS	-	expression tag	UNP A7AG62
B	3	HIS	-	expression tag	UNP A7AG62
B	4	HIS	-	expression tag	UNP A7AG62
B	5	HIS	-	expression tag	UNP A7AG62
B	6	HIS	-	expression tag	UNP A7AG62
B	7	SER	-	expression tag	UNP A7AG62
B	8	SER	-	expression tag	UNP A7AG62
B	9	GLY	-	expression tag	UNP A7AG62
B	10	VAL	-	expression tag	UNP A7AG62
B	11	ASP	-	expression tag	UNP A7AG62
B	12	LEU	-	expression tag	UNP A7AG62
B	13	GLY	-	expression tag	UNP A7AG62
B	14	THR	-	expression tag	UNP A7AG62
B	15	GLU	-	expression tag	UNP A7AG62
B	16	ASN	-	expression tag	UNP A7AG62
B	17	LEU	-	expression tag	UNP A7AG62
B	18	TYR	-	expression tag	UNP A7AG62
B	19	PHE	-	expression tag	UNP A7AG62

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	20	GLN	-	expression tag	UNP A7AG62
B	21	SER	-	expression tag	UNP A7AG62
B	22	ASN	-	expression tag	UNP A7AG62
C	1	HIS	-	expression tag	UNP A7AG62
C	2	HIS	-	expression tag	UNP A7AG62
C	3	HIS	-	expression tag	UNP A7AG62
C	4	HIS	-	expression tag	UNP A7AG62
C	5	HIS	-	expression tag	UNP A7AG62
C	6	HIS	-	expression tag	UNP A7AG62
C	7	SER	-	expression tag	UNP A7AG62
C	8	SER	-	expression tag	UNP A7AG62
C	9	GLY	-	expression tag	UNP A7AG62
C	10	VAL	-	expression tag	UNP A7AG62
C	11	ASP	-	expression tag	UNP A7AG62
C	12	LEU	-	expression tag	UNP A7AG62
C	13	GLY	-	expression tag	UNP A7AG62
C	14	THR	-	expression tag	UNP A7AG62
C	15	GLU	-	expression tag	UNP A7AG62
C	16	ASN	-	expression tag	UNP A7AG62
C	17	LEU	-	expression tag	UNP A7AG62
C	18	TYR	-	expression tag	UNP A7AG62
C	19	PHE	-	expression tag	UNP A7AG62
C	20	GLN	-	expression tag	UNP A7AG62
C	21	SER	-	expression tag	UNP A7AG62
C	22	ASN	-	expression tag	UNP A7AG62

- Molecule 2 is BICINE (three-letter code: BCN) (formula:  $C_6H_{13}NO_4$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	D	1	Total	C	N	O	0	0
			11	6	1	4		
2	C	1	Total	C	N	O	0	0
			11	6	1	4		
2	C	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Na	0	0
			2	2		
3	A	2	Total	Na	0	0
			2	2		
3	D	2	Total	Na	0	0
			2	2		
3	C	2	Total	Na	0	0
			2	2		

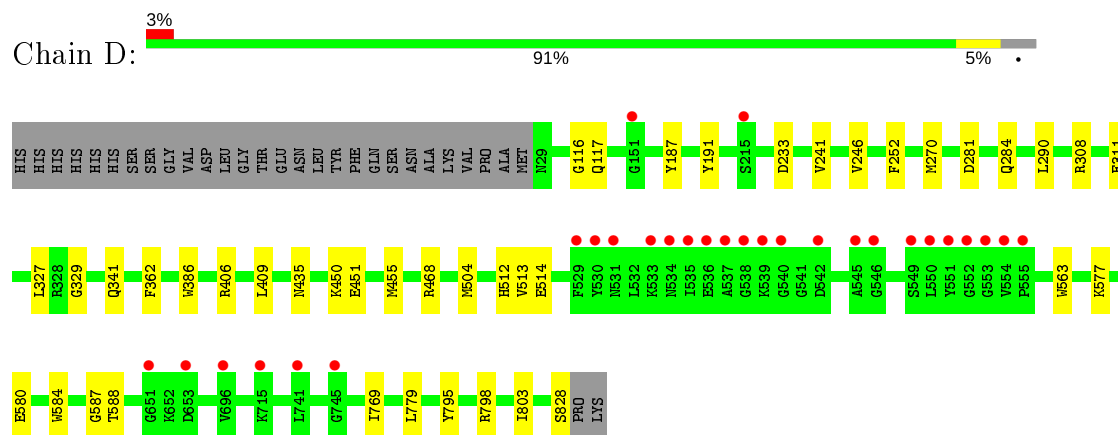
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	D	650	Total	O	0	0
			650	650		
4	A	721	Total	O	0	0
			721	721		
4	B	630	Total	O	0	0
			630	630		
4	C	716	Total	O	0	0
			716	716		

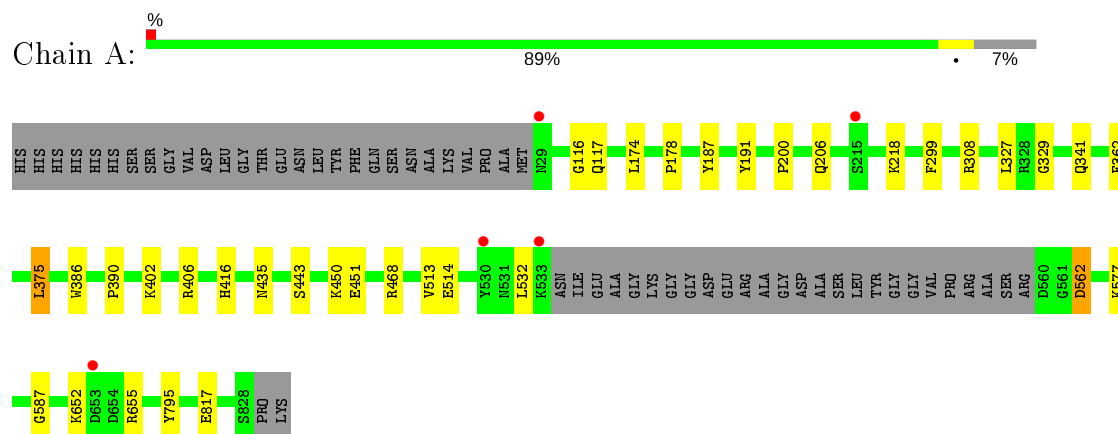
### 3 Residue-property plots [i](#)

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.

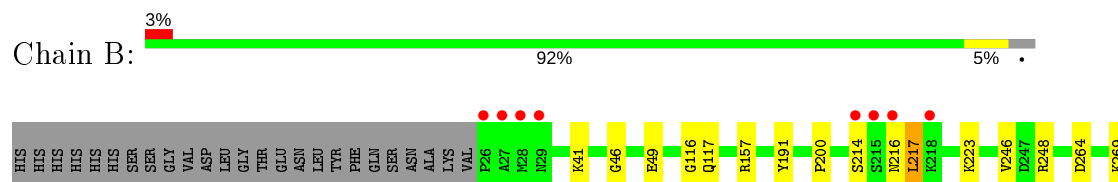
- Molecule 1: Glycosyl hydrolase family 2, TIM barrel domain protein



- Molecule 1: Glycosyl hydrolase family 2, TIM barrel domain protein

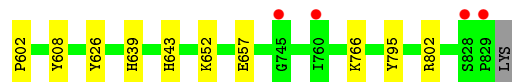
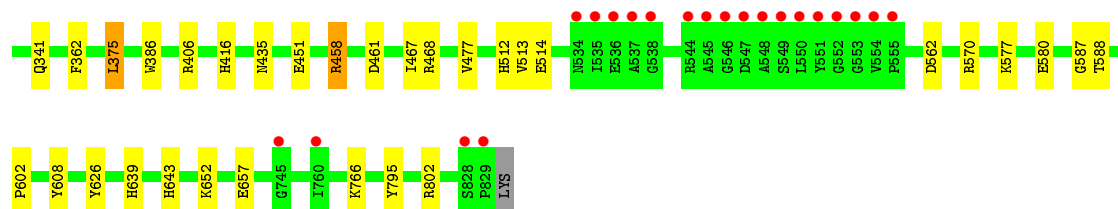
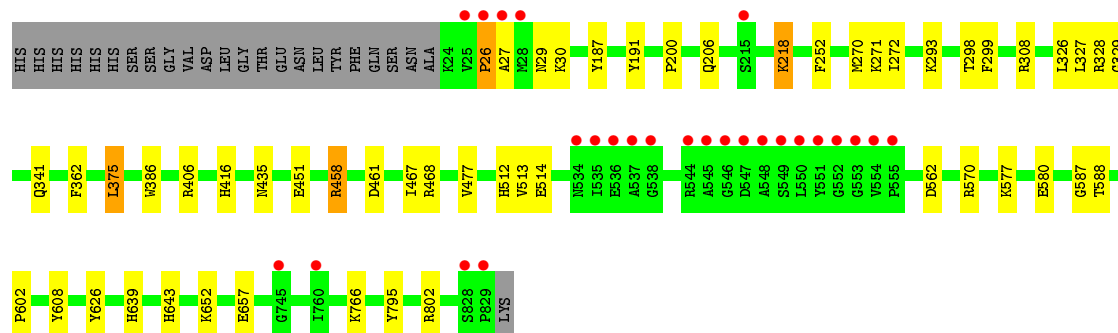
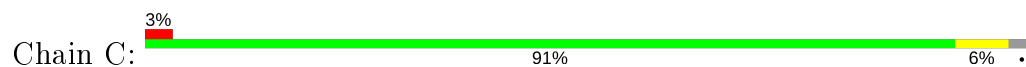


- Molecule 1: Glycosyl hydrolase family 2, TIM barrel domain protein





- Molecule 1: Glycosyl hydrolase family 2, TIM barrel domain protein





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.37Å 163.39Å 120.61Å 90.00° 103.35° 90.00°	Depositor
Resolution (Å)	29.63 – 1.90 29.63 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.9 (29.63-1.90) 96.2 (29.63-1.90)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.90 (at 1.91Å)	Xtriage
Refinement program	PHENIX 1.13_2998, PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.153 , 0.184 0.153 , 0.184	Depositor DCC
$R_{free}$ test set	2000 reflections (0.74%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	23.1	Xtriage
Anisotropy	0.226	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 52.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	28061	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.36% 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 ⓘ

### 5.1 Standard geometry ⓘ

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

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	0.40	0/6366	0.59	1/8634 (0.0%)
1	B	0.37	0/6515	0.57	1/8849 (0.0%)
1	C	0.39	0/6545	0.60	1/8885 (0.0%)
1	D	0.38	0/6520	0.58	0/8844
All	All	0.39	0/25946	0.58	3/35212 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	375	LEU	CA-CB-CG	-7.95	97.01	115.30
1	A	375	LEU	CA-CB-CG	-7.43	98.20	115.30
1	B	217	LEU	CA-CB-CG	6.19	129.54	115.30

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	6208	0	5988	19	0
1	B	6353	0	6046	19	0
1	C	6383	0	6107	33	0
1	D	6359	0	6111	20	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	C	22	0	24	3	0
2	D	11	0	12	1	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
4	A	721	0	0	7	0
4	B	630	0	0	4	1
4	C	716	0	0	9	1
4	D	650	0	0	5	2
All	All	28061	0	24288	87	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 2.

All (87) 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:562:ASP:O	1:C:766:LYS:NZ	2.08	0.85
1:B:214:SER:O	4:B:1001:HOH:O	1.97	0.82
1:B:570:ARG:NH2	1:B:657:GLU:OE1	2.21	0.73
1:A:562:ASP:OD1	4:A:1002:HOH:O	2.10	0.69
1:B:264:ASP:OD2	1:C:458:ARG:NH1	2.25	0.69
1:A:206:GLN:OE1	4:A:1001:HOH:O	2.10	0.69
1:C:26:PRO:HG2	1:C:30:LYS:HD2	1.75	0.68
1:A:406:ARG:NE	1:A:451:GLU:OE1	2.24	0.68
1:C:570:ARG:NH2	1:C:657:GLU:OE1	2.29	0.66
1:C:375:LEU:HD13	1:C:416:HIS:CE1	2.31	0.65
1:A:375:LEU:HD13	1:A:416:HIS:CE1	2.33	0.64
1:C:461:ASP:OD2	4:C:1001:HOH:O	2.16	0.63
1:B:217:LEU:HD23	4:B:1001:HOH:O	2.00	0.61
4:A:1458:HOH:O	1:B:41:LYS:HE2	2.02	0.60
1:D:580:GLU:OE1	4:D:1001:HOH:O	2.16	0.59
1:D:563:TRP:CG	2:D:901:BCN:H61	2.42	0.55
1:C:652:LYS:O	4:C:1003:HOH:O	2.18	0.54
1:C:406:ARG:NE	1:C:451:GLU:OE1	2.39	0.54
1:D:327:LEU:HB2	1:D:587:GLY:HA3	1.89	0.54
1:D:406:ARG:NE	1:D:451:GLU:OE1	2.39	0.54
1:B:406:ARG:NE	1:B:451:GLU:OE1	2.37	0.54
1:D:281:ASP:OD1	4:D:1002:HOH:O	2.18	0.54
1:C:580:GLU:OE1	4:C:1006:HOH:O	2.19	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:468:ARG:HD2	1:C:513:VAL:HG23	1.90	0.53
1:A:468:ARG:HD2	1:A:513:VAL:HG23	1.91	0.52
1:C:206:GLN:OE1	4:C:1005:HOH:O	2.19	0.52
1:C:29:ASN:N	1:C:29:ASN:OD1	2.41	0.51
1:B:327:LEU:HB2	1:B:587:GLY:HA3	1.93	0.51
1:C:270:MET:HB2	2:C:902:BCN:H11	1.93	0.51
1:C:626:TYR:OH	4:C:1002:HOH:O	2.17	0.50
1:A:327:LEU:HB2	1:A:587:GLY:HA3	1.94	0.50
1:D:233:ASP:HB2	1:A:443:SER:HB2	1.95	0.49
1:B:450:LYS:HD3	1:B:476:ILE:HG23	1.94	0.49
1:D:779:LEU:HD23	1:D:803:ILE:HD12	1.95	0.49
1:D:246:VAL:HG11	1:D:284:GLN:HB3	1.95	0.49
1:C:329:GLY:HA2	1:C:362:PHE:O	2.13	0.48
1:D:468:ARG:HD2	1:D:513:VAL:HG23	1.94	0.48
1:D:241:VAL:HG22	1:D:290:LEU:HD22	1.95	0.48
1:C:802:ARG:NE	4:C:1026:HOH:O	2.47	0.47
1:C:326:LEU:HD23	1:C:328:ARG:HH12	1.80	0.47
1:C:326:LEU:HD23	1:C:328:ARG:NH1	2.30	0.47
1:B:246:VAL:HG23	4:B:1328:HOH:O	2.14	0.47
1:C:252:PHE:HD1	2:C:902:BCN:H52	1.79	0.47
1:B:248:ARG:NH1	1:B:302:GLU:OE1	2.39	0.47
1:A:450:LYS:HE2	4:A:1198:HOH:O	2.15	0.46
1:D:504:MET:HG2	1:D:584:TRP:CZ3	2.50	0.46
1:A:116:GLY:HA2	1:A:117:GLN:HA	1.69	0.46
1:A:329:GLY:HA2	1:A:362:PHE:O	2.16	0.46
1:C:26:PRO:CG	1:C:30:LYS:HD2	2.45	0.46
1:C:580:GLU:OE2	4:C:1007:HOH:O	2.21	0.46
1:C:270:MET:HE2	1:C:272:ILE:HD11	1.98	0.46
1:D:450:LYS:HE2	4:D:1193:HOH:O	2.17	0.45
1:B:468:ARG:HD2	1:B:513:VAL:HG23	1.97	0.45
1:D:798:ARG:HD3	4:D:1014:HOH:O	2.17	0.45
1:B:200:PRO:HG3	1:B:299:PHE:HB2	1.98	0.44
1:B:223:LYS:HG2	1:B:269:LYS:HG3	2.00	0.44
1:A:178:PRO:HB3	1:A:390:PRO:HB2	2.00	0.44
1:C:327:LEU:HB2	1:C:587:GLY:HA3	2.00	0.44
1:A:532:LEU:HB2	1:A:655:ARG:HB2	2.00	0.44
1:C:271:LYS:O	2:C:902:BCN:H12	2.18	0.44
1:D:116:GLY:HA2	1:D:117:GLN:HA	1.62	0.44
1:A:817:GLU:OE2	4:A:1004:HOH:O	2.21	0.43
1:B:303:GLU:HG2	1:B:418:ASN:ND2	2.33	0.43
1:B:116:GLY:HA2	1:B:117:GLN:HA	1.66	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:652:LYS:O	4:A:1003:HOH:O	2.21	0.43
1:C:328:ARG:NH1	4:C:1041:HOH:O	2.51	0.43
1:C:467:ILE:HG23	1:C:477:VAL:HG11	2.00	0.42
1:C:218:LYS:NZ	4:C:1008:HOH:O	2.28	0.42
1:B:798:ARG:HD3	4:B:1003:HOH:O	2.19	0.42
1:D:329:GLY:HA2	1:D:362:PHE:O	2.20	0.42
1:B:602:PRO:HB3	1:B:608:TYR:CZ	2.54	0.42
1:C:293:LYS:HG2	1:C:298:THR:HG23	2.02	0.42
1:C:602:PRO:HB3	1:C:608:TYR:CZ	2.55	0.42
1:D:252:PHE:HB2	1:D:270:MET:HE3	2.00	0.42
1:C:200:PRO:HG3	1:C:299:PHE:HB2	2.01	0.41
1:A:174:LEU:HD12	1:B:46:GLY:HA3	2.03	0.41
1:C:639:HIS:CE1	1:C:643:HIS:CD2	3.08	0.41
1:D:409:LEU:HD23	1:D:455:MET:HE2	2.03	0.41
1:D:512:HIS:O	1:D:588:THR:HA	2.19	0.41
1:A:200:PRO:HG3	1:A:299:PHE:HB2	2.01	0.41
1:D:187:TYR:CD2	1:D:341:GLN:HB2	2.55	0.41
1:C:512:HIS:O	1:C:588:THR:HA	2.20	0.41
1:A:187:TYR:CD2	1:A:341:GLN:HB2	2.55	0.41
1:D:311:GLU:OE2	4:D:1003:HOH:O	2.20	0.41
1:A:402:LYS:NZ	4:A:1019:HOH:O	2.46	0.40
1:C:187:TYR:CD2	1:C:341:GLN:HB2	2.56	0.40
1:A:174:LEU:HD11	1:B:49:GLU:HG3	2.03	0.40

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)
4:D:1532:HOH:O	4:B:1554:HOH:O[2_455]	2.14	0.06
4:D:1590:HOH:O	4:C:1008:HOH:O[2_556]	2.16	0.04

## 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	770/830 (93%)	747 (97%)	21 (3%)	2 (0%)	41	31
1	B	802/830 (97%)	773 (96%)	27 (3%)	2 (0%)	47	38
1	C	804/830 (97%)	778 (97%)	22 (3%)	4 (0%)	29	18
1	D	798/830 (96%)	773 (97%)	23 (3%)	2 (0%)	41	31
All	All	3174/3320 (96%)	3071 (97%)	93 (3%)	10 (0%)	41	31

All (10) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	435	ASN
1	C	26	PRO
1	D	435	ASN
1	B	435	ASN
1	C	27	ALA
1	C	435	ASN
1	D	514	GLU
1	A	514	GLU
1	B	514	GLU
1	C	514	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	663/719 (92%)	656 (99%)	7 (1%)	73	73
1	B	662/719 (92%)	655 (99%)	7 (1%)	73	73
1	C	668/719 (93%)	661 (99%)	7 (1%)	76	76
1	D	669/719 (93%)	662 (99%)	7 (1%)	76	76
All	All	2662/2876 (93%)	2634 (99%)	28 (1%)	73	73

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	191	TYR
1	D	308	ARG
1	D	386	TRP
1	D	577	LYS
1	D	769	ILE
1	D	795	TYR
1	D	828	SER
1	A	191	TYR
1	A	218	LYS
1	A	308	ARG
1	A	386	TRP
1	A	562	ASP
1	A	577	LYS
1	A	795	TYR
1	B	157	ARG
1	B	191	TYR
1	B	216	ASN
1	B	308	ARG
1	B	386	TRP
1	B	577	LYS
1	B	795	TYR
1	C	191	TYR
1	C	218	LYS
1	C	308	ARG
1	C	386	TRP
1	C	458	ARG
1	C	577	LYS
1	C	795	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 11 ligands modelled in this entry, 8 are monoatomic - leaving 3 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	$\# Z  > 2$
2	BCN	C	902	-	7,10,10	0.57	0	8,11,11	1.34	2 (25%)
2	BCN	C	901	-	7,10,10	0.56	0	8,11,11	1.49	1 (12%)
2	BCN	D	901	-	7,10,10	0.51	0	8,11,11	1.38	1 (12%)

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	BCN	C	902	-	-	1/8/10/10	-
2	BCN	C	901	-	-	3/8/10/10	-
2	BCN	D	901	-	-	3/8/10/10	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	C	901	BCN	C2-C1-N1	-3.19	108.93	113.48
2	D	901	BCN	C2-C1-N1	-3.11	109.06	113.48
2	C	902	BCN	C1-N1-C3	-2.09	106.88	111.29
2	C	902	BCN	C1-N1-C5	-2.04	106.99	111.29

There are no chirality outliers.



All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	901	BCN	N1-C5-C6-O6
2	D	901	BCN	N1-C5-C6-O6
2	C	901	BCN	C2-C1-N1-C3
2	C	901	BCN	C2-C1-N1-C5
2	C	902	BCN	C6-C5-N1-C3
2	D	901	BCN	C6-C5-N1-C3
2	D	901	BCN	C6-C5-N1-C1

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	902	BCN	3	0
2	D	901	BCN	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 ⓘ

### 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	774/830 (93%)	-0.35	5 (0%) 89 90	17, 24, 40, 56	0
1	B	803/830 (96%)	-0.12	24 (2%) 50 53	17, 28, 47, 76	0
1	C	806/830 (97%)	-0.12	26 (3%) 47 50	16, 26, 46, 71	0
1	D	800/830 (96%)	-0.13	29 (3%) 42 45	16, 26, 46, 64	0
All	All	3183/3320 (95%)	-0.18	84 (2%) 56 58	16, 26, 45, 76	0

All (84) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	551	TYR	7.5
1	C	551	TYR	7.1
1	C	28	MET	6.9
1	B	28	MET	6.7
1	C	27	ALA	6.7
1	C	550	LEU	6.6
1	B	27	ALA	6.6
1	D	530	TYR	6.5
1	B	551	TYR	6.2
1	C	829	PRO	5.9
1	B	530	TYR	5.8
1	D	550	LEU	5.5
1	D	554	VAL	5.4
1	C	26	PRO	5.2
1	C	554	VAL	5.0
1	C	535	ILE	4.6
1	B	650	GLY	4.5
1	C	545	ALA	4.5
1	B	713	GLU	4.3
1	C	546	GLY	4.1
1	B	712	LYS	4.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	553	GLY	3.9
1	D	534	ASN	3.9
1	B	550	LEU	3.9
1	C	537	ALA	3.8
1	D	537	ALA	3.7
1	D	538	GLY	3.7
1	C	538	GLY	3.7
1	D	653	ASP	3.5
1	B	215	SER	3.5
1	C	549	SER	3.4
1	D	552	GLY	3.3
1	D	555	PRO	3.3
1	D	535	ILE	3.2
1	D	549	SER	3.1
1	D	536	GLU	3.0
1	B	711	LYS	3.0
1	D	696	VAL	3.0
1	B	651	GLY	3.0
1	C	544	ARG	3.0
1	D	545	ALA	2.9
1	B	549	SER	2.9
1	C	215	SER	2.9
1	B	26	PRO	2.9
1	C	552	GLY	2.9
1	B	29	ASN	2.8
1	D	529	PHE	2.8
1	B	540	GLY	2.8
1	D	539	LYS	2.8
1	D	553	GLY	2.8
1	C	745	GLY	2.8
1	B	714	LYS	2.8
1	D	542	ASP	2.7
1	A	530	TYR	2.6
1	D	531	ASN	2.6
1	C	555	PRO	2.6
1	B	539	LYS	2.6
1	C	536	GLU	2.5
1	D	715	LYS	2.5
1	A	29	ASN	2.5
1	C	25	VAL	2.4
1	D	151	GLY	2.4
1	C	534	ASN	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	653	ASP	2.4
1	D	540	GLY	2.4
1	A	533	LYS	2.4
1	D	533	LYS	2.3
1	B	216	ASN	2.3
1	B	728	ALA	2.3
1	D	741	LEU	2.3
1	C	828	SER	2.2
1	D	546	GLY	2.2
1	B	534	ASN	2.2
1	B	760	ILE	2.2
1	C	760	ILE	2.2
1	A	215	SER	2.2
1	B	731	ASP	2.1
1	D	651	GLY	2.1
1	D	745	GLY	2.1
1	C	548	ALA	2.1
1	B	218	LYS	2.1
1	C	547	ASP	2.1
1	D	215	SER	2.0
1	B	214	SER	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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
2	BCN	C	902	11/11	0.84	0.28	45,51,58,65	0
2	BCN	D	901	11/11	0.87	0.20	27,36,43,44	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	BCN	C	901	11/11	0.92	0.13	27,35,42,48	0
3	NA	D	902	1/1	0.99	0.07	22,22,22,22	0
3	NA	C	904	1/1	0.99	0.05	21,21,21,21	0
3	NA	C	903	1/1	0.99	0.09	22,22,22,22	0
3	NA	B	901	1/1	0.99	0.08	24,24,24,24	0
3	NA	A	901	1/1	0.99	0.09	20,20,20,20	0
3	NA	A	902	1/1	1.00	0.03	20,20,20,20	0
3	NA	B	902	1/1	1.00	0.11	21,21,21,21	0
3	NA	D	903	1/1	1.00	0.08	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.