



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:39 PM GMT

PDB ID : 4CVU  
Title : Structure of Fungal beta-mannosidase from Glycoside Hydrolase Family 2 of *Trichoderma harzianum*  
Authors : Muniz, J.R.C.; Aparicio, R.; Santos, J.C.; Nascimento, A.S.; Golubev, A.M.; Polikarpov, I.  
Deposited on : 2014-03-31  
Resolution : 1.90 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<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

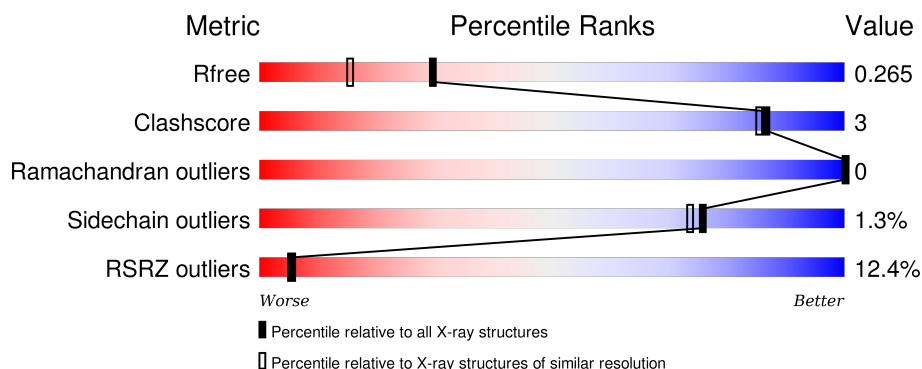
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	942	<div> <div>12%</div> <div>90%</div> <div>6%</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
3	NAG	A	1974	-	-	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	MAN	A	1950	-	-	-	X
4	MAN	A	1954	-	-	-	X
5	NAG	A	1956	-	-	-	X
5	MAN	A	1958	-	-	-	X
7	NAG	A	1977	X	-	-	-

## 2 Entry composition [i](#)

There are 14 unique types of molecules in this entry. The entry contains 8456 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-MANNOSIDASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	913	Total	C	N	O	S	0	6	0
			7241	4645	1208	1371	17			

- Molecule 2 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	3	Total	C	N	O	0	0
			39	22	2	15		

- Molecule 3 is SUGAR (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	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is a polymer of unknown type called SUGAR (8-MER).

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

- Molecule 5 is a polymer of unknown type called SUGAR (6-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	6	Total	C	N	O	0	0
			72	40	2	30		

- Molecule 6 is a polymer of unknown type called SUGAR (11-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	11	Total	C	N	O	0	0
			127	70	2	55		

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

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

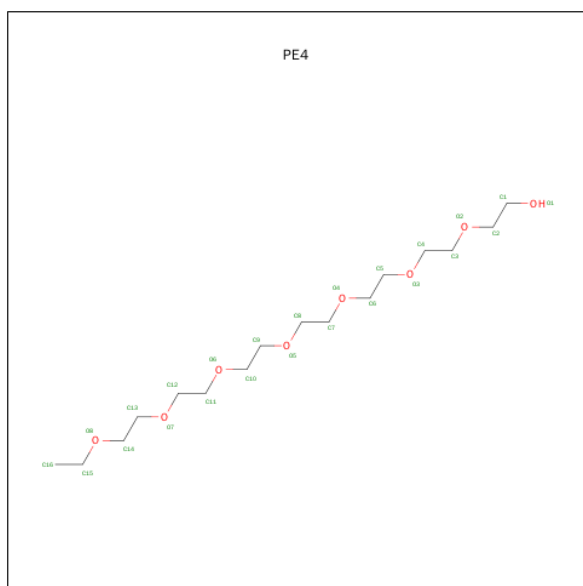
- Molecule 8 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	17	Total	Cd	0	0
			17	17		

- Molecule 9 is SODIUM ION (three-letter code: NA) (formula: Na).

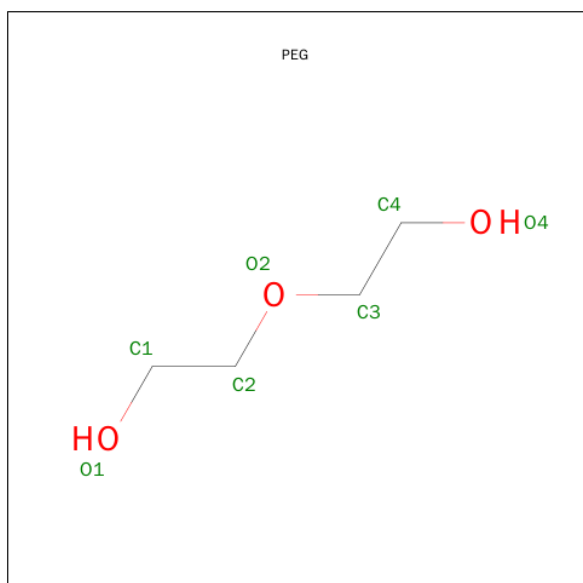
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	6	Total	Na	0	0
			6	6		

- Molecule 10 is 2-{2-[2-(2-{2-[2-(2-ETHOXY-ETHOXY)-ETHOXY]-ETHOXY}-ETHOXY)-ETHOXY]-ETHOXY}-ETHANOL (three-letter code: PE4) (formula: C<sub>16</sub>H<sub>34</sub>O<sub>8</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	A	1	Total	C	O	0	0
			12	8	4		

- Molecule 11 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	A	1	Total	C	O	0	0
			7	4	3		

- Molecule 12 is CALCIUM ION (three-letter code: CA) (formula: Ca).

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

- Molecule 13 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	1	Total	Cl	0	0
			1	1		

- Molecule 14 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	741	Total	O	0	0
			741	741		





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.46 Å   166.46 Å   121.45 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	117.70 – 1.90 48.30 – 1.90	Depositor EDS
% Data completeness (in resolution range)	98.8 (117.70-1.90) 98.8 (48.30-1.90)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.21 (at 1.90 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.234   ,   0.263 0.239   ,   0.265	Depositor DCC
$R_{free}$ test set	6641 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	25.4	Xtriage
Anisotropy	0.511	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 52.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 131627 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	8456	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.01% 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: BMA, NAG, CL, PE4, NA, CA, CD, PEG, 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	0.67	0/7479	0.72	4/10243 (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
7	A	1	0
All	All	1	1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	723	ASP	CB-CG-OD1	6.79	124.41	118.30
1	A	409	ASP	CB-CG-OD1	6.43	124.09	118.30
1	A	723	ASP	CB-CG-OD2	-6.22	112.70	118.30
1	A	188	ARG	NE-CZ-NH2	5.04	122.82	120.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	A	1977	NAG	C3

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	424	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	7241	0	6832	39	0
2	A	39	0	34	1	0
3	A	70	0	65	0	0
4	A	94	0	79	1	0
5	A	72	0	61	0	0
6	A	127	0	105	0	0
7	A	28	0	23	0	0
8	A	17	0	0	0	0
9	A	6	0	0	0	0
10	A	12	0	14	0	0
11	A	7	0	10	0	0
12	A	1	0	0	0	0
13	A	1	0	0	1	0
14	A	741	0	0	7	0
All	All	8456	0	7223	40	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 3.

All (40) 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)
4:A:1948:NAG:H81	4:A:1953:MAN:O6	1.81	0.80
1:A:194:GLU:HG3	1:A:425:GLY:HA2	1.67	0.75
1:A:597:PHE:HB2	1:A:682:MET:CE	2.26	0.66
1:A:565:HIS:NE2	13:A:2003:CL:CL	2.58	0.65
1:A:597:PHE:HB2	1:A:682:MET:HE1	1.80	0.64
1:A:844:THR:O	14:A:3586:HOH:O	2.16	0.62
1:A:879:ALA:O	1:A:908:PRO:HA	2.03	0.58
1:A:656:TYR:CE1	1:A:854:PRO:HG2	2.41	0.55
1:A:444:TRP:CD1	1:A:444:TRP:C	2.80	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:860:SER:OG	14:A:3647:HOH:O	2.18	0.53
1:A:272:VAL:HG12	1:A:328:VAL:HG22	1.91	0.52
1:A:617:PHE:CD1	2:A:1943:NAG:H82	2.45	0.51
1:A:243:ASN:HD21	1:A:534:TYR:H	1.57	0.51
1:A:107:LEU:O	1:A:136:PHE:HA	2.11	0.51
1:A:610[A]:VAL:CG1	1:A:614:GLU:HB2	2.42	0.50
1:A:572:PHE:CD1	1:A:587:PRO:HD3	2.46	0.49
1:A:597:PHE:HB2	1:A:682:MET:HE2	1.96	0.48
1:A:194:GLU:HG3	1:A:425:GLY:CA	2.42	0.47
1:A:594:GLU:CD	14:A:3402:HOH:O	2.53	0.47
1:A:335:SER:CB	14:A:3290:HOH:O	2.63	0.47
1:A:185:TYR:HB2	1:A:198:PHE:HB3	1.98	0.46
1:A:747:ASN:ND2	14:A:3595:HOH:O	2.39	0.46
1:A:242:ASN:O	1:A:243:ASN:HB2	2.16	0.45
1:A:430:ASP:OD1	1:A:478:HIS:CE1	2.69	0.45
1:A:427:TYR:CE2	1:A:448:GLN:HG2	2.51	0.45
1:A:713:GLN:OE1	1:A:713:GLN:N	2.46	0.45
1:A:762:TRP:CE3	1:A:833:PRO:HB3	2.53	0.44
1:A:125[A]:VAL:HG12	1:A:137:TYR:HB3	1.99	0.44
1:A:860:SER:CB	14:A:3647:HOH:O	2.65	0.43
1:A:273:ILE:HG22	1:A:275[B]:THR:HG23	2.01	0.42
1:A:368:THR:HA	1:A:369:PRO:HD3	1.87	0.42
1:A:33:LEU:HG	1:A:216:ARG:HA	2.01	0.42
1:A:559:TYR:N	1:A:559:TYR:CD1	2.86	0.42
1:A:252:TRP:CD1	1:A:310:PRO:HG2	2.55	0.42
1:A:444:TRP:C	1:A:444:TRP:HD1	2.22	0.41
1:A:370:GLY:HA3	1:A:566:ILE:O	2.20	0.41
1:A:610[A]:VAL:HG13	1:A:614:GLU:HB2	2.02	0.41
1:A:386:GLY:HA2	1:A:416:PHE:CD2	2.55	0.41
1:A:68:GLY:HA3	14:A:3051:HOH:O	2.20	0.41
1:A:243:ASN:HD21	1:A:534:TYR:N	2.19	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	915/942 (97%)	891 (97%)	24 (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	777/811 (96%)	766 (99%)	11 (1%)	74	71

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	70	SER
1	A	122	ASP
1	A	216	ARG
1	A	349[A]	ARG
1	A	349[B]	ARG
1	A	444	TRP
1	A	499	GLN
1	A	559	TYR
1	A	699	ARG
1	A	712	TRP
1	A	907	ARG

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	123	HIS
1	A	243	ASN

### 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 ⓘ

30 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	1943	1,2	14,14,15	0.64	0	15,19,21	1.16	1 (6%)
2	NAG	A	1944	2	14,14,15	0.62	0	15,19,21	0.94	1 (6%)
2	BMA	A	1945	2	11,11,12	0.50	0	14,15,17	0.76	0
4	NAG	A	1947	1,4	14,14,15	0.79	1 (7%)	15,19,21	0.94	0
4	NAG	A	1948	4	14,14,15	0.65	0	15,19,21	1.37	2 (13%)
4	BMA	A	1949	4	11,11,12	0.76	0	14,15,17	1.17	2 (14%)
4	MAN	A	1950	4	11,11,12	0.58	0	14,15,17	0.87	0
4	MAN	A	1951	4	11,11,12	0.51	0	14,15,17	1.07	1 (7%)
4	MAN	A	1952	4	11,11,12	0.59	0	14,15,17	1.97	1 (7%)
4	MAN	A	1953	4	11,11,12	0.74	0	14,15,17	1.59	2 (14%)
4	MAN	A	1954	4	11,11,12	0.66	0	14,15,17	0.91	1 (7%)
5	NAG	A	1955	1,5	14,14,15	0.50	0	15,19,21	0.88	1 (6%)
5	NAG	A	1956	5	14,14,15	0.72	0	15,19,21	1.30	1 (6%)
5	BMA	A	1957	5	11,11,12	0.58	0	14,15,17	1.15	1 (7%)
5	MAN	A	1958	5	11,11,12	0.60	0	14,15,17	1.25	2 (14%)
5	MAN	A	1959	5	11,11,12	0.59	0	14,15,17	0.97	0
5	MAN	A	1960	5	11,11,12	0.97	0	15,15,17	5.63	3 (20%)
6	NAG	A	1961	1,6	14,14,15	0.64	0	15,19,21	1.01	1 (6%)
6	NAG	A	1962	6	14,14,15	0.75	0	15,19,21	0.97	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	BMA	A	1963	6	11,11,12	0.52	0	14,15,17	0.93	1 (7%)
6	MAN	A	1964	6	11,11,12	0.59	0	14,15,17	1.19	3 (21%)
6	MAN	A	1965	6	11,11,12	1.00	0	14,15,17	1.16	1 (7%)
6	MAN	A	1966	6	11,11,12	0.60	0	14,15,17	1.13	2 (14%)
6	MAN	A	1967	6	11,11,12	1.37	1 (9%)	14,15,17	3.50	3 (21%)
6	MAN	A	1968	6	11,11,12	0.73	1 (9%)	14,15,17	0.82	0
6	MAN	A	1969	6	11,11,12	0.62	0	14,15,17	0.76	1 (7%)
6	MAN	A	1970	6	11,11,12	0.72	0	14,15,17	1.57	4 (28%)
6	MAN	A	1971	6	11,11,12	0.69	0	14,15,17	1.38	1 (7%)
7	NAG	A	1976	1,7	14,14,15	0.93	1 (7%)	15,19,21	1.16	1 (6%)
7	NAG	A	1977	7	14,14,15	25.93	1 (7%)	15,19,21	11.78	8 (53%)

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	1943	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	1944	2	-	0/6/23/26	0/1/1/1
2	BMA	A	1945	2	-	0/2/19/22	0/1/1/1
4	NAG	A	1947	1,4	-	0/6/23/26	0/1/1/1
4	NAG	A	1948	4	-	0/6/23/26	0/1/1/1
4	BMA	A	1949	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1950	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1951	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1952	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1953	4	-	0/2/19/22	0/1/1/1
4	MAN	A	1954	4	-	0/2/19/22	0/1/1/1
5	NAG	A	1955	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	1956	5	-	0/6/23/26	0/1/1/1
5	BMA	A	1957	5	-	0/2/19/22	0/1/1/1
5	MAN	A	1958	5	-	0/2/19/22	0/1/1/1
5	MAN	A	1959	5	-	0/2/19/22	0/1/1/1
5	MAN	A	1960	5	-	0/2/18/22	0/1/1/1
6	NAG	A	1961	1,6	-	0/6/23/26	0/1/1/1
6	NAG	A	1962	6	-	0/6/23/26	0/1/1/1
6	BMA	A	1963	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1964	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1965	6	-	0/2/19/22	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	MAN	A	1966	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1967	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1968	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1969	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1970	6	-	0/2/19/22	0/1/1/1
6	MAN	A	1971	6	-	0/2/19/22	0/1/1/1
7	NAG	A	1976	1,7	-	0/6/23/26	0/1/1/1
7	NAG	A	1977	7	1/1/5/7	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	1967	MAN	O3-C3	-3.95	1.33	1.43
7	A	1976	NAG	O5-C1	-2.06	1.40	1.43
6	A	1968	MAN	O5-C5	-2.01	1.39	1.43
4	A	1947	NAG	C1-C2	2.15	1.55	1.52
7	A	1977	NAG	O3-C3	97.00	3.75	1.43

All (46) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	A	1977	NAG	O3-C3-C2	-44.35	21.26	109.11
7	A	1977	NAG	O3-C3-C4	-6.80	95.03	110.34
5	A	1960	MAN	C1-C2-C3	-6.23	100.55	111.23
5	A	1958	MAN	O2-C2-C3	-2.80	104.49	110.12
6	A	1966	MAN	C2-C3-C4	-2.79	106.30	111.04
4	A	1949	BMA	O3-C3-C4	-2.66	104.36	110.34
6	A	1970	MAN	C1-C2-C3	-2.62	106.44	109.54
5	A	1957	BMA	O4-C4-C3	-2.60	104.49	110.34
5	A	1956	NAG	C3-C4-C5	-2.54	105.78	110.20
7	A	1977	NAG	O7-C7-C8	-2.47	117.53	122.06
4	A	1953	MAN	O2-C2-C1	-2.39	104.41	109.21
6	A	1962	NAG	O7-C7-C8	-2.25	117.94	122.06
6	A	1970	MAN	O4-C4-C3	-2.18	105.42	110.34
6	A	1963	BMA	O6-C6-C5	-2.14	104.25	111.33
7	A	1977	NAG	O7-C7-N2	-2.07	117.64	121.86
6	A	1964	MAN	O2-C2-C1	-2.07	105.06	109.21
4	A	1949	BMA	O4-C4-C3	-2.06	105.70	110.34
6	A	1964	MAN	C2-C3-C4	-2.04	107.58	111.04
5	A	1958	MAN	C1-C2-C3	-2.01	107.16	109.54
6	A	1967	MAN	O3-C3-C4	-2.00	105.82	110.34
6	A	1966	MAN	C1-O5-C5	2.03	114.83	112.25

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	1964	MAN	O4-C4-C5	2.06	114.70	109.24
6	A	1967	MAN	C1-O5-C5	2.15	114.98	112.25
7	A	1977	NAG	C3-C2-N2	2.17	115.75	110.56
2	A	1944	NAG	C1-O5-C5	2.22	115.07	112.25
4	A	1948	NAG	C8-C7-N2	2.23	120.38	116.11
6	A	1969	MAN	C1-O5-C5	2.34	115.22	112.25
5	A	1955	NAG	C1-O5-C5	2.40	115.30	112.25
7	A	1977	NAG	C3-C4-C5	2.41	114.39	110.20
6	A	1965	MAN	C1-O5-C5	2.50	115.42	112.25
2	A	1943	NAG	C1-O5-C5	2.61	115.56	112.25
6	A	1970	MAN	C3-C4-C5	2.74	114.97	110.20
6	A	1961	NAG	C1-O5-C5	2.75	115.73	112.25
6	A	1970	MAN	O5-C1-C2	2.86	115.49	110.86
4	A	1951	MAN	C1-O5-C5	3.02	116.08	112.25
4	A	1954	MAN	C1-O5-C5	3.07	116.15	112.25
6	A	1971	MAN	C1-O5-C5	3.07	116.15	112.25
4	A	1948	NAG	C2-N2-C7	3.41	127.42	123.04
7	A	1976	NAG	C1-O5-C5	3.41	116.58	112.25
4	A	1953	MAN	C1-C2-C3	3.68	113.89	109.54
7	A	1977	NAG	C2-N2-C7	4.44	128.75	123.04
7	A	1977	NAG	C8-C7-N2	4.56	124.83	116.11
5	A	1960	MAN	C3-C4-C5	5.25	115.22	109.93
4	A	1952	MAN	C1-O5-C5	6.22	120.14	112.25
6	A	1967	MAN	O3-C3-C2	12.37	132.35	110.00
5	A	1960	MAN	O5-C1-C2	20.03	126.31	110.72

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
7	A	1977	NAG	C3

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1943	NAG	1	0
4	A	1948	NAG	1	0
4	A	1953	MAN	1	0

## 5.6 Ligand geometry

Of 32 ligands modelled in this entry, 25 are monoatomic - leaving 7 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	1946	1	14,14,15	0.90	0	15,19,21	1.46	1 (6%)
3	NAG	A	1972	1	14,14,15	0.46	0	15,19,21	1.26	1 (6%)
3	NAG	A	1973	1	14,14,15	0.43	0	15,19,21	1.33	1 (6%)
3	NAG	A	1974	1	14,14,15	0.52	0	15,19,21	1.05	1 (6%)
3	NAG	A	1975	1	14,14,15	0.51	0	15,19,21	1.05	1 (6%)
10	PE4	A	2000	-	11,11,23	0.55	0	10,10,22	0.43	0
11	PEG	A	2001	-	6,6,6	0.52	0	5,5,5	0.24	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
3	NAG	A	1946	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1972	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1973	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1974	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1975	1	-	0/6/23/26	0/1/1/1
10	PE4	A	2000	-	-	0/9/9/21	0/0/0/0
11	PEG	A	2001	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1975	NAG	O7-C7-C8	-2.03	118.34	122.06
3	A	1974	NAG	C1-O5-C5	2.54	115.47	112.25
3	A	1973	NAG	C1-O5-C5	2.95	116.00	112.25
3	A	1972	NAG	C1-O5-C5	3.71	116.95	112.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1946	NAG	C1-O5-C5	4.93	118.51	112.25

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	913/942 (96%)	0.79	113 (12%) 5 5	28, 45, 69, 105	1 (0%)

All (113) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	749	THR	7.3
1	A	839	LYS	6.5
1	A	750	ALA	6.4
1	A	334	GLY	6.1
1	A	838	SER	5.7
1	A	389	LEU	5.5
1	A	563	SER	5.4
1	A	751	ASN	5.4
1	A	419	LEU	5.2
1	A	267	HIS	4.8
1	A	421	VAL	4.7
1	A	432	ILE	4.6
1	A	443	LEU	4.6
1	A	428	LEU	4.5
1	A	810	SER	4.4
1	A	306	ASP	4.4
1	A	265	PRO	4.3
1	A	708	LEU	4.2
1	A	807	ASN	4.1
1	A	268	ALA	4.1
1	A	308	HIS	4.1
1	A	168	ALA	4.0
1	A	811	ASP	4.0
1	A	805	THR	3.9
1	A	142	ALA	3.9
1	A	96	GLY	3.8
1	A	333	ALA	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	103	ILE	3.7
1	A	840	THR	3.7
1	A	564	GLY	3.6
1	A	411	VAL	3.4
1	A	250	ALA	3.4
1	A	141	SER	3.4
1	A	276	ASP	3.3
1	A	392	PRO	3.3
1	A	396	TRP	3.3
1	A	390	ILE	3.3
1	A	37	LYS	3.2
1	A	336	LYS	3.2
1	A	748	TYR	3.2
1	A	481	LEU	3.2
1	A	249	THR	3.1
1	A	148	SER	3.1
1	A	722	TYR	3.1
1	A	720	LEU	3.1
1	A	752	GLU	3.1
1	A	427	TYR	3.0
1	A	146	CYS	3.0
1	A	93	PRO	3.0
1	A	391	PRO	2.9
1	A	282	SER	2.8
1	A	501	ALA	2.8
1	A	737	TYR	2.8
1	A	145	SER	2.8
1	A	794	ASN	2.8
1	A	363	ILE	2.8
1	A	837	SER	2.8
1	A	729	LEU	2.8
1	A	337	THR	2.8
1	A	30	ILE	2.7
1	A	504	ALA	2.7
1	A	102	ASP	2.6
1	A	407	LEU	2.6
1	A	307	ALA	2.6
1	A	287	GLY	2.6
1	A	562	THR	2.6
1	A	447	PHE	2.6
1	A	332	SER	2.6
1	A	484	TRP	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	841	THR	2.6
1	A	712	TRP	2.5
1	A	408	PHE	2.5
1	A	223	GLY	2.5
1	A	482	ALA	2.5
1	A	97	LEU	2.5
1	A	387	ALA	2.5
1	A	416	PHE	2.4
1	A	746	TRP	2.4
1	A	783	ASN	2.4
1	A	430	ASP	2.4
1	A	561	THR	2.4
1	A	559	TYR	2.4
1	A	809	GLN	2.4
1	A	836	ALA	2.4
1	A	433	TYR	2.4
1	A	846	HIS	2.4
1	A	172	VAL	2.4
1	A	705	TYR	2.3
1	A	99	LYS	2.3
1	A	709	ASN	2.3
1	A	368	THR	2.3
1	A	547	ASP	2.3
1	A	277	ALA	2.3
1	A	285	TYR	2.3
1	A	442	LEU	2.3
1	A	369	PRO	2.3
1	A	170	ASP	2.2
1	A	386	GLY	2.2
1	A	104	ALA	2.2
1	A	806	ALA	2.2
1	A	109	PHE	2.2
1	A	335	SER	2.2
1	A	444	TRP	2.1
1	A	789	LYS	2.1
1	A	264	LEU	2.1
1	A	247	ASP	2.1
1	A	136	PHE	2.1
1	A	394	ALA	2.1
1	A	660	ASP	2.1
1	A	100	HIS	2.1
1	A	711	ILE	2.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	728	VAL	2.0
1	A	320	ASN	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
4	MAN	A	1950	11/12	0.88	0.27	7.60	57,60,64,64	0
4	MAN	A	1954	11/12	0.85	0.26	3.22	52,60,62,64	0
5	MAN	A	1958	11/12	0.91	0.16	3.07	43,46,47,48	0
5	NAG	A	1956	14/15	0.87	0.15	2.44	44,48,50,58	0
4	NAG	A	1947	14/15	0.85	0.18	1.92	40,43,47,49	0
5	MAN	A	1959	11/12	0.89	0.17	1.77	43,47,53,56	0
2	NAG	A	1943	14/15	0.92	0.12	1.44	35,38,42,45	0
6	MAN	A	1966	11/12	0.83	0.21	1.44	52,57,61,65	0
5	NAG	A	1955	14/15	0.90	0.17	0.89	44,46,47,47	0
6	NAG	A	1962	14/15	0.91	0.12	-0.62	41,44,47,48	0
6	NAG	A	1961	14/15	0.91	0.12	-0.78	42,45,51,51	0
6	MAN	A	1968	11/12	0.94	0.12	-0.81	46,50,52,53	0
6	MAN	A	1969	11/12	0.91	0.10	-2.15	49,51,52,52	0
4	BMA	A	1949	11/12	0.88	0.25	-	50,55,57,59	0
7	NAG	A	1976	14/15	0.91	0.20	-	46,51,55,57	0
6	MAN	A	1964	11/12	0.90	0.11	-	43,44,47,48	0
4	MAN	A	1951	11/12	0.84	0.18	-	57,61,64,69	0
6	MAN	A	1971	11/12	0.89	0.17	-	51,53,54,56	0
6	MAN	A	1970	11/12	0.89	0.15	-	48,55,62,63	0
4	MAN	A	1952	11/12	0.65	0.43	-	63,65,67,68	0
4	NAG	A	1948	14/15	0.81	0.19	-	44,48,53,54	0
5	BMA	A	1957	11/12	0.88	0.21	-	48,51,55,58	0
6	BMA	A	1963	11/12	0.95	0.12	-	43,45,50,50	0
7	NAG	A	1977	14/15	0.85	0.28	-	67,73,78,80	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
6	MAN	A	1965	11/12	0.93	0.18	-	45,48,53,57	0
2	BMA	A	1945	11/12	0.81	0.34	-	54,57,59,62	0
2	NAG	A	1944	14/15	0.92	0.20	-	46,50,51,52	0
5	MAN	A	1960	11/12	0.71	0.24	-	65,66,68,70	0
6	MAN	A	1967	11/12	0.79	0.19	-	47,51,53,53	0
4	MAN	A	1953	11/12	0.86	0.27	-	51,60,62,64	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, 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
3	NAG	A	1974	14/15	0.84	0.30	8.30	54,59,68,70	0
10	PE4	A	2000	12/24	0.79	0.17	1.84	57,61,66,66	0
3	NAG	A	1973	14/15	0.79	0.30	1.76	56,60,63,68	0
3	NAG	A	1975	14/15	0.90	0.19	1.18	40,43,46,47	0
3	NAG	A	1946	14/15	0.90	0.14	0.24	39,44,48,48	0
8	CD	A	1990	1/1	0.96	0.10	-2.10	77,77,77,77	0
8	CD	A	1981	1/1	0.98	0.07	-	50,50,50,50	0
8	CD	A	1992	1/1	0.96	0.17	-	87,87,87,87	1
8	CD	A	1979	1/1	0.98	0.08	-	43,43,43,43	0
9	NA	A	1998	1/1	0.89	0.17	-	40,40,40,40	0
13	CL	A	2003	1/1	0.89	0.15	-	49,49,49,49	0
9	NA	A	1995	1/1	0.70	0.30	-	54,54,54,54	0
9	NA	A	1997	1/1	0.93	0.28	-	30,30,30,30	0
8	CD	A	1983	1/1	0.96	0.08	-	52,52,52,52	0
9	NA	A	1999	1/1	0.95	0.19	-	33,33,33,33	0
8	CD	A	1987	1/1	0.98	0.08	-	81,81,81,81	0
8	CD	A	1978	1/1	0.99	0.10	-	39,39,39,39	0
8	CD	A	1993	1/1	0.91	0.39	-	44,44,44,44	0
8	CD	A	1988	1/1	0.97	0.07	-	74,74,74,74	0
9	NA	A	2004	1/1	0.97	0.33	-	37,37,37,37	0
8	CD	A	1985	1/1	0.98	0.09	-	51,51,51,51	0
8	CD	A	1989	1/1	0.93	0.05	-	80,80,80,80	0
8	CD	A	1984	1/1	0.94	0.08	-	76,76,76,76	0
8	CD	A	1982	1/1	0.91	0.06	-	72,72,72,72	0
9	NA	A	1996	1/1	0.95	0.28	-	48,48,48,48	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
8	CD	A	1986	1/1	0.98	0.05	-	59,59,59,59	0
3	NAG	A	1972	14/15	0.87	0.14	-	54,59,64,67	0
8	CD	A	1980	1/1	0.98	0.10	-	58,58,58,58	0
11	PEG	A	2001	7/7	0.78	0.19	-	58,60,62,64	0
8	CD	A	1994	1/1	0.92	0.61	-	44,44,44,44	0
12	CA	A	2002	1/1	0.95	0.08	-	54,54,54,54	0
8	CD	A	1991	1/1	0.74	0.20	-	125,125,125,125	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.