



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 11:16 AM GMT

PDB ID : 3OG2  
Title : Native crystal structure of Trichoderma reesei beta-galactosidase  
Authors : Maksimainen, M.; Rouvinen, J.  
Deposited on : 2010-08-16  
Resolution : 1.20 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

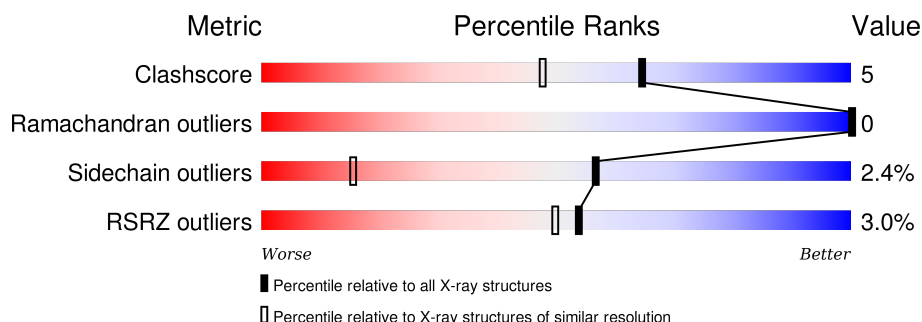
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.20 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))
Clashscore	102246	1607 (1.26-1.14)
Ramachandran outliers	100387	1540 (1.26-1.14)
Sidechain outliers	100360	1538 (1.26-1.14)
RSRZ outliers	91569	1500 (1.26-1.14)

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	1003	

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	MAN	A	1031	-	-	-	X
3	GLC	A	1042	-	-	-	X
3	GLC	A	1043	-	-	-	X
5	NAG	A	1044	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	NAG	A	1045	-	-	-	X
5	NAG	A	1047	-	-	-	X
6	GOL	A	1049	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9055 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-galactosidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	986	Total	C	N	O	S	0	27	0
			7773	5016	1298	1450	9			

- Molecule 2 is a polymer of unknown type called SUGAR (7-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	7	Total	C	N	O	0	0
			83	46	2	35		

- Molecule 3 is a polymer of unknown type called SUGAR (8-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	8	Total	C	N	O	0	0
			94	52	2	40		

- Molecule 4 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	4	Total	C	O	0	0
			44	24	20		

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		
6	A	1	Total	C	O	0	0
			6	3	3		

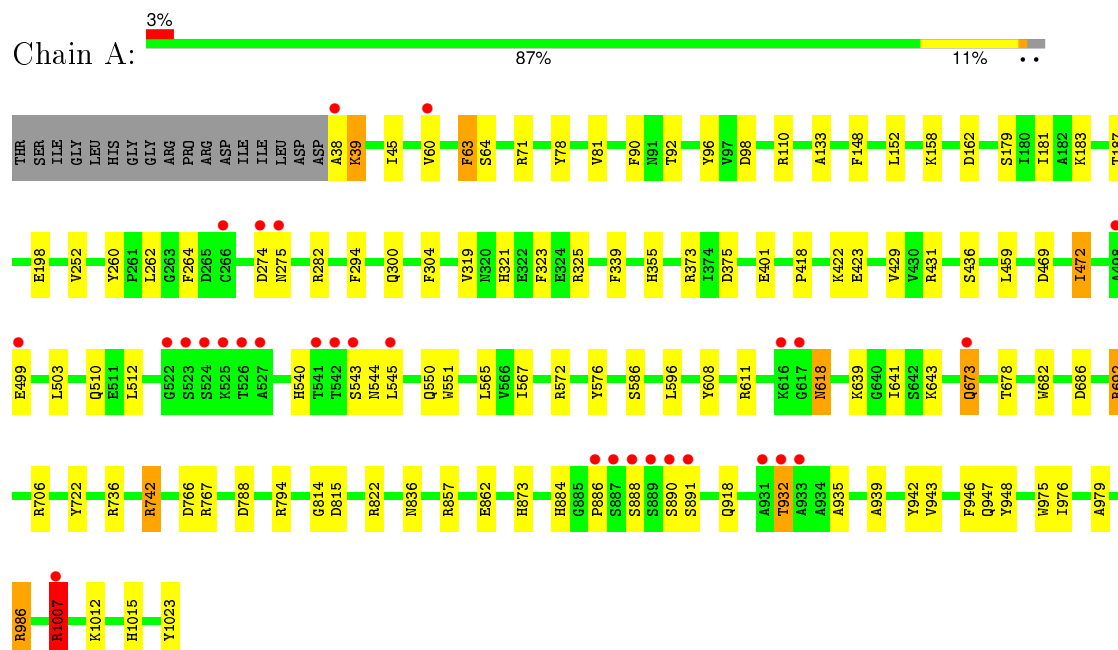
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	993	Total	O	0	0
			993	993		

### 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 ( $RSRZ > 2$ ). 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-galactosidase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.30 Å 69.10 Å 81.50 Å 109.10° 97.30° 114.50°	Depositor
Resolution (Å)	20.00 – 1.20 47.87 – 1.20	Depositor EDS
% Data completeness (in resolution range)	86.2 (20.00-1.20) 76.0 (47.87-1.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.55 (at 1.20 Å)	Xtriage
Refinement program	SHELXL-97	Depositor
R, $R_{free}$	0.139 , 0.171 (Not available) , (Not available)	Depositor DCC
$R_{free}$ test set	NotAvailable	DCC
Wilson B-factor (Å <sup>2</sup> )	8.8	Xtriage
Anisotropy	0.453	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 63.2	EDS
Estimated twinning fraction	0.014 for k,h,-h-k-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 719611 reflections	Xtriage
$F_o, F_c$ correlation	0.98	EDS
Total number of atoms	9055	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% 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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, GOL, BMA, NAG, MAN

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	0.71	1/8059 (0.0%)	1.25	66/10981 (0.6%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1023	TYR	C-OXT	5.92	1.34	1.23

All (66) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	942	TYR	CB-CG-CD2	11.84	128.10	121.00
1	A	794	ARG	NE-CZ-NH2	11.22	125.91	120.30
1	A	1007	ARG	CD-NE-CZ	11.22	139.31	123.60
1	A	71	ARG	NE-CZ-NH2	-10.91	114.84	120.30
1	A	1007	ARG	NE-CZ-NH1	10.07	125.34	120.30
1	A	942	TYR	CB-CG-CD1	-9.81	115.11	121.00
1	A	373	ARG	NE-CZ-NH1	9.18	124.89	120.30
1	A	986	ARG	NE-CZ-NH2	8.93	124.76	120.30
1	A	857	ARG	NE-CZ-NH1	8.73	124.66	120.30
1	A	110	ARG	NE-CZ-NH2	-8.29	116.16	120.30
1	A	71	ARG	NE-CZ-NH1	7.94	124.27	120.30
1	A	572	ARG	NE-CZ-NH1	7.71	124.15	120.30
1	A	722	TYR	CB-CG-CD2	-7.65	116.41	121.00
1	A	576	TYR	CB-CG-CD2	-7.57	116.46	121.00
1	A	942	TYR	CG-CD1-CE1	7.54	127.33	121.30
1	A	572	ARG	NE-CZ-NH2	-7.54	116.53	120.30
1	A	692	ARG	NE-CZ-NH1	7.43	124.02	120.30
1	A	1007	ARG	NE-CZ-NH2	-7.30	116.65	120.30
1	A	339	PHE	CB-CG-CD2	-7.16	115.79	120.80
1	A	767	ARG	NE-CZ-NH1	-7.14	116.73	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	282	ARG	NE-CZ-NH1	-7.07	116.77	120.30
1	A	742	ARG	NE-CZ-NH1	-6.98	116.81	120.30
1	A	766	ASP	CB-CG-OD1	6.93	124.54	118.30
1	A	431	ARG	NE-CZ-NH2	-6.55	117.03	120.30
1	A	862	GLU	CB-CA-C	-6.54	97.33	110.40
1	A	706	ARG	NE-CZ-NH2	-6.44	117.08	120.30
1	A	110	ARG	NE-CZ-NH1	6.40	123.50	120.30
1	A	551	TRP	CD1-CG-CD2	-6.39	101.19	106.30
1	A	576	TYR	CB-CG-CD1	6.35	124.81	121.00
1	A	942	TYR	CG-CD2-CE2	-6.31	116.25	121.30
1	A	96	TYR	CB-CG-CD2	-6.26	117.24	121.00
1	A	815	ASP	CB-CG-OD1	-6.26	112.67	118.30
1	A	686	ASP	CB-CG-OD1	6.23	123.90	118.30
1	A	822	ARG	NE-CZ-NH2	-6.19	117.21	120.30
1	A	325	ARG	NE-CZ-NH1	6.11	123.36	120.30
1	A	325	ARG	NE-CZ-NH2	-6.05	117.28	120.30
1	A	608	TYR	CB-CG-CD2	-6.04	117.38	121.00
1	A	551	TRP	CB-CG-CD1	5.98	134.77	127.00
1	A	63	PHE	CB-CG-CD1	5.96	124.97	120.80
1	A	788	ASP	CB-CG-OD1	5.86	123.58	118.30
1	A	162	ASP	CB-CG-OD2	-5.84	113.05	118.30
1	A	294	PHE	CB-CG-CD2	-5.82	116.73	120.80
1	A	932	THR	C-N-CA	-5.81	107.18	121.70
1	A	736	ARG	NE-CZ-NH2	5.80	123.20	120.30
1	A	767	ARG	NE-CZ-NH2	5.72	123.16	120.30
1	A	611	ARG	NE-CZ-NH2	-5.67	117.47	120.30
1	A	862	GLU	CG-CD-OE1	5.63	129.55	118.30
1	A	572	ARG	CD-NE-CZ	5.56	131.38	123.60
1	A	152	LEU	CB-CG-CD2	5.55	120.44	111.00
1	A	862	GLU	CG-CD-OE2	-5.51	107.28	118.30
1	A	469	ASP	CB-CG-OD1	5.47	123.22	118.30
1	A	611	ARG	NE-CZ-NH1	5.46	123.03	120.30
1	A	339	PHE	CB-CG-CD1	5.41	124.59	120.80
1	A	722	TYR	CB-CG-CD1	5.38	124.23	121.00
1	A	90	PHE	CB-CG-CD2	-5.37	117.04	120.80
1	A	608	TYR	CB-CG-CD1	5.29	124.18	121.00
1	A	975	TRP	CA-CB-CG	-5.26	103.70	113.70
1	A	551	TRP	CE2-CD2-CG	5.26	111.51	107.30
1	A	98	ASP	CB-CG-OD1	5.24	123.01	118.30
1	A	946	PHE	CB-CG-CD2	-5.20	117.16	120.80
1	A	706	ARG	CG-CD-NE	-5.19	100.89	111.80
1	A	282	ARG	NH1-CZ-NH2	5.19	125.11	119.40

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	148	PHE	CB-CG-CD2	-5.17	117.18	120.80
1	A	63	PHE	CB-CG-CD2	-5.13	117.21	120.80
1	A	375	ASP	CB-CG-OD1	5.13	122.92	118.30
1	A	431	ARG	NE-CZ-NH1	5.09	122.84	120.30

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	7773	0	7625	76	0
2	A	83	0	70	1	0
3	A	94	0	78	5	0
4	A	44	0	37	3	0
5	A	56	0	51	1	0
6	A	12	0	16	0	0
7	A	993	0	0	13	0
All	All	9055	0	7877	82	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 (82) 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:264[B]:PHE:CG	1:A:304[B]:PHE:CE2	2.03	1.44
1:A:264[B]:PHE:CB	1:A:304[B]:PHE:CE2	2.28	1.15
1:A:264[B]:PHE:CD1	1:A:304[B]:PHE:HE2	1.67	1.13
1:A:264[B]:PHE:CG	1:A:304[B]:PHE:HE2	1.55	1.09
1:A:264[B]:PHE:CB	1:A:304[B]:PHE:CZ	2.42	1.02
1:A:264[B]:PHE:HB3	1:A:304[B]:PHE:CE2	2.05	0.90
1:A:264[B]:PHE:HB2	1:A:304[B]:PHE:CZ	2.08	0.89
1:A:264[B]:PHE:CD2	1:A:304[B]:PHE:CE2	2.62	0.87
1:A:422:LYS:HG2	1:A:423:GLU:HG3	1.57	0.86

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:264[B]:PHE:CD1	1:A:304[B]:PHE:CE2	2.52	0.86
1:A:264[B]:PHE:HB3	1:A:304[B]:PHE:CZ	2.12	0.84
1:A:264[B]:PHE:CD2	1:A:304[B]:PHE:CD2	2.66	0.84
1:A:1007:ARG:HG3	1:A:1007:ARG:HH11	1.47	0.78
3:A:1043:GLC:H61	7:A:1818:HOH:O	1.83	0.78
1:A:512:LEU:HD11	1:A:550:GLN:HG2	1.67	0.75
4:A:1035:MAN:C2	4:A:1036:MAN:H5	2.22	0.70
1:A:158[B]:LYS:HE2	7:A:1567:HOH:O	1.92	0.67
1:A:60[B]:VAL:HG21	1:A:596:LEU:CD1	2.25	0.66
1:A:673:GLN:HE21	1:A:673:GLN:HA	1.60	0.66
1:A:429:VAL:HG22	1:A:472[A]:ILE:HD13	1.79	0.65
4:A:1035:MAN:H2	4:A:1036:MAN:H5	1.77	0.65
1:A:264[B]:PHE:CG	1:A:304[B]:PHE:CD2	2.77	0.63
1:A:1012:LYS:HE3	7:A:1704:HOH:O	1.99	0.61
1:A:262[B]:LEU:HD22	1:A:323[B]:PHE:CE1	2.37	0.60
1:A:304[B]:PHE:CZ	3:A:1042:GLC:H4	2.37	0.59
1:A:60[B]:VAL:HG21	1:A:596:LEU:HD12	1.83	0.59
3:A:1043:GLC:H62	7:A:1278:HOH:O	2.03	0.58
1:A:943:VAL:HG22	1:A:976[A]:ILE:CD1	2.34	0.58
1:A:886:PRO:HD2	1:A:891:SER:OG	2.05	0.57
1:A:935:ALA:HB2	1:A:986:ARG:HD2	1.86	0.57
1:A:510:GLN:HA	1:A:510:GLN:OE1	2.02	0.56
1:A:262[B]:LEU:HD22	1:A:323[B]:PHE:HE1	1.70	0.56
1:A:873:HIS:H	1:A:873:HIS:CD2	2.24	0.55
1:A:673:GLN:NE2	1:A:673:GLN:HA	2.22	0.54
1:A:133:ALA:HB1	1:A:181[A]:ILE:HD13	1.88	0.54
1:A:510:GLN:HG2	7:A:2036:HOH:O	2.10	0.51
1:A:78:TYR:O	1:A:81[A]:VAL:HG23	2.10	0.51
2:A:1031:MAN:H3	7:A:1825:HOH:O	2.09	0.51
1:A:38:ALA:N	1:A:39:LYS:HE2	2.25	0.51
1:A:565:LEU:HG	1:A:567[B]:ILE:HD11	1.92	0.51
1:A:678:THR:HG23	7:A:1647:HOH:O	2.10	0.50
1:A:401:GLU:HB2	1:A:418:PRO:HG2	1.94	0.49
1:A:304[B]:PHE:CE1	3:A:1042:GLC:H4	2.46	0.49
1:A:422:LYS:CG	1:A:423:GLU:HG3	2.36	0.49
1:A:544:ASN:O	1:A:545:LEU:HB2	2.12	0.49
1:A:459:LEU:HD13	1:A:540:HIS:HD2	1.79	0.47
1:A:321:HIS:HE1	7:A:1825:HOH:O	1.98	0.47
1:A:943:VAL:HG22	1:A:976[A]:ILE:HD12	1.96	0.46
1:A:692:ARG:NH1	7:A:1795:HOH:O	2.49	0.46
1:A:565:LEU:CG	1:A:567[B]:ILE:HD11	2.46	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:459:LEU:HD13	1:A:540:HIS:CD2	2.51	0.46
1:A:503:LEU:HD23	1:A:567[B]:ILE:HD12	1.98	0.46
1:A:45[B]:ILE:HD12	1:A:187:THR:HA	1.98	0.45
1:A:939:ALA:HA	1:A:979:ALA:O	2.17	0.45
1:A:264[B]:PHE:CD2	1:A:304[B]:PHE:HD2	2.33	0.45
1:A:158[B]:LYS:HD2	1:A:814:GLY:HA2	1.98	0.45
1:A:682:TRP:O	1:A:884:HIS:HD2	2.00	0.45
1:A:260:TYR:HA	1:A:300:GLN:HB2	1.99	0.44
1:A:304[B]:PHE:HZ	3:A:1042:GLC:H4	1.79	0.44
1:A:436:SER:OG	5:A:1046:NAG:H5	2.18	0.44
1:A:1007:ARG:NH1	1:A:1007:ARG:HG3	2.17	0.44
1:A:264[B]:PHE:HB3	1:A:304[B]:PHE:CD2	2.49	0.44
4:A:1035:MAN:H2	4:A:1036:MAN:C5	2.46	0.44
1:A:836:ASN:ND2	7:A:1996:HOH:O	2.51	0.43
1:A:319:VAL:HG13	1:A:323[A]:PHE:CD2	2.54	0.43
1:A:836:ASN:ND2	7:A:1549:HOH:O	2.49	0.43
1:A:918:GLN:HG3	1:A:1007:ARG:HD3	2.01	0.43
1:A:565:LEU:HD21	1:A:567[B]:ILE:HD11	2.01	0.42
1:A:873:HIS:HE1	7:A:1239:HOH:O	2.02	0.42
1:A:618:ASN:HD21	1:A:643:LYS:HD2	1.85	0.42
1:A:618:ASN:ND2	1:A:643:LYS:HD2	2.35	0.41
1:A:60[B]:VAL:HG21	1:A:596:LEU:HD13	1.99	0.41
1:A:262[B]:LEU:HD12	1:A:262[B]:LEU:HA	1.67	0.41
1:A:64:SER:HA	1:A:92:THR:O	2.20	0.41
1:A:565:LEU:CD2	1:A:567[B]:ILE:HD11	2.51	0.41
1:A:274:ASP:O	1:A:275:ASN:HB2	2.21	0.41
1:A:38:ALA:HB2	1:A:252:VAL:O	2.22	0.40
1:A:429:VAL:HG22	1:A:472[B]:ILE:HG12	2.03	0.40
1:A:422:LYS:HE2	1:A:423:GLU:OE2	2.21	0.40
1:A:499:GLU:N	1:A:499:GLU:OE1	2.49	0.40
1:A:179[B]:SER:OG	1:A:183:LYS:NZ	2.54	0.40
1:A:183:LYS:HG3	7:A:1498:HOH:O	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	1012/1003 (101%)	983 (97%)	29 (3%)	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	832/819 (102%)	812 (98%)	20 (2%)	57	15

All (20) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	39	LYS
1	A	63	PHE
1	A	198	GLU
1	A	355	HIS
1	A	472[A]	ILE
1	A	472[B]	ILE
1	A	543	SER
1	A	586	SER
1	A	618	ASN
1	A	639	LYS
1	A	641	ILE
1	A	673	GLN
1	A	742	ARG
1	A	888	SER
1	A	890	SER
1	A	932	THR
1	A	947	GLN
1	A	948	TYR
1	A	1007	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	1015	HIS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	168	HIS
1	A	225	ASN
1	A	413	ASN
1	A	540	HIS
1	A	618	ASN
1	A	664	HIS
1	A	673	GLN
1	A	747	GLN
1	A	836	ASN
1	A	870	GLN
1	A	873	HIS
1	A	884	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 ⓘ

19 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	NAG	A	1024	1,2	14,14,15	0.98	1 (7%)	15,19,21	1.11	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	A	1025	2	14,14,15	0.64	0	15,19,21	0.70	0
2	BMA	A	1026	2	11,11,12	0.79	0	14,15,17	0.89	0
2	MAN	A	1027	2	11,11,12	0.91	0	14,15,17	1.23	2 (14%)
2	MAN	A	1028	2	11,11,12	0.80	0	14,15,17	1.75	3 (21%)
2	MAN	A	1029	2	11,11,12	1.04	1 (9%)	14,15,17	1.54	5 (35%)
2	MAN	A	1031	2	11,11,12	1.25	1 (9%)	14,15,17	2.60	4 (28%)
3	NAG	A	1032	1,3	14,14,15	0.74	0	15,19,21	0.82	0
3	NAG	A	1033	3	14,14,15	0.57	0	15,19,21	1.70	4 (26%)
3	BMA	A	1034	3,4	11,11,12	0.77	0	14,15,17	1.57	3 (21%)
4	MAN	A	1035	3,4	11,11,12	0.51	0	14,15,17	1.85	4 (28%)
4	MAN	A	1036	4	11,11,12	0.78	0	14,15,17	1.91	2 (14%)
4	MAN	A	1037	4	11,11,12	0.83	1 (9%)	14,15,17	1.56	4 (28%)
4	MAN	A	1038	4	11,11,12	1.00	1 (9%)	14,15,17	2.53	5 (35%)
3	MAN	A	1039	3	11,11,12	0.85	0	14,15,17	2.50	3 (21%)
3	MAN	A	1040	3	11,11,12	0.74	0	14,15,17	0.99	1 (7%)
3	MAN	A	1041	3	11,11,12	0.89	1 (9%)	14,15,17	1.78	3 (21%)
3	GLC	A	1042	3	11,11,12	1.00	1 (9%)	14,15,17	2.68	4 (28%)
3	GLC	A	1043	3	11,11,12	0.68	0	14,15,17	2.91	5 (35%)

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	NAG	A	1024	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	1025	2	-	0/6/23/26	0/1/1/1
2	BMA	A	1026	2	-	0/2/19/22	0/1/1/1
2	MAN	A	1027	2	-	0/2/19/22	0/1/1/1
2	MAN	A	1028	2	-	0/2/19/22	0/1/1/1
2	MAN	A	1029	2	-	0/2/19/22	0/1/1/1
2	MAN	A	1031	2	-	0/2/19/22	0/1/1/1
3	NAG	A	1032	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1033	3	-	0/6/23/26	0/1/1/1
3	BMA	A	1034	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	1035	3,4	-	0/2/19/22	0/1/1/1
4	MAN	A	1036	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1037	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1038	4	-	0/2/19/22	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MAN	A	1039	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1040	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1041	3	-	0/2/19/22	0/1/1/1
3	GLC	A	1042	3	-	0/2/19/22	0/1/1/1
3	GLC	A	1043	3	-	0/2/19/22	0/1/1/1

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1042	GLC	O5-C1	-2.32	1.39	1.43
4	A	1037	MAN	C2-C3	2.28	1.55	1.52
2	A	1031	MAN	C2-C3	2.37	1.55	1.52
3	A	1041	MAN	C2-C3	2.39	1.55	1.52
2	A	1029	MAN	C2-C3	2.50	1.55	1.52
4	A	1038	MAN	C2-C3	2.58	1.56	1.52
2	A	1024	NAG	C1-C2	2.67	1.56	1.52

All (53) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1041	MAN	O3-C3-C2	-4.30	102.23	110.00
3	A	1033	NAG	C2-N2-C7	-4.16	117.70	123.04
2	A	1031	MAN	O3-C3-C4	-3.32	102.86	110.34
4	A	1038	MAN	C1-C2-C3	-3.27	105.68	109.54
3	A	1039	MAN	C6-C5-C4	-3.26	104.98	113.02
4	A	1038	MAN	C2-C3-C4	-3.09	105.80	111.04
2	A	1029	MAN	O3-C3-C2	-3.03	104.53	110.00
3	A	1034	BMA	C1-C2-C3	-2.96	106.04	109.54
2	A	1031	MAN	C6-C5-C4	-2.93	105.78	113.02
3	A	1033	NAG	C3-C2-N2	-2.93	103.55	110.56
3	A	1041	MAN	O6-C6-C5	-2.69	102.45	111.33
4	A	1036	MAN	C1-C2-C3	-2.58	106.49	109.54
2	A	1031	MAN	O6-C6-C5	-2.55	102.91	111.33
4	A	1038	MAN	O2-C2-C3	-2.44	105.22	110.12
3	A	1034	BMA	C1-O5-C5	-2.41	109.19	112.25
4	A	1037	MAN	O3-C3-C2	-2.39	105.68	110.00
4	A	1035	MAN	O3-C3-C4	-2.38	104.98	110.34
2	A	1029	MAN	O3-C3-C4	-2.36	105.03	110.34
4	A	1037	MAN	O6-C6-C5	-2.27	103.83	111.33
2	A	1029	MAN	O6-C6-C5	-2.24	103.93	111.33
4	A	1037	MAN	O5-C1-C2	-2.17	107.34	110.86
4	A	1038	MAN	O3-C3-C4	-2.14	105.53	110.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1042	GLC	O3-C3-C2	-2.08	106.23	110.00
2	A	1029	MAN	C3-C4-C5	-2.08	106.57	110.20
2	A	1027	MAN	O3-C3-C2	-2.08	106.24	110.00
2	A	1029	MAN	C1-C2-C3	-2.07	107.09	109.54
2	A	1024	NAG	C4-C3-C2	-2.07	108.02	111.23
2	A	1028	MAN	O2-C2-C1	-2.06	105.08	109.21
3	A	1033	NAG	O7-C7-N2	-2.04	117.71	121.86
3	A	1033	NAG	O4-C4-C3	-2.00	105.83	110.34
3	A	1039	MAN	C3-C4-C5	2.17	113.97	110.20
2	A	1028	MAN	C1-C2-C3	2.35	112.33	109.54
4	A	1035	MAN	O5-C5-C6	2.40	112.55	107.35
3	A	1043	GLC	O5-C1-C2	2.55	115.00	110.86
2	A	1027	MAN	C1-O5-C5	2.72	115.70	112.25
4	A	1037	MAN	C1-O5-C5	2.78	115.77	112.25
3	A	1034	BMA	O3-C3-C2	2.82	115.08	110.00
3	A	1040	MAN	C1-O5-C5	2.90	115.93	112.25
4	A	1035	MAN	C1-C2-C3	2.91	112.99	109.54
3	A	1042	GLC	O2-C2-C3	2.94	116.04	110.12
3	A	1043	GLC	C1-O5-C5	3.12	116.21	112.25
3	A	1041	MAN	O4-C4-C3	3.39	117.97	110.34
3	A	1043	GLC	O6-C6-C5	3.86	124.07	111.33
4	A	1035	MAN	C1-O5-C5	4.31	117.72	112.25
3	A	1042	GLC	C1-C2-C3	5.01	115.46	109.54
2	A	1028	MAN	C1-O5-C5	5.11	118.73	112.25
3	A	1043	GLC	O5-C5-C6	5.61	119.49	107.35
4	A	1036	MAN	C1-O5-C5	5.73	119.52	112.25
3	A	1042	GLC	C1-O5-C5	6.91	121.02	112.25
3	A	1043	GLC	C1-C2-C3	6.97	117.79	109.54
4	A	1038	MAN	C1-O5-C5	7.21	121.39	112.25
3	A	1039	MAN	C1-O5-C5	7.26	121.46	112.25
2	A	1031	MAN	C1-O5-C5	7.44	121.68	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1031	MAN	1	0
4	A	1035	MAN	3	0
4	A	1036	MAN	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1042	GLC	3	0
3	A	1043	GLC	2	0

## 5.6 Ligand geometry

6 ligands 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	$\# Z  > 2$
5	NAG	A	1044	1	14,14,15	0.71	0	15,19,21	1.46	3 (20%)
5	NAG	A	1045	1	14,14,15	0.46	0	15,19,21	1.43	3 (20%)
5	NAG	A	1046	1,5	14,14,15	0.62	0	15,19,21	2.37	5 (33%)
5	NAG	A	1047	5	14,14,15	0.57	0	15,19,21	1.01	1 (6%)
6	GOL	A	1048	-	5,5,5	0.76	0	5,5,5	0.43	0
6	GOL	A	1049	-	5,5,5	0.22	0	5,5,5	1.40	1 (20%)

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
5	NAG	A	1044	1	-	0/6/23/26	0/1/1/1
5	NAG	A	1045	1	-	0/6/23/26	0/1/1/1
5	NAG	A	1046	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1047	5	-	0/6/23/26	0/1/1/1
6	GOL	A	1048	-	-	0/4/4/4	0/0/0/0
6	GOL	A	1049	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1045	NAG	O7-C7-C8	-3.34	115.93	122.06
5	A	1046	NAG	O7-C7-C8	-2.97	116.61	122.06
5	A	1044	NAG	O7-C7-C8	-2.72	117.07	122.06
5	A	1047	NAG	O7-C7-C8	-2.42	117.63	122.06
5	A	1045	NAG	O3-C3-C2	-2.31	104.53	109.11
5	A	1046	NAG	O3-C3-C4	2.08	115.02	110.34
5	A	1044	NAG	O5-C5-C6	2.17	112.06	107.35
5	A	1045	NAG	C8-C7-N2	2.27	120.46	116.11
5	A	1044	NAG	O7-C7-N2	2.31	126.56	121.86
6	A	1049	GOL	O2-C2-C3	2.50	120.13	108.65
5	A	1046	NAG	C2-N2-C7	2.94	126.81	123.04
5	A	1046	NAG	C4-C3-C2	3.93	117.34	111.23
5	A	1046	NAG	C1-O5-C5	6.10	119.99	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1046	NAG	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	986/1003 (98%)	-0.30	30 (3%) 54 50	7, 12, 33, 69	0

All (30) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	525	LYS	5.5
1	A	526	THR	5.3
1	A	933	ALA	5.2
1	A	888	SER	5.0
1	A	886	PRO	4.4
1	A	543	SER	4.1
1	A	932	THR	4.1
1	A	931	ALA	4.1
1	A	889	SER	3.8
1	A	545	LEU	3.7
1	A	498	ALA	3.7
1	A	523	SER	3.3
1	A	38	ALA	2.9
1	A	890	SER	2.9
1	A	274	ASP	2.7
1	A	887	SER	2.7
1	A	60[A]	VAL	2.6
1	A	524	SER	2.5
1	A	275	ASN	2.4
1	A	499	GLU	2.4
1	A	541	THR	2.4
1	A	616	LYS	2.3
1	A	522	GLY	2.3
1	A	527	ALA	2.2
1	A	542	THR	2.1
1	A	1007	ARG	2.1
1	A	617	GLY	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	266[A]	CYS	2.0
1	A	891	SER	2.0
1	A	673	GLN	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	GLC	A	1043	11/12	0.79	0.34	13.19	66,75,88,91	0
2	MAN	A	1031	11/12	0.92	0.13	12.12	18,20,26,27	0
3	GLC	A	1042	11/12	0.81	0.18	3.50	35,50,55,66	0
3	MAN	A	1040	11/12	0.98	0.12	1.13	17,20,25,30	0
2	MAN	A	1028	11/12	0.99	0.06	1.10	11,11,17,28	0
3	NAG	A	1032	14/15	0.97	0.07	0.07	11,14,28,37	0
2	NAG	A	1025	14/15	0.99	0.04	-1.19	8,10,13,15	0
2	NAG	A	1024	14/15	0.99	0.04	-1.58	9,11,22,24	0
4	MAN	A	1036	11/12	0.58	0.41	-	41,55,72,74	0
4	MAN	A	1038	11/12	0.79	0.31	-	32,37,51,59	0
3	BMA	A	1034	11/12	0.98	0.07	-	14,15,17,18	0
4	MAN	A	1037	11/12	0.83	0.33	-	34,40,66,76	0
2	BMA	A	1026	11/12	0.98	0.07	-	10,12,15,17	0
3	MAN	A	1041	11/12	0.94	0.15	-	19,25,36,36	0
3	NAG	A	1033	14/15	0.97	0.09	-	12,19,45,46	0
2	MAN	A	1029	11/12	0.94	0.25	-	20,26,47,52	0
4	MAN	A	1035	11/12	0.95	0.17	-	21,26,34,39	0
2	MAN	A	1027	11/12	0.99	0.04	-	9,10,14,22	0
3	MAN	A	1039	11/12	0.96	0.13	-	18,25,38,45	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(Å <sup>2</sup> )	Q<0.9
5	NAG	A	1047	14/15	0.80	0.24	17.61	55,86,102,108	0
6	GOL	A	1049	6/6	0.80	0.17	11.27	23,48,51,59	0
5	NAG	A	1045	14/15	0.91	0.21	7.17	27,42,59,65	0
5	NAG	A	1044	14/15	0.91	0.12	5.29	23,35,50,55	0
5	NAG	A	1046	14/15	0.89	0.16	0.96	29,43,58,68	0
6	GOL	A	1048	6/6	0.99	0.04	-1.84	11,12,12,14	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.