



wwPDB X-ray Structure Validation Summary Report

Feb 27, 2014 – 12:29 PM GMT

PDB ID : 3OGS
Title : Complex structure of beta-galactosidase from Trichoderma reesei with IPTG
Authors : Maksimainen, M.; Rouvinen, J.
Deposited on : 2010-08-17
Resolution : 1.75 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

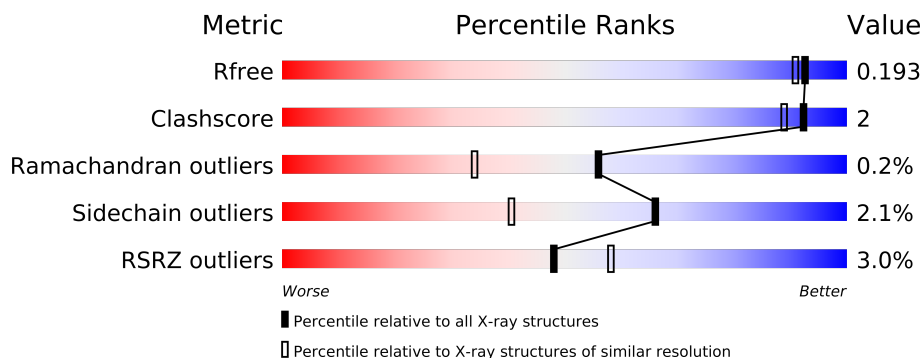
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.75 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}	66092	1134 (1.76-1.76)
Clashscore	79885	1304 (1.76-1.76)
Ramachandran outliers	78287	1288 (1.76-1.76)
Sidechain outliers	78261	1288 (1.76-1.76)
RSRZ outliers	66119	1135 (1.76-1.76)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	1003	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	NAG	A	1041	-	X
5	NAG	A	1042	-	X

2 Entry composition i

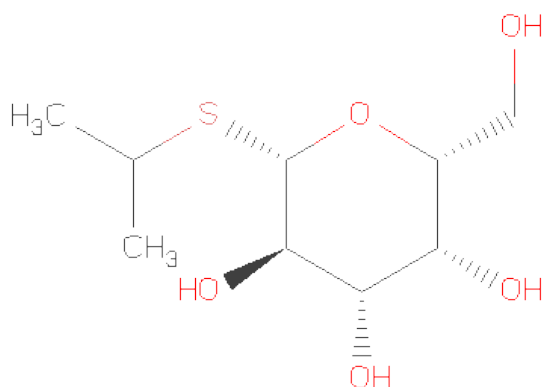
There are 7 unique types of molecules in this entry. The entry contains 8819 atoms, of which 0 are hydrogen and 0 are deuterium.

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
			Total	C	N	O	S			
1	A	986	7634	4906	1289	1431	8	0	3	0

- Molecule 2 is ISOPROPYL-1-BETA-D-THIOGALACTOSIDE (three-letter code: IPT) (formula: C₉H₁₈O₅S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
2	A	1	15	9	5	1	0	0

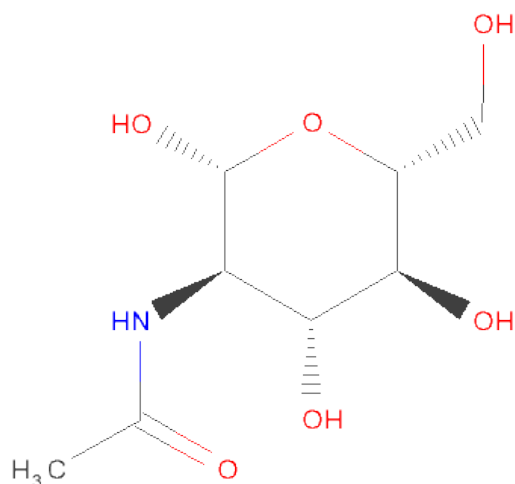
- Molecule 3 is a polymer of unknown type called SUGAR (7-MER).

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

- Molecule 4 is a polymer of unknown type called SUGAR (9-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	9	Total	C	N	O	0	0
			105	58	2	45		

- 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		

- Molecule 6 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	2	Total	C	N	O	0	0
			28	16	2	10		

- Molecule 7 is water.

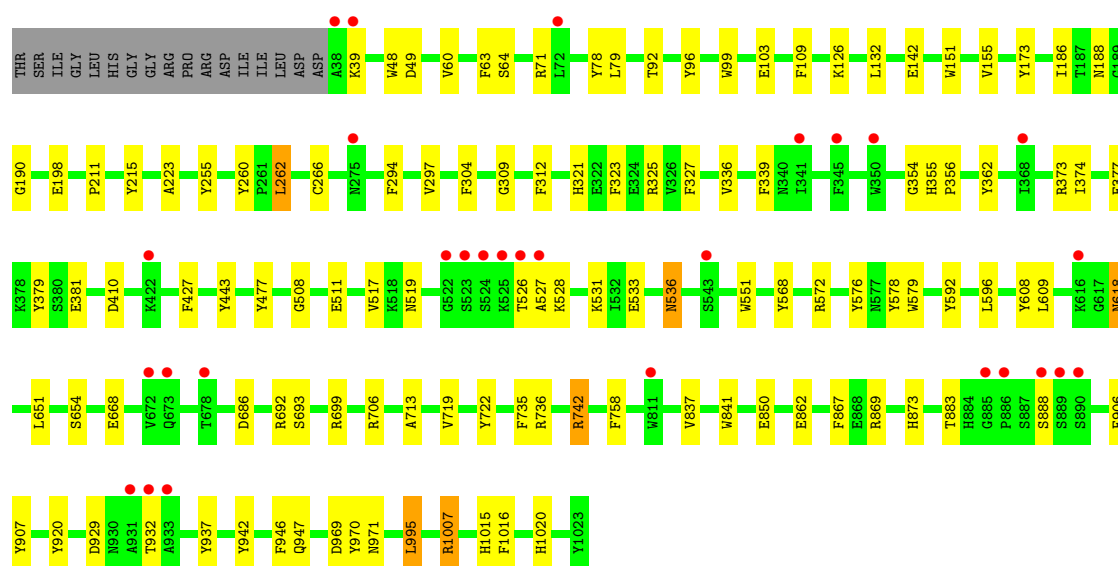
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	926	Total	O	0	0
			926	926		

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($\text{RSRZ} > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Beta-galactosidase

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	69.50Å 70.30Å 82.40Å 108.50° 97.80° 114.40°	Depositor
Resolution (Å)	43.54 – 1.75 43.54 – 1.75	Depositor EDS
% Data completeness (in resolution range)	95.0 (43.54-1.75) 91.2 (43.54-1.75)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.28 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.145 , 0.200 0.160 , 0.193	Depositor DCC
R_{free} test set	5893 reflections (5.26%)	DCC
Wilson B-factor (Å ²)	18.2	Xtriage
Anisotropy	0.590	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 54.3	EDS
Estimated twinning fraction	0.019 for k,h,-h-k-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 117858 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	8819	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, BMA, NAG, IPT, 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	1.78	85/7862 (1.1%)	0.99	18/10715 (0.2%)

The worst 5 of 85 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	260	TYR	CD1-CE1	8.49	1.52	1.39
1	A	937	TYR	CD1-CE1	8.14	1.51	1.39
1	A	109	PHE	CE2-CZ	7.80	1.52	1.37
1	A	103	GLU	CB-CG	7.63	1.66	1.52
1	A	362	TYR	CE2-CZ	7.43	1.48	1.38

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	869	ARG	NE-CZ-NH1	9.01	124.80	120.30
1	A	869	ARG	NE-CZ-NH2	-8.48	116.06	120.30
1	A	49	ASP	CB-CG-OD1	8.34	125.81	118.30
1	A	572	ARG	NE-CZ-NH1	6.96	123.78	120.30
1	A	373	ARG	NE-CZ-NH1	6.36	123.48	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the

chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	7634	0	7435	26	0
2	A	15	0	18	0	0
3	A	83	0	70	0	0
4	A	105	0	88	2	0
5	A	28	0	26	0	0
6	A	28	0	25	0	0
7	A	926	0	0	15	0
All	All	8819	0	7662	28	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 2.

The worst 5 of 28 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:1020:HIS:HB2	7:A:1807:HOH:O	1.20	1.34
1:A:188:ASN:HB3	7:A:1881:HOH:O	1.48	1.14
1:A:531:LYS:HB3	7:A:1575:HOH:O	1.62	0.97
1:A:971:ASN:HB3	7:A:1816:HOH:O	1.67	0.95
4:A:1040:GLC:H61	7:A:1952:HOH:O	1.79	0.82

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	987/1003 (98%)	957 (97%)	28 (3%)	2 (0%)	56 34

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	527	ALA
1	A	519	ASN

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	808/819 (99%)	791 (98%)	17 (2%)	66 40

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	536	ASN
1	A	618	ASN
1	A	947	GLN
1	A	528	LYS
1	A	995	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	618	ASN
1	A	1015	HIS
1	A	873	HIS
1	A	412	GLN
1	A	836	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

18 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
3	NAG	A	1025	1,3	12,14,15	1.37	1 (8%)	15,19,21	1.72	2 (13%)
3	NAG	A	1026	3	12,14,15	0.98	0	15,19,21	0.96	1 (6%)
3	BMA	A	1027	3	10,11,12	1.29	1 (10%)	11,15,17	1.47	3 (27%)
3	MAN	A	1028	3	10,11,12	1.33	1 (10%)	11,15,17	1.51	2 (18%)
3	MAN	A	1029	3	10,11,12	1.15	1 (10%)	11,15,17	1.60	3 (27%)
3	MAN	A	1030	3	10,11,12	1.00	1 (10%)	11,15,17	1.56	3 (27%)
3	MAN	A	1031	3	10,11,12	0.71	0	11,15,17	1.61	3 (27%)
4	NAG	A	1032	1,4	12,14,15	1.29	1 (8%)	15,19,21	1.56	5 (33%)
4	NAG	A	1033	4	12,14,15	0.80	0	15,19,21	1.56	3 (20%)
4	BMA	A	1034	4	10,11,12	1.00	0	11,15,17	1.75	3 (27%)
4	MAN	A	1035	4	10,11,12	0.62	0	11,15,17	3.40	2 (18%)
4	MAN	A	1036	4	10,11,12	0.57	0	11,15,17	1.38	1 (9%)
4	MAN	A	1037	4	10,11,12	0.81	0	11,15,17	2.71	5 (45%)
4	MAN	A	1038	4	10,11,12	0.79	0	11,15,17	1.78	2 (18%)
4	MAN	A	1039	4	10,11,12	0.84	0	11,15,17	2.36	5 (45%)
4	GLC	A	1040	4	10,11,12	0.63	0	11,15,17	1.96	2 (18%)
6	NAG	A	1043	1,6	12,14,15	0.53	0	15,19,21	2.34	4 (26%)
6	NAG	A	1044	6	12,14,15	0.71	0	15,19,21	1.10	1 (6%)

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	NAG	A	1025	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1026	3	-	0/6/23/26	0/1/1/1
3	BMA	A	1027	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1028	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1029	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1030	3	-	0/2/19/22	0/1/1/1
3	MAN	A	1031	3	-	0/2/19/22	0/1/1/1
4	NAG	A	1032	1,4	-	0/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	1033	4	-	0/6/23/26	0/1/1/1
4	BMA	A	1034	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1035	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
4	MAN	A	1039	4	-	0/2/19/22	0/1/1/1
4	GLC	A	1040	4	-	0/2/19/22	0/1/1/1
6	NAG	A	1043	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	1044	6	-	0/6/23/26	0/1/1/1

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1025	NAG	C2-N2	3.02	1.49	1.46
4	A	1032	NAG	C3-C2	2.87	1.58	1.52
3	A	1027	BMA	O4-C4	2.67	1.49	1.43
3	A	1030	MAN	O5-C5	-2.53	1.40	1.45
3	A	1029	MAN	O2-C2	2.41	1.48	1.43

The worst 5 of 50 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1035	MAN	O5-C5-C6	10.03	117.51	106.98
6	A	1043	NAG	O5-C5-C6	7.66	115.01	106.98
4	A	1037	MAN	C6-C5-C4	-6.70	96.82	113.00
4	A	1040	GLC	O5-C5-C6	5.06	112.29	106.98
4	A	1038	MAN	O5-C5-C6	4.86	112.09	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.6 Ligand geometry ⓘ

3 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	$\# Z > 2$
2	IPT	A	1024	-	15,15,15	1.25	1 (6%)	21,21,21	1.70	6 (28%)
5	NAG	A	1041	1	12,14,15	0.67	0	15,19,21	1.38	1 (6%)
5	NAG	A	1042	1	12,14,15	0.79	1 (8%)	15,19,21	1.46	4 (26%)

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	IPT	A	1024	-	-	0/6/26/26	0/1/1/1
5	NAG	A	1041	1	-	0/6/23/26	0/1/1/1
5	NAG	A	1042	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	1024	IPT	C4-C5	2.99	1.59	1.53
5	A	1042	NAG	O5-C5	-2.23	1.41	1.45

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	1041	NAG	O5-C5-C4	4.29	116.10	110.65
2	A	1024	IPT	C1-O5-C5	3.90	120.39	112.70
5	A	1042	NAG	O3-C3-C2	-3.02	102.75	109.09
2	A	1024	IPT	C2-C1-S1	-2.92	107.46	110.97
5	A	1042	NAG	O7-C7-C8	-2.81	116.55	122.04

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 95th 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 ²)	Q<0.9
1	A	986/1003 (98%)	-0.05	29 (2%) 49 59	10, 17, 33, 64	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	933	ALA	7.5
1	A	526	THR	6.9
1	A	525	LYS	6.3
1	A	932	THR	6.3
1	A	523	SER	5.0

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 ⓘ

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, 95th 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
3	MAN	A	1030	11/12	0.26	30.08	30,35,41,47	0
6	NAG	A	1044	14/15	0.34	10.55	46,56,63,63	0
4	NAG	A	1033	14/15	0.13	8.47	16,23,37,39	0
4	MAN	A	1037	11/12	0.10	7.25	22,25,31,34	0
4	GLC	A	1040	11/12	0.16	5.82	41,47,48,49	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	MAN	A	1038	11/12	0.13	2.60	17,22,26,27	0
4	MAN	A	1039	11/12	0.10	1.54	24,30,35,38	0
3	MAN	A	1031	11/12	0.08	1.28	24,26,28,30	0
4	BMA	A	1034	11/12	0.10	0.97	19,21,25,29	0
6	NAG	A	1043	14/15	0.14	0.91	32,37,44,46	0
4	NAG	A	1032	14/15	0.10	0.05	13,18,24,30	0
3	NAG	A	1025	14/15	0.08	-0.18	13,15,25,27	0
3	MAN	A	1028	11/12	0.07	-0.25	13,15,18,25	0
3	MAN	A	1029	11/12	0.06	-1.06	15,16,28,35	0
3	NAG	A	1026	14/15	0.06	-1.23	13,15,21,22	0
3	BMA	A	1027	11/12	0.06	-16.00	13,16,19,19	0
4	MAN	A	1036	11/12	0.42	-	55,62,64,65	0
4	MAN	A	1035	11/12	0.22	-	34,40,45,48	0

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. 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, 95th 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	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	NAG	A	1042	14/15	0.25	7.25	33,39,48,52	0
5	NAG	A	1041	14/15	0.21	3.09	39,48,56,59	0
2	IPT	A	1024	15/15	0.11	0.13	9,12,26,27	0

6.5 Other polymers

There are no such residues in this entry.