



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 01:18 AM GMT

PDB ID : 2CN3  
Title : CRYSTAL STRUCTURES OF CLOSTRIDIUM THERMOCELLUM XY-  
LOGLUCANASE  
Authors : Martinez-Fleites, C.; Taylor, E.J.; Guerreiro, C.I.P.D.; Prates, J.A.M.; Fer-  
reira, L.M.A.; Fontes, C.M.G.A.; Baumann, M.J.; Brumer, H.; Davies, G.J.  
Deposited on : 2006-05-17  
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

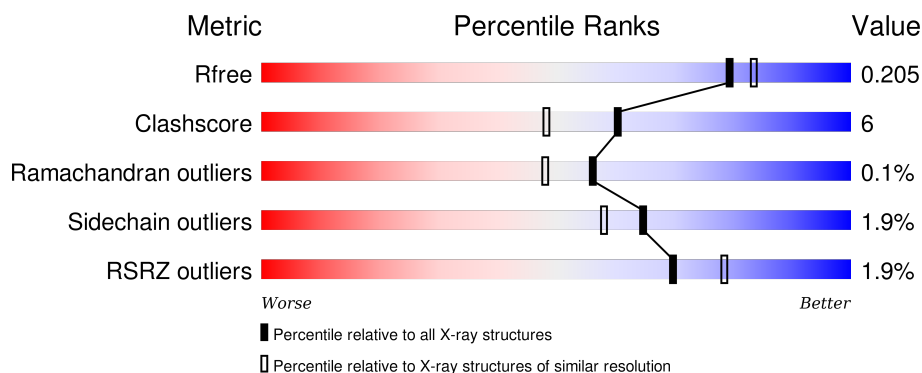
# 1 Overall quality at a glance ⓘ

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	737	<div> <div>2%</div> <div>87%</div> <div>11%</div> <div>..</div> </div>
1	B	737	<div> <div>2%</div> <div>88%</div> <div>10%</div> <div>..</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	GLC	A	1761	X	-	-	-
2	GLC	B	1761	X	-	-	-
2	GLC	B	1769	X	-	-	-
3	GLC	A	1769	X	-	-	-
3	XYS	A	1774	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12865 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-1,4-XYLOGLUCAN HYDROLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	728	Total	C	N	O	S	0	0	0
			5645	3608	938	1075	24			
1	B	728	Total	C	N	O	S	0	0	0
			5647	3609	938	1076	24			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	70	ALA	ASP	ENGINEERED MUTATION	UNP Q70DK5
B	70	ALA	ASP	ENGINEERED MUTATION	UNP Q70DK5

- Molecule 2 is a polymer of unknown type called SUGAR (8-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	8	Total	C	O	0	0
			83	45	38		
2	B	8	Total	C	O	0	0
			83	45	38		
2	B	8	Total	C	O	0	0
			83	45	38		

- Molecule 3 is a polymer of unknown type called SUGAR (9-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	9	Total	C	O	0	0
			94	51	43		

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Ca	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

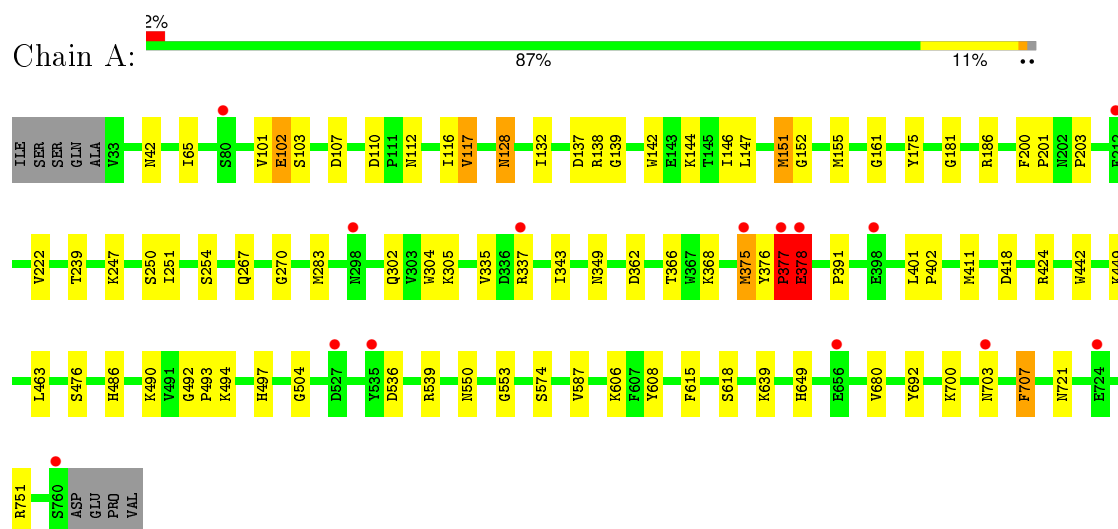
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	611	Total	O	0	0
			611	611		
5	B	617	Total	O	0	0
			617	617		

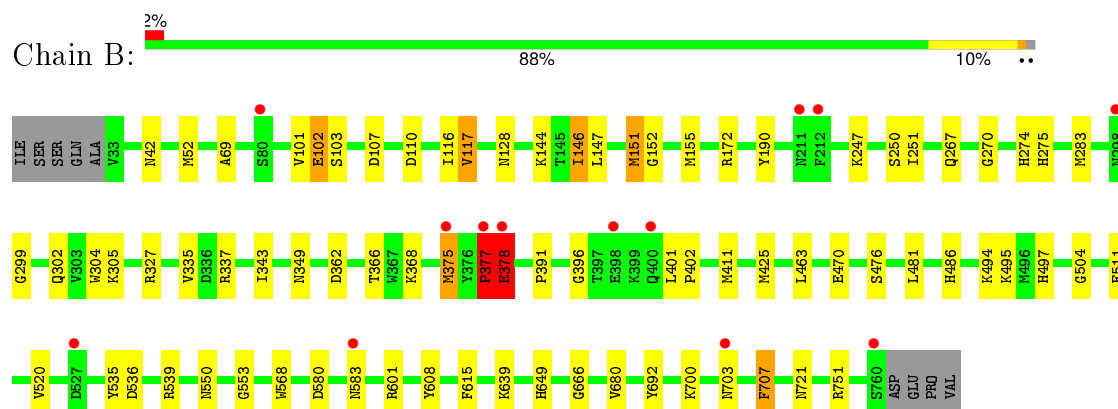
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: BETA-1,4-XYLOGLUCAN HYDROLASE



#### • Molecule 1: BETA-1,4-XYLOGLUCAN HYDROLASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	141.44Å 141.44Å 193.39Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.96 – 1.95 19.96 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.96-1.95) 100.0 (19.96-1.95)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.07 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.178 , 0.204 0.178 , 0.205	Depositor DCC
$R_{free}$ test set	7140 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	16.0	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.39 , 51.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	1 of 142337 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12865	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.92 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 9.2594e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, CA, BGC, XYS, GAL

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	0.53	0/5811	0.61	2/7926 (0.0%)
1	B	0.51	0/5814	0.61	2/7931 (0.0%)
All	All	0.52	0/11625	0.61	4/15857 (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
1	B	0	2
2	A	1	0
2	B	2	0
3	A	1	0
All	All	4	3

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	378	GLU	N-CA-C	-7.31	91.25	111.00
1	A	378	GLU	N-CA-C	-7.02	92.06	111.00
1	B	275	HIS	N-CA-C	6.77	129.27	111.00
1	A	137	ASP	N-CA-C	6.70	129.08	111.00

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	1761	GLC	C1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atom
3	A	1769	GLC	C1
2	B	1761	GLC	C1
2	B	1769	GLC	C1

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	377	PRO	Peptide
1	B	274	HIS	Peptide
1	B	377	PRO	Peptide

## 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	5645	0	5400	67	0
1	B	5647	0	5406	60	0
2	A	83	0	69	0	0
2	B	166	0	138	0	0
3	A	94	0	78	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	611	0	0	5	0
5	B	617	0	0	4	0
All	All	12865	0	11091	127	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 6.

All (127) 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:117:VAL:HG23	1:B:151:MET:HE2	1.27	1.12
1:B:283:MET:HE2	1:B:305:LYS:HB2	1.31	1.07
1:A:103:SER:HB3	1:A:117:VAL:CG1	2.01	0.91
1:B:103:SER:HB3	1:B:117:VAL:CG1	2.04	0.88

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:144:LYS:HE2	1:A:146:ILE:HD11	1.54	0.88
1:B:117:VAL:HG23	1:B:151:MET:CE	2.04	0.87
1:B:700:LYS:HE2	1:B:703:ASN:HA	1.59	0.84
1:A:700:LYS:HE3	1:A:703:ASN:HA	1.60	0.82
1:B:337:ARG:HD3	5:B:2026:HOH:O	1.78	0.82
1:A:375:MET:O	1:A:378:GLU:HB2	1.79	0.81
1:A:362:ASP:OD2	1:A:366:THR:HG22	1.82	0.79
1:B:283:MET:CE	1:B:305:LYS:HB2	2.12	0.77
1:B:362:ASP:OD2	1:B:366:THR:HG22	1.86	0.76
1:A:117:VAL:HG23	1:A:151:MET:CE	2.16	0.76
1:A:302:GLN:HE21	1:A:304:TRP:HE1	1.33	0.75
1:A:283:MET:HE2	1:A:305:LYS:HB2	1.67	0.75
1:B:102:GLU:HB3	1:B:117:VAL:HG22	1.68	0.74
1:B:251:ILE:H	1:B:267:GLN:HE22	1.32	0.74
1:A:550:ASN:HD22	1:A:553:GLY:H	1.34	0.74
1:A:302:GLN:NE2	1:A:304:TRP:HE1	1.86	0.73
1:A:251:ILE:H	1:A:267:GLN:HE22	1.34	0.72
1:B:302:GLN:HE21	1:B:304:TRP:HE1	1.39	0.70
1:B:550:ASN:HD22	1:B:553:GLY:H	1.38	0.69
1:A:490:LYS:HD3	5:A:2577:HOH:O	1.91	0.69
1:B:375:MET:O	1:B:378:GLU:HB2	1.93	0.68
1:B:283:MET:HE2	1:B:305:LYS:CB	2.19	0.67
1:B:302:GLN:NE2	1:B:304:TRP:HE1	1.92	0.67
1:A:102:GLU:HB3	1:A:117:VAL:HG22	1.77	0.67
1:B:42:ASN:HD21	1:B:751:ARG:HH12	1.44	0.66
1:B:103:SER:HB3	1:B:117:VAL:HG11	1.78	0.65
1:B:536:ASP:O	1:B:539:ARG:HG3	1.96	0.65
1:A:117:VAL:HG23	1:A:151:MET:HE3	1.79	0.64
1:B:375:MET:N	1:B:375:MET:SD	2.61	0.64
1:A:375:MET:SD	1:A:375:MET:N	2.62	0.63
1:B:368:LYS:HE3	5:B:2312:HOH:O	1.98	0.62
1:A:103:SER:HB3	1:A:117:VAL:HG11	1.80	0.62
1:B:349:ASN:ND2	1:B:411:MET:H	1.98	0.61
1:B:377:PRO:O	1:B:402:PRO:HB3	2.00	0.60
1:B:247:LYS:O	1:B:270:GLY:HA2	2.01	0.60
1:B:377:PRO:HB2	1:B:378:GLU:HG2	1.84	0.58
1:B:377:PRO:HB3	1:B:401:LEU:HB3	1.85	0.58
1:A:377:PRO:O	1:A:402:PRO:HB3	2.02	0.58
1:A:377:PRO:HB2	1:A:378:GLU:HG2	1.86	0.58
1:A:476:SER:OG	1:A:486:HIS:HE1	1.86	0.58
1:A:247:LYS:O	1:A:270:GLY:HA2	2.05	0.57

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:250:SER:H	1:B:267:GLN:NE2	2.01	0.57
1:A:335:VAL:HG22	1:A:343:ILE:HG22	1.87	0.57
1:A:349:ASN:ND2	1:A:411:MET:H	2.03	0.57
1:A:476:SER:O	1:A:497:HIS:HE1	1.88	0.56
1:B:497:HIS:HD2	5:B:2433:HOH:O	1.89	0.55
1:B:476:SER:OG	1:B:486:HIS:HE1	1.90	0.55
1:B:144:LYS:HE2	1:B:146:ILE:HD11	1.87	0.55
1:A:497:HIS:HD2	5:A:2405:HOH:O	1.89	0.54
1:B:396:GLY:HA3	1:B:495:LYS:HE2	1.89	0.54
1:A:117:VAL:CG2	1:A:151:MET:HE3	2.38	0.54
1:A:391:PRO:HG2	1:A:494:LYS:HD3	1.90	0.53
1:A:536:ASP:O	1:A:539:ARG:HG3	2.08	0.53
1:B:335:VAL:HG22	1:B:343:ILE:HG22	1.91	0.52
1:A:112:ASN:HA	1:A:138:ARG:HG2	1.91	0.52
1:B:580:ASP:HB3	1:B:583:ASN:HB2	1.92	0.52
1:A:128:ASN:H	1:A:128:ASN:HD22	1.58	0.51
1:B:550:ASN:ND2	1:B:553:GLY:H	2.08	0.51
1:B:463:LEU:HD12	1:B:504:GLY:O	2.11	0.51
1:B:152:GLY:CA	1:B:155:MET:HG2	2.41	0.51
1:A:132:ILE:HG12	1:A:147:LEU:HD11	1.93	0.51
1:B:103:SER:N	1:B:117:VAL:HG13	2.26	0.50
1:A:337:ARG:HD3	5:A:2025:HOH:O	2.11	0.50
1:A:250:SER:H	1:A:267:GLN:NE2	2.09	0.50
1:A:337:ARG:HG2	1:A:337:ARG:O	2.11	0.50
1:B:391:PRO:HG2	1:B:494:LYS:HD3	1.92	0.50
1:A:42:ASN:HD21	1:A:751:ARG:HH12	1.58	0.50
1:A:117:VAL:HG23	1:A:151:MET:HE2	1.93	0.49
1:A:107:ASP:HB3	1:A:110:ASP:O	2.12	0.48
1:A:550:ASN:ND2	1:A:553:GLY:H	2.09	0.48
1:B:707:PHE:CD2	1:B:721:ASN:HB3	2.49	0.48
1:B:172:ARG:HG2	1:B:190:TYR:CZ	2.49	0.48
1:A:377:PRO:HB3	1:A:401:LEU:HB3	1.96	0.48
1:B:520:VAL:HG11	1:B:568:TRP:CD1	2.49	0.47
1:A:639:LYS:HG3	1:A:680:VAL:HG12	1.95	0.47
1:B:299:GLY:HA3	1:B:327:ARG:O	2.14	0.47
1:B:128:ASN:HD22	1:B:128:ASN:H	1.63	0.47
1:A:103:SER:N	1:A:117:VAL:HG13	2.30	0.47
1:A:376:TYR:HA	1:A:377:PRO:C	2.36	0.47
1:B:639:LYS:HG3	1:B:680:VAL:HG12	1.97	0.46
1:B:250:SER:H	1:B:267:GLN:HE22	1.63	0.46
1:A:101:VAL:HG13	1:A:116:ILE:HG23	1.97	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:649:HIS:HD2	1:B:692:TYR:OH	1.98	0.46
1:A:200:PHE:HA	1:A:201:PRO:HD3	1.80	0.45
1:B:349:ASN:HD21	1:B:411:MET:H	1.60	0.45
1:A:139:GLY:HA2	1:A:142:TRP:CZ2	2.51	0.45
1:A:463:LEU:HD12	1:A:504:GLY:O	2.17	0.45
1:A:250:SER:H	1:A:267:GLN:HE22	1.65	0.45
1:A:368:LYS:HE2	1:A:442:TRP:HE1	1.83	0.44
1:A:368:LYS:HG2	5:A:2282:HOH:O	2.16	0.44
1:B:470:GLU:HG2	5:B:2177:HOH:O	2.17	0.44
1:B:52:MET:SD	1:B:69:ALA:HB2	2.57	0.44
1:B:251:ILE:H	1:B:267:GLN:NE2	2.09	0.44
1:A:490:LYS:HA	1:A:490:LYS:HD2	1.78	0.44
1:A:608:TYR:HB3	1:A:615:PHE:CZ	2.53	0.44
1:B:147:LEU:HD12	1:B:151:MET:HE1	1.99	0.44
1:B:128:ASN:ND2	1:B:128:ASN:H	2.16	0.44
1:A:117:VAL:HG23	1:A:151:MET:HG2	2.00	0.43
1:A:649:HIS:HD2	1:A:692:TYR:OH	2.01	0.43
1:B:608:TYR:HB3	1:B:615:PHE:CZ	2.53	0.43
1:A:587:VAL:HG22	5:A:2190:HOH:O	2.18	0.43
1:B:107:ASP:HB3	1:B:110:ASP:O	2.18	0.43
1:A:128:ASN:H	1:A:128:ASN:ND2	2.16	0.43
1:B:101:VAL:HG13	1:B:116:ILE:HG23	2.00	0.42
1:B:368:LYS:HE2	1:B:368:LYS:HB3	1.94	0.42
1:B:615:PHE:CE2	1:B:666:GLY:HA3	2.53	0.42
1:B:511:GLU:OE1	1:B:601:ARG:HD3	2.19	0.42
1:A:239:THR:HA	1:A:254:SER:O	2.18	0.42
1:A:606:LYS:HA	1:A:618:SER:O	2.20	0.42
1:A:161:GLY:HA2	1:A:222:VAL:O	2.19	0.42
1:A:418:ASP:OD2	1:A:424:ARG:HD2	2.19	0.42
1:A:152:GLY:CA	1:A:155:MET:HG2	2.49	0.42
1:A:103:SER:CB	1:A:117:VAL:CG1	2.87	0.42
1:A:117:VAL:CG2	1:A:151:MET:HG2	2.50	0.41
1:B:146:ILE:CD1	1:B:146:ILE:N	2.83	0.41
1:A:476:SER:O	1:A:497:HIS:CE1	2.72	0.41
1:A:181:GLY:HA2	1:A:203:PRO:HB2	2.03	0.41
1:A:707:PHE:CD2	1:A:721:ASN:HB3	2.55	0.41
1:B:42:ASN:HD21	1:B:751:ARG:NH1	2.14	0.40
1:A:175:TYR:CZ	1:A:186:ARG:HD3	2.55	0.40
1:A:349:ASN:HD21	1:A:411:MET:H	1.69	0.40
1:A:492:GLY:HA2	1:A:493:PRO:HD3	1.95	0.40
1:B:377:PRO:HB3	1:B:401:LEU:CB	2.49	0.40

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	726/737 (98%)	702 (97%)	23 (3%)	1 (0%)	56	48
1	B	726/737 (98%)	702 (97%)	23 (3%)	1 (0%)	56	48
All	All	1452/1474 (98%)	1404 (97%)	46 (3%)	2 (0%)	56	48

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	377	PRO
1	B	377	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	593/602 (98%)	582 (98%)	11 (2%)	65	58
1	B	594/602 (99%)	583 (98%)	11 (2%)	65	58
All	All	1187/1204 (99%)	1165 (98%)	22 (2%)	65	58

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	65	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	102	GLU
1	A	117	VAL
1	A	128	ASN
1	A	151	MET
1	A	375	MET
1	A	377	PRO
1	A	378	GLU
1	A	449	LYS
1	A	574	SER
1	A	707	PHE
1	B	102	GLU
1	B	117	VAL
1	B	146	ILE
1	B	151	MET
1	B	375	MET
1	B	377	PRO
1	B	378	GLU
1	B	425	MET
1	B	481	LEU
1	B	535	TYR
1	B	707	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (26) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	42	ASN
1	A	90	HIS
1	A	128	ASN
1	A	169	ASN
1	A	267	GLN
1	A	298	ASN
1	A	302	GLN
1	A	349	ASN
1	A	486	HIS
1	A	497	HIS
1	A	550	ASN
1	A	590	ASN
1	A	649	HIS
1	B	42	ASN
1	B	90	HIS
1	B	128	ASN
1	B	169	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	267	GLN
1	B	298	ASN
1	B	302	GLN
1	B	349	ASN
1	B	486	HIS
1	B	497	HIS
1	B	550	ASN
1	B	590	ASN
1	B	649	HIS

### 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 ⓘ

33 carbohydrates 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	GLC	A	1761	2	12,12,12	0.54	0	17,17,17	0.61	0
2	BGC	A	1762	2	11,11,12	0.60	0	14,15,17	0.64	0
2	BGC	A	1763	2	11,11,12	0.55	0	14,15,17	1.24	0
2	BGC	A	1764	2	11,11,12	0.60	0	14,15,17	1.00	1 (7%)
2	XYS	A	1765	2	9,9,10	1.12	1 (11%)	12,12,14	1.15	1 (8%)
2	XYS	A	1766	2	9,9,10	0.98	1 (11%)	12,12,14	1.06	1 (8%)
2	XYS	A	1767	2	9,9,10	1.32	1 (11%)	12,12,14	1.23	2 (16%)
2	GAL	A	1768	2	11,11,12	0.65	0	14,15,17	0.70	1 (7%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GLC	A	1769	3	12,12,12	0.55	0	17,17,17	0.87	0
3	BGC	A	1770	3	11,11,12	0.62	0	14,15,17	1.09	2 (14%)
3	BGC	A	1771	3	11,11,12	0.72	0	14,15,17	1.05	1 (7%)
3	BGC	A	1772	3	11,11,12	0.58	0	14,15,17	1.09	1 (7%)
3	XYS	A	1773	3	9,9,10	1.09	1 (11%)	12,12,14	1.30	2 (16%)
3	XYS	A	1774	3	9,9,10	1.26	1 (11%)	12,12,14	1.04	1 (8%)
3	XYS	A	1775	3	9,9,10	1.09	1 (11%)	12,12,14	1.04	1 (8%)
3	GAL	A	1776	3	11,11,12	0.60	0	14,15,17	0.81	1 (7%)
3	GAL	A	1777	3	11,11,12	0.60	0	14,15,17	1.51	2 (14%)
2	GLC	B	1761	2	12,12,12	0.49	0	17,17,17	0.58	0
2	BGC	B	1762	2	11,11,12	0.56	0	14,15,17	0.65	0
2	BGC	B	1763	2	11,11,12	0.57	0	14,15,17	1.24	4 (28%)
2	BGC	B	1764	2	11,11,12	0.60	0	14,15,17	1.02	1 (7%)
2	XYS	B	1765	2	9,9,10	1.12	1 (11%)	12,12,14	0.96	1 (8%)
2	XYS	B	1766	2	9,9,10	1.07	1 (11%)	12,12,14	1.05	1 (8%)
2	XYS	B	1767	2	9,9,10	1.30	1 (11%)	12,12,14	1.24	1 (8%)
2	GAL	B	1768	2	11,11,12	0.67	0	14,15,17	0.66	0
2	GLC	B	1769	2	12,12,12	0.61	0	17,17,17	0.95	0
2	BGC	B	1770	2	11,11,12	0.63	0	14,15,17	1.16	1 (7%)
2	BGC	B	1771	2	11,11,12	0.70	0	14,15,17	0.86	0
2	BGC	B	1772	2	11,11,12	0.61	0	14,15,17	1.16	1 (7%)
2	XYS	B	1773	2	9,9,10	1.14	1 (11%)	12,12,14	0.96	1 (8%)
2	XYS	B	1774	2	9,9,10	1.18	1 (11%)	12,12,14	1.22	1 (8%)
2	XYS	B	1775	2	9,9,10	1.15	1 (11%)	12,12,14	1.08	1 (8%)
2	GAL	B	1776	2	11,11,12	0.56	0	14,15,17	0.79	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	GLC	A	1761	2	1/1/5/5	0/2/22/22	0/1/1/1
2	BGC	A	1762	2	-	0/2/19/22	0/1/1/1
2	BGC	A	1763	2	-	0/2/19/22	0/1/1/1
2	BGC	A	1764	2	-	0/2/19/22	0/1/1/1
2	XYS	A	1765	2	-	0/0/14/17	0/1/1/1
2	XYS	A	1766	2	-	0/0/14/17	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	XYS	A	1767	2	-	0/0/14/17	0/1/1/1
2	GAL	A	1768	2	-	0/2/19/22	0/1/1/1
3	GLC	A	1769	3	1/1/5/5	0/2/22/22	0/1/1/1
3	BGC	A	1770	3	-	0/2/19/22	0/1/1/1
3	BGC	A	1771	3	-	0/2/19/22	0/1/1/1
3	BGC	A	1772	3	-	0/2/19/22	0/1/1/1
3	XYS	A	1773	3	-	0/0/14/17	0/1/1/1
3	XYS	A	1774	3	-	0/0/14/17	0/1/1/1
3	XYS	A	1775	3	-	0/0/14/17	0/1/1/1
3	GAL	A	1776	3	-	0/2/19/22	0/1/1/1
3	GAL	A	1777	3	-	0/2/19/22	0/1/1/1
2	GLC	B	1761	2	1/1/5/5	0/2/22/22	0/1/1/1
2	BGC	B	1762	2	-	0/2/19/22	0/1/1/1
2	BGC	B	1763	2	-	0/2/19/22	0/1/1/1
2	BGC	B	1764	2	-	0/2/19/22	0/1/1/1
2	XYS	B	1765	2	-	0/0/14/17	0/1/1/1
2	XYS	B	1766	2	-	0/0/14/17	0/1/1/1
2	XYS	B	1767	2	-	0/0/14/17	0/1/1/1
2	GAL	B	1768	2	-	0/2/19/22	0/1/1/1
2	GLC	B	1769	2	1/1/5/5	0/2/22/22	0/1/1/1
2	BGC	B	1770	2	-	0/2/19/22	0/1/1/1
2	BGC	B	1771	2	-	0/2/19/22	0/1/1/1
2	BGC	B	1772	2	-	0/2/19/22	0/1/1/1
2	XYS	B	1773	2	-	0/0/14/17	0/1/1/1
2	XYS	B	1774	2	-	0/0/14/17	0/1/1/1
2	XYS	B	1775	2	-	0/0/14/17	0/1/1/1
2	GAL	B	1776	2	-	0/2/19/22	0/1/1/1

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1767	XYS	O5-C1	-3.43	1.36	1.42
2	A	1767	XYS	O5-C1	-3.42	1.36	1.42
3	A	1774	XYS	O5-C1	-3.39	1.36	1.42
2	B	1774	XYS	O5-C1	-3.15	1.36	1.42
2	A	1765	XYS	O5-C1	-3.06	1.36	1.42
2	B	1775	XYS	O5-C1	-3.04	1.37	1.42
2	B	1765	XYS	O5-C1	-2.96	1.37	1.42
3	A	1773	XYS	O5-C1	-2.90	1.37	1.42
2	B	1773	XYS	O5-C1	-2.90	1.37	1.42
3	A	1775	XYS	O5-C1	-2.85	1.37	1.42
2	B	1766	XYS	O5-C1	-2.78	1.37	1.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1766	XYS	O5-C1	-2.57	1.37	1.42

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1773	XYS	O2-C2-C3	-2.60	104.89	110.12
2	B	1763	BGC	C6-C5-C4	-2.13	107.77	113.02
2	B	1763	BGC	O4-C4-C5	-2.10	103.66	109.24
2	A	1767	XYS	C1-C2-C3	-2.10	107.05	109.54
2	B	1763	BGC	C3-C4-C5	2.02	113.71	110.20
2	A	1768	GAL	C1-C2-C3	2.03	111.94	109.54
2	B	1764	BGC	C1-O5-C5	2.06	114.86	112.25
2	B	1763	BGC	C1-O5-C5	2.12	114.93	112.25
3	A	1776	GAL	C1-C2-C3	2.23	112.18	109.54
3	A	1774	XYS	C5-O5-C1	2.25	115.20	111.57
2	B	1773	XYS	C5-O5-C1	2.31	115.29	111.57
3	A	1777	GAL	O5-C5-C6	2.35	112.44	107.35
2	A	1764	BGC	C1-O5-C5	2.44	115.35	112.25
3	A	1770	BGC	C1-C2-C3	2.45	112.44	109.54
3	A	1770	BGC	C1-O5-C5	2.48	115.39	112.25
2	B	1765	XYS	C5-O5-C1	2.61	115.78	111.57
2	B	1774	XYS	C5-O5-C1	2.67	115.88	111.57
2	B	1767	XYS	C5-O5-C1	2.86	116.18	111.57
2	A	1765	XYS	C5-O5-C1	2.86	116.18	111.57
2	B	1775	XYS	C5-O5-C1	2.90	116.24	111.57
2	A	1766	XYS	C5-O5-C1	2.91	116.27	111.57
2	A	1767	XYS	C5-O5-C1	2.92	116.28	111.57
2	B	1770	BGC	C1-C2-C3	2.97	113.05	109.54
3	A	1773	XYS	C5-O5-C1	3.05	116.49	111.57
3	A	1771	BGC	C1-C2-C3	3.09	113.20	109.54
3	A	1775	XYS	C5-O5-C1	3.11	116.59	111.57
3	A	1772	BGC	C1-O5-C5	3.31	116.45	112.25
2	B	1766	XYS	C5-O5-C1	3.34	116.95	111.57
2	B	1772	BGC	C1-O5-C5	3.60	116.82	112.25
3	A	1777	GAL	C1-C2-C3	4.68	115.08	109.54

All (4) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	1761	GLC	C1
2	A	1761	GLC	C1
3	A	1769	GLC	C1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atom
2	B	1769	GLC	C1

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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 ⓘ

### 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	728/737 (98%)	-0.12	14 (1%) 70 78	7, 14, 22, 32	0
1	B	728/737 (98%)	-0.16	13 (1%) 71 80	7, 14, 24, 32	0
All	All	1456/1474 (98%)	-0.14	27 (1%) 70 78	7, 14, 23, 32	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	377	PRO	6.0
1	A	375	MET	5.5
1	B	375	MET	5.5
1	B	760	SER	4.8
1	B	377	PRO	4.6
1	A	760	SER	4.4
1	A	527	ASP	3.0
1	A	378	GLU	3.0
1	A	212	PHE	3.0
1	A	703	ASN	2.9
1	B	703	ASN	2.9
1	B	527	ASP	2.9
1	A	298	ASN	2.8
1	A	337	ARG	2.6
1	A	535	TYR	2.6
1	B	298	ASN	2.5
1	B	212	PHE	2.4
1	B	378	GLU	2.3
1	A	80	SER	2.3
1	B	400	GLN	2.2
1	B	80	SER	2.1
1	B	398	GLU	2.1
1	B	583	ASN	2.1
1	B	211	ASN	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	724	GLU	2.1
1	A	656	GLU	2.0
1	A	398	GLU	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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	XYS	A	1774	9/10	0.90	0.13	6.41	19,21,22,24	0
2	XYS	B	1774	9/10	0.94	0.10	1.28	18,20,21,23	0
3	XYS	A	1773	9/10	0.97	0.10	0.87	10,10,12,15	0
2	BGC	B	1770	11/12	0.97	0.08	0.14	8,9,10,11	0
2	GLC	B	1769	12/12	0.98	0.07	-0.26	9,10,13,14	0
2	XYS	B	1773	9/10	0.97	0.08	-0.37	12,12,14,16	0
3	BGC	A	1770	11/12	0.98	0.08	-0.39	8,9,10,10	0
3	GLC	A	1769	12/12	0.98	0.07	-0.54	8,10,10,13	0
2	BGC	B	1764	11/12	0.98	0.07	-0.58	10,11,11,13	0
2	XYS	A	1766	9/10	0.98	0.07	-0.99	10,11,13,13	0
2	BGC	A	1764	11/12	0.99	0.06	-1.12	8,9,10,12	0
2	XYS	B	1766	9/10	0.98	0.07	-1.61	12,13,14,15	0
2	XYS	A	1767	9/10	0.98	0.07	-	13,15,16,17	0
3	GAL	A	1777	11/12	0.84	0.21	-	25,26,27,27	11
2	BGC	A	1762	11/12	0.97	0.07	-	10,12,14,15	0
2	XYS	B	1765	9/10	0.94	0.13	-	18,19,20,22	0
2	GLC	B	1761	12/12	0.93	0.16	-	16,18,19,22	0
2	GAL	B	1768	11/12	0.92	0.20	-	20,23,25,27	0
2	XYS	B	1775	9/10	0.89	0.20	-	33,33,34,35	0
2	GAL	B	1776	11/12	0.89	0.17	-	21,24,26,27	0
3	BGC	A	1772	11/12	0.87	0.15	-	22,26,29,29	0
2	XYS	A	1765	9/10	0.96	0.11	-	17,17,19,19	0
3	XYS	A	1775	9/10	0.93	0.22	-	29,30,30,31	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	XYS	B	1767	9/10	0.96	0.07	-	15,16,17,18	0
2	BGC	B	1772	11/12	0.88	0.17	-	24,28,31,32	0
2	GAL	A	1768	11/12	0.91	0.20	-	21,23,24,25	0
3	GAL	A	1776	11/12	0.91	0.17	-	21,24,26,27	0
2	GLC	A	1761	12/12	0.94	0.13	-	14,18,20,23	0
2	BGC	B	1763	11/12	0.97	0.07	-	11,11,12,13	0
2	BGC	A	1763	11/12	0.98	0.07	-	10,10,11,11	0
2	BGC	B	1762	11/12	0.97	0.07	-	14,14,16,17	0
3	BGC	A	1771	11/12	0.97	0.08	-	11,13,16,18	0
2	BGC	B	1771	11/12	0.96	0.08	-	11,14,16,20	0

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	CA	B	1777	1/1	1.00	0.03	-2.26	12,12,12,12	0
4	CA	A	1778	1/1	1.00	0.03	-3.18	12,12,12,12	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.