



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 20, 2016 – 11:15 PM EDT

PDB ID : 5HJO  
Title : Murine endoplasmic reticulum alpha-glucosidase II with bound substrate analogue  
Authors : Caputo, A.T.; Roversi, P.; Alonzi, D.S.; Kiappes, J.L.; Zitzmann, N.  
Deposited on : 2016-01-13  
Resolution : 2.29 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027939  
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) : rb-20027939

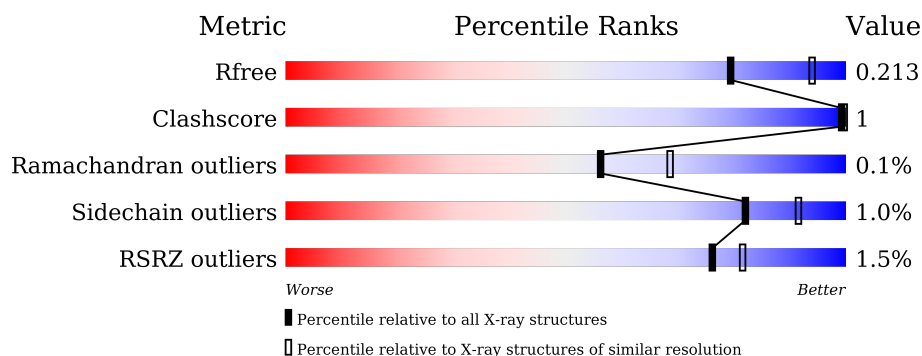
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.29 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	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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	857	<div> <div style="width: 97%;"></div> <div>97%</div> </div>
1	C	857	<div> <div style="width: 96%;"></div> <div>96%</div> </div>
2	B	83	<div> <div style="width: 98%;"></div> <div>98%</div> </div>
2	D	83	<div> <div style="width: 95%;"></div> <div>95%</div> </div>

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

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	EDO	C	1007	-	-	-	X
8	FMT	A	1009	-	-	-	X
8	FMT	C	1011	-	-	-	X
8	FMT	C	1012	-	-	-	X
8	FMT	C	1016	-	-	-	X

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 30426 atoms, of which 14497 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 Neutral alpha-glucosidase AB.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	854	Total	C	H	N	O	S	0	2	0
			13547	4419	6645	1195	1260	28			
1	C	854	Total	C	H	N	O	S	0	3	0
			13554	4422	6647	1195	1262	28			

There are 198 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	VAL	deletion	UNP Q8BHN3
A	?	-	PRO	deletion	UNP Q8BHN3
A	?	-	PHE	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	VAL	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	VAL	deletion	UNP Q8BHN3
A	?	-	TRP	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ILE	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ASN	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	PHE	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	ARG	deletion	UNP Q8BHN3
A	?	-	GLN	deletion	UNP Q8BHN3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	PRO	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	ASN	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	GLN	deletion	UNP Q8BHN3
A	?	-	PRO	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	THR	deletion	UNP Q8BHN3
A	?	-	PRO	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	PRO	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	THR	deletion	UNP Q8BHN3
A	?	-	GLN	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	THR	deletion	UNP Q8BHN3
A	?	-	ALA	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3
A	?	-	THR	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	PHE	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	LYS	deletion	UNP Q8BHN3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	MET	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	ASP	deletion	UNP Q8BHN3
A	?	-	TYR	deletion	UNP Q8BHN3
A	?	-	LEU	deletion	UNP Q8BHN3
A	?	-	GLN	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	SER	deletion	UNP Q8BHN3
A	?	-	GLY	deletion	UNP Q8BHN3
A	?	-	GLU	deletion	UNP Q8BHN3
A	945	THR	-	expression tag	UNP Q8BHN3
A	946	SER	-	expression tag	UNP Q8BHN3
A	947	VAL	-	expression tag	UNP Q8BHN3
A	948	LEU	-	expression tag	UNP Q8BHN3
A	949	ILE	-	expression tag	UNP Q8BHN3
A	950	LEU	-	expression tag	UNP Q8BHN3
A	951	ARG	-	expression tag	UNP Q8BHN3
A	952	LYS	-	expression tag	UNP Q8BHN3
A	953	PRO	-	expression tag	UNP Q8BHN3
A	954	GLY	-	expression tag	UNP Q8BHN3
A	955	VAL	-	expression tag	UNP Q8BHN3
A	956	SER	-	expression tag	UNP Q8BHN3
A	957	VAL	-	expression tag	UNP Q8BHN3
A	958	ALA	-	expression tag	UNP Q8BHN3
A	959	SER	-	expression tag	UNP Q8BHN3
A	960	ASP	-	expression tag	UNP Q8BHN3
A	961	TRP	-	expression tag	UNP Q8BHN3
A	962	SER	-	expression tag	UNP Q8BHN3
A	963	ILE	-	expression tag	UNP Q8BHN3
A	964	HIS	-	expression tag	UNP Q8BHN3
A	965	LEU	-	expression tag	UNP Q8BHN3
A	966	ARG	-	expression tag	UNP Q8BHN3
C	?	-	VAL	deletion	UNP Q8BHN3
C	?	-	PRO	deletion	UNP Q8BHN3
C	?	-	PHE	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	VAL	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	VAL	deletion	UNP Q8BHN3
C	?	-	TRP	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	ILE	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	ASN	deletion	UNP Q8BHN3
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	PHE	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	ARG	deletion	UNP Q8BHN3
C	?	-	GLN	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	PRO	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	ASN	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3
C	?	-	GLN	deletion	UNP Q8BHN3
C	?	-	PRO	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3
C	?	-	THR	deletion	UNP Q8BHN3
C	?	-	PRO	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	PRO	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	THR	deletion	UNP Q8BHN3
C	?	-	GLN	deletion	UNP Q8BHN3

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	THR	deletion	UNP Q8BHN3
C	?	-	ALA	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	THR	deletion	UNP Q8BHN3
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	PHE	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	LYS	deletion	UNP Q8BHN3
C	?	-	MET	deletion	UNP Q8BHN3
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	ASP	deletion	UNP Q8BHN3
C	?	-	TYR	deletion	UNP Q8BHN3
C	?	-	LEU	deletion	UNP Q8BHN3
C	?	-	GLN	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	SER	deletion	UNP Q8BHN3
C	?	-	GLY	deletion	UNP Q8BHN3
C	?	-	GLU	deletion	UNP Q8BHN3
C	945	THR	-	expression tag	UNP Q8BHN3
C	946	SER	-	expression tag	UNP Q8BHN3
C	947	VAL	-	expression tag	UNP Q8BHN3
C	948	LEU	-	expression tag	UNP Q8BHN3
C	949	ILE	-	expression tag	UNP Q8BHN3
C	950	LEU	-	expression tag	UNP Q8BHN3
C	951	ARG	-	expression tag	UNP Q8BHN3
C	952	LYS	-	expression tag	UNP Q8BHN3
C	953	PRO	-	expression tag	UNP Q8BHN3
C	954	GLY	-	expression tag	UNP Q8BHN3
C	955	VAL	-	expression tag	UNP Q8BHN3
C	956	SER	-	expression tag	UNP Q8BHN3
C	957	VAL	-	expression tag	UNP Q8BHN3
C	958	ALA	-	expression tag	UNP Q8BHN3
C	959	SER	-	expression tag	UNP Q8BHN3
C	960	ASP	-	expression tag	UNP Q8BHN3
C	961	TRP	-	expression tag	UNP Q8BHN3

*Continued on next page...*



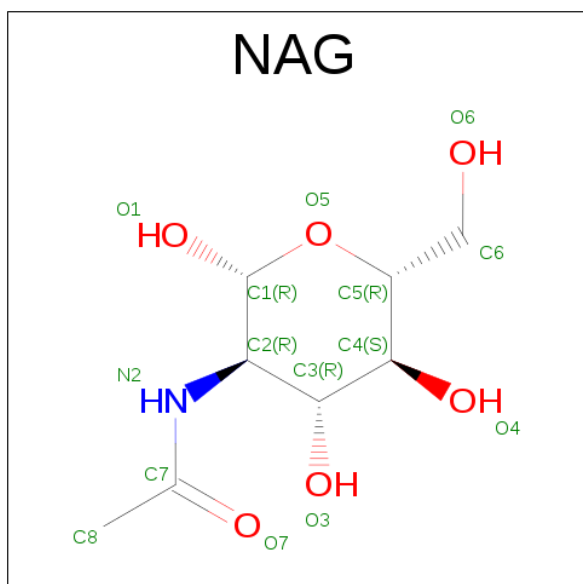
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	962	SER	-	expression tag	UNP Q8BHN3
C	963	ILE	-	expression tag	UNP Q8BHN3
C	964	HIS	-	expression tag	UNP Q8BHN3
C	965	LEU	-	expression tag	UNP Q8BHN3
C	966	ARG	-	expression tag	UNP Q8BHN3

- Molecule 2 is a protein called Glucosidase 2 subunit beta.

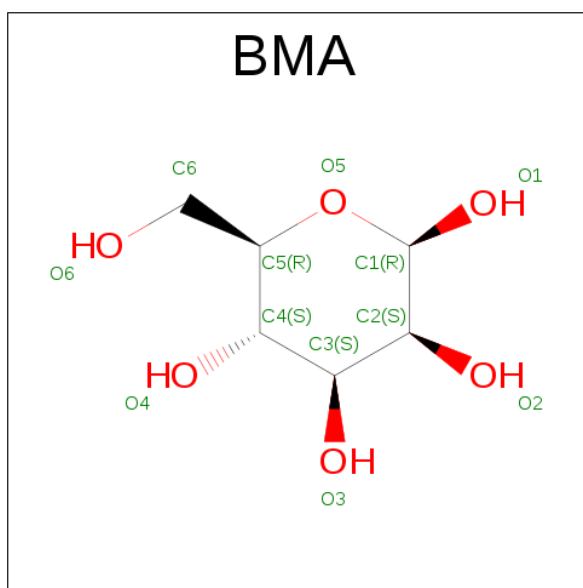
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	83	Total	C	H	N	O	S	0	0	0
			1149	368	533	101	137	10			
2	D	83	Total	C	H	N	O	S	0	0	0
			1149	368	533	101	137	10			

- Molecule 3 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



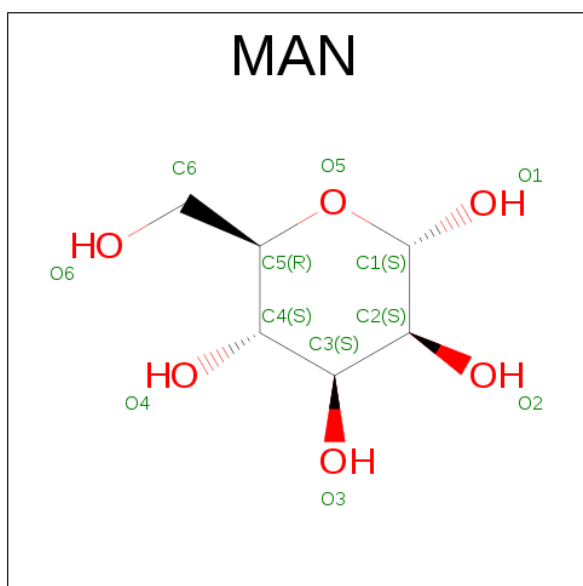
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
3	A	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
3	C	1	Total	C	H	N	O	0	0
			26	8	12	1	5		
3	C	1	Total	C	H	N	O	0	0
			26	8	12	1	5		

- Molecule 4 is BETA-D-MANNOSE (three-letter code: BMA) (formula:  $C_6H_{12}O_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			20	6	9	5		
4	C	1	Total	C	H	O	0	0
			20	6	9	5		

- Molecule 5 is ALPHA-D-MANNOSE (three-letter code: MAN) (formula:  $C_6H_{12}O_6$ ).



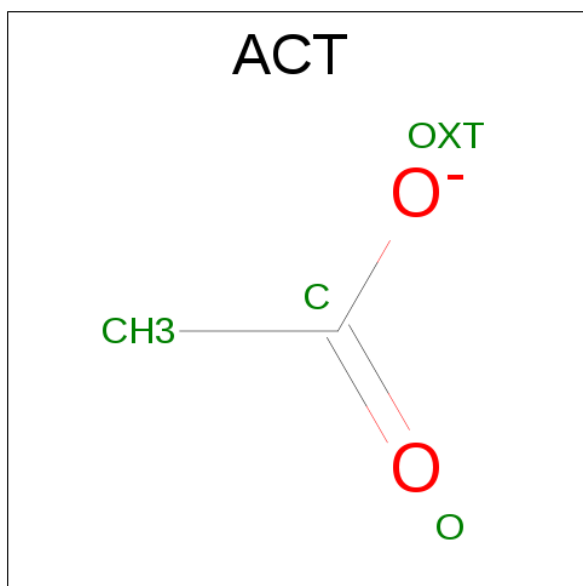
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	H	O	0	0
			21	6	10	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	C	1	Total	C	H	O	0	0
			21	6	10	5		

- Molecule 6 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



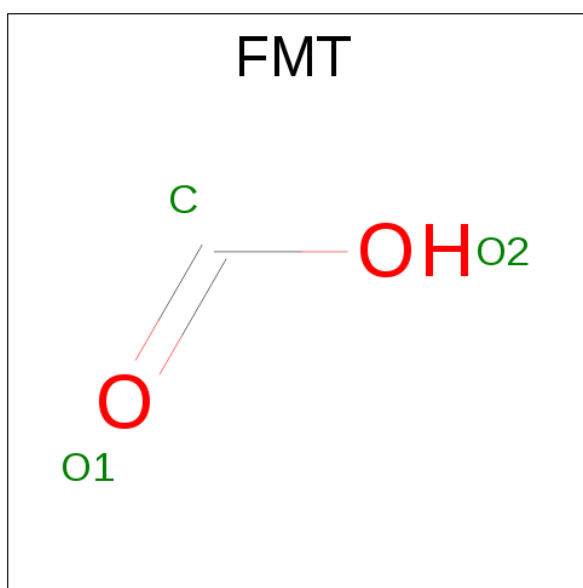
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	H	O	0	0
			7	2	3	2		
6	C	1	Total	C	H	O	0	0
			7	2	3	2		

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	H	O	0	0
			10	2	6	2		
7	C	1	Total	C	H	O	0	0
			10	2	6	2		

- Molecule 8 is FORMIC ACID (three-letter code: FMT) (formula:  $\text{CH}_2\text{O}_2$ ).



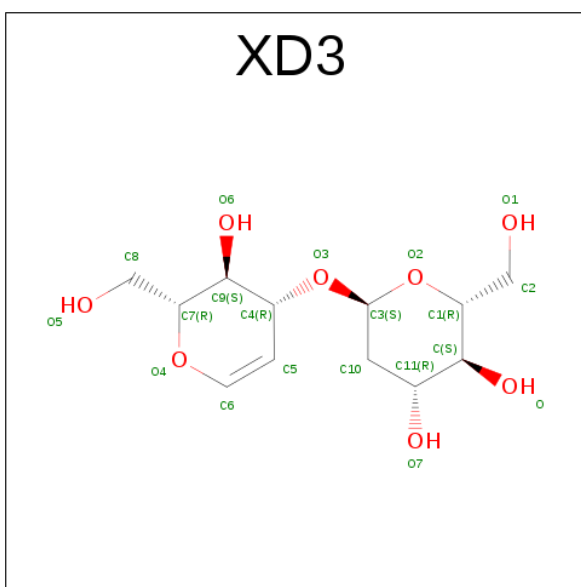
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	H	O	0	0
			4	1	1	2		
8	A	1	Total	C	H	O	0	0
			5	1	2	2		

*Continued on next page...*

*Continued from previous page...*

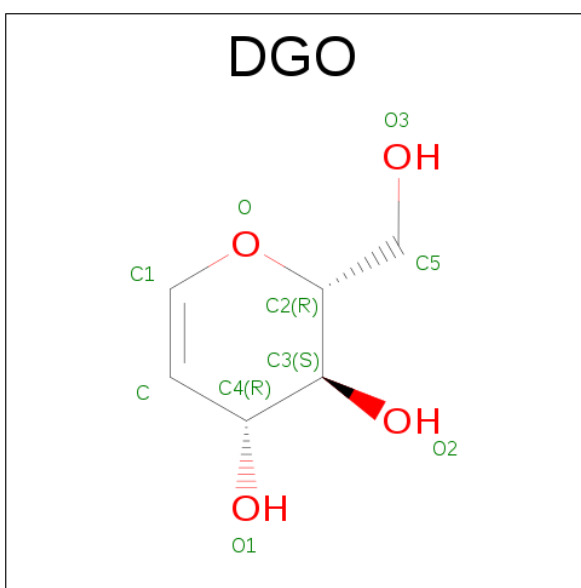
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	H	O	0	0
			5	1	2	2		
8	A	1	Total	C	H	O	0	0
			5	1	2	2		
8	A	1	Total	C	H	O	0	0
			4	1	1	2		
8	A	1	Total	C	H	O	0	0
			4	1	1	2		
8	A	1	Total	C	H	O	0	0
			5	1	2	2		
8	A	1	Total	C	H	O	0	0
			5	1	2	2		
8	A	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			4	1	1	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			4	1	1	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		
8	C	1	Total	C	H	O	0	0
			5	1	2	2		

- Molecule 9 is 1,5-anhydro-2-deoxy-3-O-(2-deoxy- $\alpha$ -D-arabino-hexopyranosyl)-D-arabino-hex-1-enitol (three-letter code: XD3) (formula: C<sub>12</sub>H<sub>20</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	A	1	Total	C	O	0	0
			20	12	8		
9	C	1	Total	C	O	0	0
			20	12	8		

- Molecule 10 is D-glucal (three-letter code: DGO) (formula:  $C_6H_{10}O_4$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	B	1	Total	C	O	0	0
			10	6	4		
10	D	1	Total	C	O	0	0
			10	6	4		

- Molecule 11 is CALCIUM ION (three-letter code: CA) (formula: Ca).

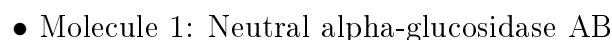
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	B	2	Total 2	Ca 2	0	0
11	D	2	Total 2	Ca 2	0	0

- Molecule 12 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	299	Total 299	O 299	0	0
12	B	35	Total 35	O 35	0	0
12	C	288	Total 288	O 288	0	0
12	D	26	Total 26	O 26	0	0



- Molecule 1: Neutral alpha-glucosidase AB





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	127.05Å 176.44Å 103.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	103.10 – 2.29 103.10 – 2.29	Depositor EDS
% Data completeness (in resolution range)	96.6 (103.10-2.29) 96.6 (103.10-2.29)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.21	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.63 (at 2.29Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.204 , 0.236 0.218 , 0.213	Depositor DCC
$R_{free}$ test set	4999 reflections (5.18%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.3	Xtriage
Anisotropy	0.664	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 43.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	30426	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.52% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: BMA, NAG, FMT, EDO, DGO, XD3, ACT, CA, 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	0.47	0/7118	0.69	3/9686 (0.0%)
1	C	0.51	0/7126	0.70	3/9697 (0.0%)
2	B	0.52	0/628	0.69	0/854
2	D	0.59	0/628	0.74	0/854
All	All	0.50	0/15500	0.70	6/21091 (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	C	0	1

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	885	ARG	CG-CD-NE	10.21	133.23	111.80
1	A	513	LYS	CD-CE-NZ	-6.41	96.96	111.70
1	C	497	SER	CA-C-N	5.54	127.27	116.20
1	A	467	ARG	CB-CG-CD	5.43	125.73	111.60
1	C	467	ARG	CB-CG-CD	5.19	125.09	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	497	SER	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	6902	6645	6651	6	0
1	C	6907	6647	6655	8	0
2	B	616	533	533	1	0
2	D	616	533	533	2	0
3	A	28	24	24	0	0
3	C	28	24	24	0	0
4	A	11	9	9	0	0
4	C	11	9	9	0	0
5	A	11	10	10	0	0
5	C	11	10	10	0	0
6	A	4	3	3	0	0
6	C	4	3	3	0	0
7	A	4	6	6	0	0
7	C	4	6	6	0	0
8	A	30	17	19	0	0
8	C	30	18	18	0	0
9	A	20	0	0	0	0
9	C	20	0	0	0	0
10	B	10	0	0	0	0
10	D	10	0	0	1	0
11	B	2	0	0	0	0
11	D	2	0	0	0	0
12	A	299	0	0	0	0
12	B	35	0	0	0	0
12	C	288	0	0	0	0
12	D	26	0	0	0	0
All	All	15929	14497	14513	17	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 1.

The worst 5 of 17 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:182:ARG:NH1	1:A:263:PRO:HG2	2.27	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:682:LEU:HD23	1:A:711:LEU:HD11	1.93	0.49
1:C:182:ARG:NH1	1:C:263:PRO:HG2	2.27	0.49
1:C:493:ILE:HG23	1:C:499:TYR:CE2	2.49	0.48
1:A:320:SER:O	1:A:627:PHE:HA	2.13	0.47

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	850/857 (99%)	827 (97%)	23 (3%)	0	100	100
1	C	851/857 (99%)	824 (97%)	26 (3%)	1 (0%)	56	68
2	B	81/83 (98%)	80 (99%)	1 (1%)	0	100	100
2	D	81/83 (98%)	79 (98%)	2 (2%)	0	100	100
All	All	1863/1880 (99%)	1810 (97%)	52 (3%)	1 (0%)	56	68

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	576	VAL

### 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	747/748 (100%)	740 (99%)	7 (1%)	84	93
1	C	748/748 (100%)	739 (99%)	9 (1%)	78	89
2	B	72/72 (100%)	72 (100%)	0	100	100
2	D	72/72 (100%)	71 (99%)	1 (1%)	74	86
All	All	1639/1640 (100%)	1622 (99%)	17 (1%)	82	91

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

Mol	Chain	Res	Type
1	C	424	ASN
1	C	446	ASP
1	C	706	ARG
1	A	859	GLN
1	C	714	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	904	HIS
1	C	714	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 40 ligands modelled in this entry, 4 are monoatomic - leaving 36 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	NAG	A	1001	1,3	14,14,15	0.26	0	15,19,21	0.60	0
3	NAG	A	1002	3,4	14,14,15	0.37	0	15,19,21	0.45	0
4	BMA	A	1003	3,5	11,11,12	0.26	0	15,15,17	0.55	0
5	MAN	A	1004	4	11,11,12	0.31	0	15,15,17	0.70	0
6	ACT	A	1005	-	0,3,3	0.00	-	0,3,3	0.00	-
7	EDO	A	1006	-	3,3,3	0.82	0	2,2,2	0.30	0
8	FMT	A	1007	-	0,2,2	0.00	-	0,1,1	0.00	-
8	FMT	A	1008	-	2,2,2	1.58	1 (50%)	1,1,1	1.65	0
8	FMT	A	1009	-	2,2,2	1.43	1 (50%)	1,1,1	1.55	0
8	FMT	A	1010	-	2,2,2	1.53	1 (50%)	1,1,1	1.46	0
8	FMT	A	1011	-	0,2,2	0.00	-	0,1,1	0.00	-
8	FMT	A	1012	-	0,2,2	0.00	-	0,1,1	0.00	-
8	FMT	A	1013	-	2,2,2	1.56	1 (50%)	1,1,1	1.10	0
8	FMT	A	1014	-	2,2,2	1.29	0	1,1,1	1.63	0
8	FMT	A	1015	-	2,2,2	1.77	1 (50%)	1,1,1	1.41	0
8	FMT	A	1016	-	2,2,2	1.52	1 (50%)	1,1,1	1.61	0
9	XD3	A	1017	-	20,21,21	0.22	0	23,29,29	1.87	1 (4%)
10	DGO	B	201	-	10,10,10	0.34	0	11,13,13	1.38	1 (9%)
9	XD3	C	1001	-	20,21,21	0.18	0	23,29,29	1.87	1 (4%)
3	NAG	C	1002	1,3	14,14,15	0.26	0	15,19,21	0.64	0
3	NAG	C	1003	3,4	14,14,15	0.37	0	15,19,21	0.47	0
4	BMA	C	1004	3,5	11,11,12	0.27	0	15,15,17	0.57	0
5	MAN	C	1005	4	11,11,12	0.31	0	15,15,17	0.73	0
6	ACT	C	1006	-	0,3,3	0.00	-	0,3,3	0.00	-
7	EDO	C	1007	-	3,3,3	0.83	0	2,2,2	0.31	0
8	FMT	C	1008	-	0,2,2	0.00	-	0,1,1	0.00	-
8	FMT	C	1009	-	2,2,2	1.54	1 (50%)	1,1,1	1.61	0
8	FMT	C	1010	-	2,2,2	1.25	0	1,1,1	1.48	0
8	FMT	C	1011	-	2,2,2	1.48	1 (50%)	1,1,1	1.52	0
8	FMT	C	1012	-	2,2,2	1.30	0	1,1,1	1.95	0
8	FMT	C	1013	-	0,2,2	0.00	-	0,1,1	0.00	-
8	FMT	C	1014	-	2,2,2	1.58	1 (50%)	1,1,1	1.13	0
8	FMT	C	1015	-	2,2,2	1.12	0	1,1,1	1.70	0
8	FMT	C	1016	-	2,2,2	1.32	0	1,1,1	1.59	0
8	FMT	C	1017	-	2,2,2	1.40	0	1,1,1	1.48	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
10	DGO	D	201	-	10,10,10	0.30	0	11,13,13	1.50	1 (9%)

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	1001	1,3	-	0/6/23/26	0/1/1/1
3	NAG	A	1002	3,4	-	0/6/23/26	0/1/1/1
4	BMA	A	1003	3,5	-	0/2/19/22	0/1/1/1
5	MAN	A	1004	4	-	0/2/19/22	0/1/1/1
6	ACT	A	1005	-	-	0/0/0/0	0/0/0/0
7	EDO	A	1006	-	-	0/1/1/1	0/0/0/0
8	FMT	A	1007	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1008	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1009	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1010	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1011	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1012	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1013	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1014	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1015	-	-	0/0/0/0	0/0/0/0
8	FMT	A	1016	-	-	0/0/0/0	0/0/0/0
9	XD3	A	1017	-	-	0/8/38/38	0/2/2/2
10	DGO	B	201	-	-	0/2/16/16	0/1/1/1
9	XD3	C	1001	-	-	0/8/38/38	0/2/2/2
3	NAG	C	1002	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	1003	3,4	-	0/6/23/26	0/1/1/1
4	BMA	C	1004	3,5	-	0/2/19/22	0/1/1/1
5	MAN	C	1005	4	-	0/2/19/22	0/1/1/1
6	ACT	C	1006	-	-	0/0/0/0	0/0/0/0
7	EDO	C	1007	-	-	0/1/1/1	0/0/0/0
8	FMT	C	1008	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1009	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1010	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1011	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1012	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1013	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1014	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1015	-	-	0/0/0/0	0/0/0/0
8	FMT	C	1016	-	-	0/0/0/0	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
8	FMT	C	1017	-	-	0/0/0/0	0/0/0/0
10	DGO	D	201	-	-	0/2/16/16	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	A	1009	FMT	O2-C	2.00	1.39	1.28
8	C	1011	FMT	O2-C	2.01	1.39	1.28
8	A	1013	FMT	O2-C	2.06	1.39	1.28
8	A	1010	FMT	O2-C	2.06	1.39	1.28
8	A	1016	FMT	O2-C	2.09	1.39	1.28

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	A	1017	XD3	O4-C7-C8	-8.68	97.83	107.00
9	C	1001	XD3	O4-C7-C8	-8.68	97.84	107.00
10	B	201	DGO	O-C2-C5	3.53	110.72	107.00
10	D	201	DGO	O-C2-C5	4.14	111.37	107.00

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
10	D	201	DGO	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.



## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	854/857 (99%)	0.11	11 (1%) 79 84	28, 43, 73, 104	0
1	C	854/857 (99%)	0.06	5 (0%) 90 93	29, 42, 67, 96	0
2	B	83/83 (100%)	0.50	7 (8%) 14 19	31, 49, 89, 104	0
2	D	83/83 (100%)	0.41	6 (7%) 18 26	31, 47, 93, 106	0
All	All	1874/1880 (99%)	0.12	29 (1%) 76 81	28, 42, 73, 106	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	81	GLY	9.1
2	D	82	TYR	5.0
1	C	814	TYR	4.3
2	B	82	TYR	4.3
2	B	43	THR	4.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
8	FMT	C	1012	3/3	0.91	0.27	6.88	56,57,59,59	0
8	FMT	C	1016	3/3	0.90	0.23	3.51	69,69,71,73	0
8	FMT	A	1009	3/3	0.90	0.18	2.61	63,63,65,65	0
8	FMT	C	1011	3/3	0.84	0.25	2.55	59,60,62,64	0
7	EDO	C	1007	4/4	0.91	0.19	2.21	39,45,47,47	0
8	FMT	A	1012	3/3	0.93	0.22	1.49	50,53,54,57	0
8	FMT	A	1010	3/3	0.81	0.20	1.48	60,60,63,65	0
8	FMT	A	1015	3/3	0.87	0.18	1.25	48,53,55,56	0
8	FMT	A	1007	3/3	0.88	0.16	0.86	55,60,60,63	0
7	EDO	A	1006	4/4	0.93	0.15	0.84	35,43,45,46	0
8	FMT	C	1010	3/3	0.86	0.15	0.77	54,54,57,59	0
6	ACT	A	1005	4/4	0.81	0.16	0.61	42,42,50,51	0
11	CA	B	203	1/1	0.99	0.14	0.27	36,36,36,36	0
8	FMT	C	1014	3/3	0.87	0.15	0.24	55,56,63,65	0
8	FMT	A	1011	3/3	0.96	0.16	0.23	49,51,52,55	0
10	DGO	D	201	10/10	0.91	0.15	0.10	40,46,50,53	0
9	XD3	A	1017	20/20	0.97	0.13	-0.07	23,38,47,49	0
9	XD3	C	1001	20/20	0.96	0.14	-0.11	33,40,48,50	0
8	FMT	A	1016	3/3	0.91	0.14	-0.31	53,53,56,57	0
3	NAG	A	1001	14/15	0.95	0.14	-0.54	43,50,56,58	0
10	DGO	B	201	10/10	0.94	0.12	-0.71	36,43,46,48	0
6	ACT	C	1006	4/4	0.83	0.14	-0.77	38,40,56,56	0
8	FMT	C	1008	3/3	0.91	0.13	-0.84	52,56,56,61	0
3	NAG	C	1002	14/15	0.96	0.12	-0.99	37,44,48,49	0
8	FMT	C	1017	3/3	0.91	0.13	-1.01	68,68,69,69	0
8	FMT	A	1008	3/3	0.91	0.11	-1.23	49,51,52,53	0
11	CA	D	203	1/1	0.99	0.11	-1.29	36,36,36,36	0
8	FMT	C	1009	3/3	0.94	0.10	-1.37	51,57,61,61	0
11	CA	B	202	1/1	0.97	0.09	-1.47	39,39,39,39	0
8	FMT	A	1013	3/3	0.94	0.11	-1.50	40,41,44,45	0
11	CA	D	202	1/1	0.99	0.10	-1.52	37,37,37,37	0
8	FMT	C	1013	3/3	0.96	0.12	-1.93	48,54,54,55	0
5	MAN	A	1004	11/12	0.71	0.23	-	101,105,106,107	0
3	NAG	C	1003	14/15	0.93	0.11	-	52,56,65,73	0
4	BMA	A	1003	11/12	0.83	0.15	-	86,96,98,102	0
4	BMA	C	1004	11/12	0.81	0.16	-	80,87,89,93	0
5	MAN	C	1005	11/12	0.86	0.16	-	97,102,103,103	0
3	NAG	A	1002	14/15	0.92	0.12	-	63,64,72,77	0
8	FMT	C	1015	3/3	0.95	0.09	-	49,50,54,54	0
8	FMT	A	1014	3/3	0.92	0.22	-	62,66,70,70	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.