



# Full wwPDB X-ray Structure Validation Report i

Feb 27, 2014 – 04:36 PM GMT

PDB ID : 4AX7  
Title : Hypocrea jecorina Cel6A D221A mutant soaked with 4-Methylumbelliferyl  
-beta-D-cellobioside  
Authors : Wu, M.; Nerinckx, W.; Piens, K.; Ishida, T.; Hansson, H.; Stahlberg, J.;  
Sandgren, M.  
Deposited on : 2012-06-11  
Resolution : 1.70 Å(reported)

This is a full wwPDB 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/ValidationPDFNotes.html>

---

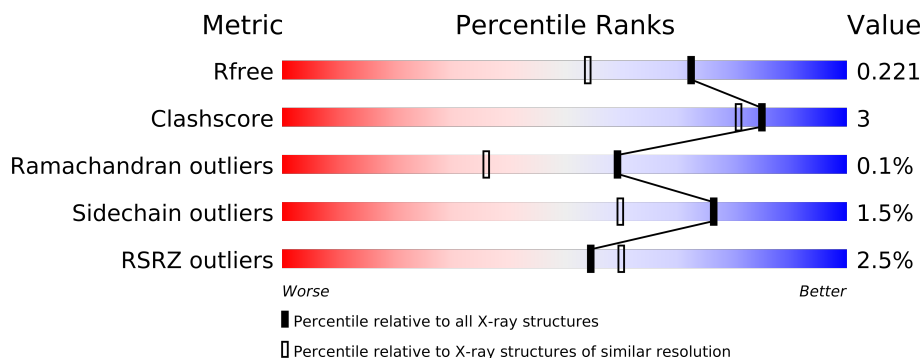
The following versions of software and data (see [references](#)) were used in the production of this report:

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 1.70 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	2456 (1.70-1.70)
Clashscore	79885	2929 (1.70-1.70)
Ramachandran outliers	78287	2878 (1.70-1.70)
Sidechain outliers	78261	2878 (1.70-1.70)
RSRZ outliers	66119	2456 (1.70-1.70)

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. 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	363	
1	B	363	
1	C	363	
1	D	363	

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

Mol	Type	Chain	Res	Geometry	Electron density
2	NAG	B	501	-	X
2	NAG	C	501	-	X
2	NAG	D	501	-	X
3	MAN	A	503	-	X

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Geometry	Electron density
3	MAN	A	504	-	X
3	MAN	A	505	-	X
3	MAN	A	508	-	X
3	MAN	A	509	-	X
3	MAN	B	503	-	X
3	MAN	B	504	-	X
3	MAN	B	505	-	X
3	MAN	B	507	-	X
3	MAN	B	509	-	X
3	MAN	C	504	-	X
3	MAN	C	505	-	X
3	MAN	D	505	-	X
3	MAN	D	509	-	X
4	4MU	B	600	-	X
4	4MU	C	600	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12750 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 EXOGLUCANASE 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	363	Total	C	N	O	S	0	2	0
			2758	1750	466	532	10			
1	B	360	Total	C	N	O	S	0	3	0
			2740	1740	462	528	10			
1	C	363	Total	C	N	O	S	0	3	0
			2764	1755	466	533	10			
1	D	360	Total	C	N	O	S	0	3	0
			2737	1738	461	528	10			

There are 4 discrepancies between the modelled and reference sequences:

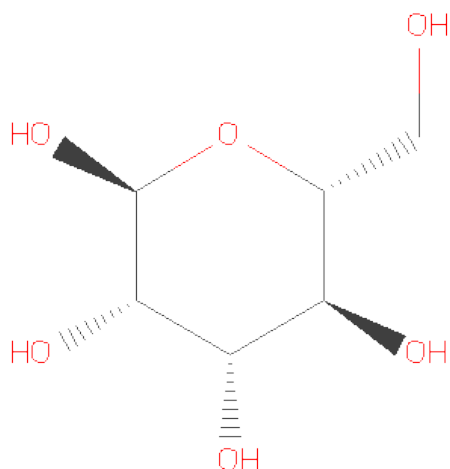
Chain	Residue	Modelled	Actual	Comment	Reference
A	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987
B	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987
C	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987
D	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

- Molecule 3 is SUGAR (ALPHA-D-MANNOSE) (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>).



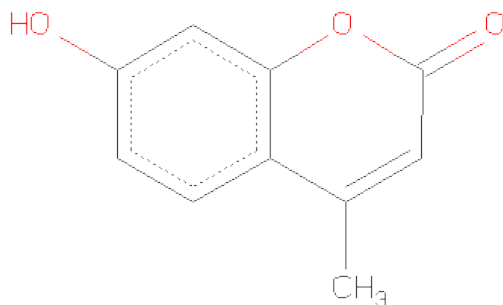
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	A	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		
3	B	1	Total	C	O	0	0
			11	6	5		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	C	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		
3	D	1	Total	C	O	0	0
			11	6	5		

- Molecule 4 is 7-HYDROXY-4-METHYL-2H-CHROMEN-2-ONE (three-letter code: 4MU) (formula: C<sub>10</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	10	3		
4	B	1	Total	C	O	0	0
			13	10	3		
4	C	1	Total	C	O	0	0
			13	10	3		
4	D	1	Total	C	O	0	0
			13	10	3		

- Molecule 5 is a polymer of unknown type called SUGAR (4-MER).

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

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987
C	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	2	Total	C	O	0	0
			22	12	10		
6	D	2	Total	C	O	0	0
			22	12	10		

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987
D	221	ALA	ASP	ENGINEERED MUTATION	UNP P07987

- Molecule 7 is water.

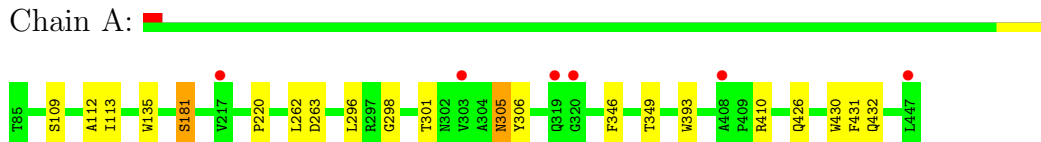
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	329	Total	O	0	0
			329	329		
7	B	287	Total	O	0	0
			287	287		
7	C	285	Total	O	0	0
			285	285		
7	D	257	Total	O	0	0
			257	257		

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

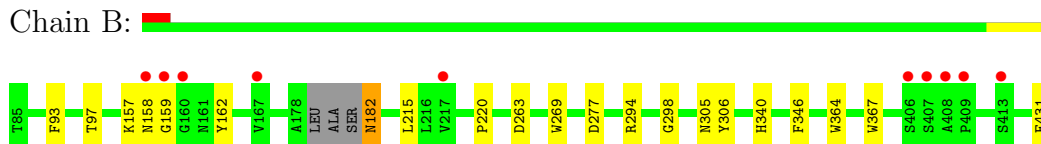
#### • Molecule 1: EXOGLUCANASE 2

Chain A:



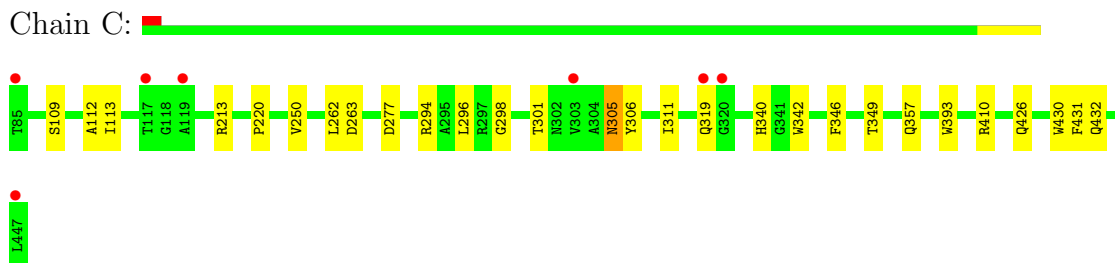
#### • Molecule 1: EXOGLUCANASE 2

Chain B:



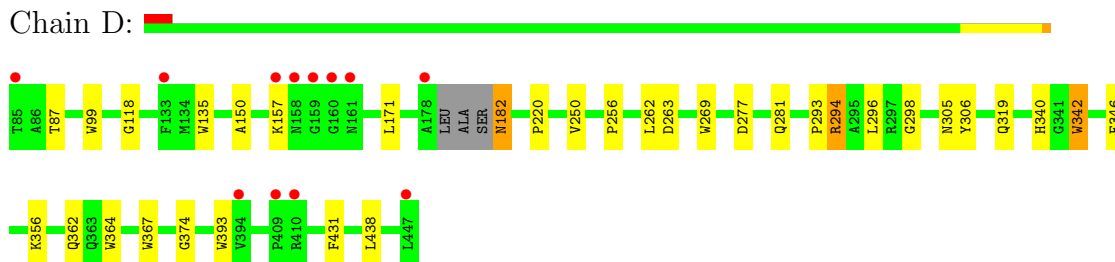
#### • Molecule 1: EXOGLUCANASE 2

Chain C:



#### • Molecule 1: EXOGLUCANASE 2

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	48.82Å 74.68Å 93.30Å 103.97° 103.43° 90.00°	Depositor
Resolution (Å)	87.90 – 1.70 29.86 – 1.70	Depositor EDS
% Data completeness (in resolution range)	96.4 (87.90-1.70) 96.4 (29.86-1.70)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.61 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.189 , 0.221 0.189 , 0.221	Depositor DCC
$R_{free}$ test set	6627 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	15.0	Xtriage
Anisotropy	0.252	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 32.4	EDS
Estimated twinning fraction	0.013 for h,-k,-h-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 131662 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	12750	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 53.07 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 4.4432e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities 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: 4MU, BGC, 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	0.53	1/2834 (0.0%)	0.55	0/3887
1	B	0.52	3/2821 (0.1%)	0.54	0/3866
1	C	0.52	2/2840 (0.1%)	0.54	0/3894
1	D	0.53	7/2818 (0.2%)	0.54	0/3862
All	All	0.52	13/11313 (0.1%)	0.55	0/15509

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	364	TRP	CD2-CE2	5.24	1.47	1.41
1	B	269	TRP	CD2-CE2	5.22	1.47	1.41
1	D	269	TRP	CD2-CE2	5.17	1.47	1.41
1	D	364	TRP	CD2-CE2	5.16	1.47	1.41
1	D	135	TRP	CD2-CE2	5.11	1.47	1.41
1	C	393	TRP	CD2-CE2	5.10	1.47	1.41
1	A	393	TRP	CD2-CE2	5.10	1.47	1.41
1	C	342	TRP	CD2-CE2	5.08	1.47	1.41
1	D	342	TRP	CD2-CE2	5.07	1.47	1.41
1	B	367	TRP	CD2-CE2	5.06	1.47	1.41
1	D	99	TRP	CD2-CE2	5.06	1.47	1.41
1	D	367	TRP	CD2-CE2	5.05	1.47	1.41
1	D	393	TRP	CD2-CE2	5.05	1.47	1.41

There are no bond angle outliers.

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	2758	0	2618	14	0
1	B	2740	0	2604	10	0
1	C	2764	0	2628	15	0
1	D	2737	0	2601	19	0
2	A	28	0	26	0	0
2	B	28	0	26	0	0
2	C	28	0	26	0	0
2	D	28	0	26	0	0
3	A	77	0	70	1	0
3	B	77	0	70	0	0
3	C	66	0	60	0	0
3	D	77	0	70	1	0
4	A	13	0	7	1	0
4	B	13	0	7	1	0
4	C	13	0	7	0	0
4	D	13	0	7	0	0
5	A	44	0	37	3	0
5	C	44	0	37	1	0
6	B	22	0	19	0	0
6	D	22	0	19	0	0
7	A	329	0	0	1	0
7	B	287	0	0	1	0
7	C	285	0	0	1	0
7	D	257	0	0	4	0
All	All	12750	0	10965	62	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 3.

All (62) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:305:ASN:HD22	1:A:306:TYR:H	1.21	0.87
1:B:305:ASN:HD22	1:B:306:TYR:H	1.25	0.83
1:D:305:ASN:HD22	1:D:306:TYR:H	1.25	0.81

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:C:305:ASN:HD22	1:C:306:TYR:H	1.29	0.81
1:B:305:ASN:ND2	1:B:306:TYR:H	1.84	0.76
1:C:301:THR:OG1	1:C:349[B]:THR:HG22	1.88	0.74
5:A:603:BGC:H3	7:A:2112:HOH:O	1.89	0.73
1:A:301:THR:OG1	1:A:349[B]:THR:HG22	1.89	0.71
1:D:305:ASN:ND2	1:D:306:TYR:H	1.91	0.69
1:A:305:ASN:ND2	1:A:306:TYR:H	1.93	0.66
1:D:356:LYS:HE3	7:D:2192:HOH:O	1.97	0.65
1:B:277:ASP:OD1	1:B:340:HIS:HE1	1.84	0.60
3:A:509:MAN:H62	1:D:150:ALA:HB1	1.84	0.59
1:D:182:ASN:HB2	7:D:2071:HOH:O	2.03	0.57
1:C:298:GLY:HA3	1:C:346:PHE:O	2.06	0.56
1:C:410:ARG:HH21	1:C:426:GLN:HE21	1.55	0.55
1:C:305:ASN:ND2	1:C:306:TYR:H	2.02	0.54
1:D:118:GLY:O	3:D:509:MAN:H2	2.07	0.54
1:A:410:ARG:HH21	1:A:426:GLN:HE21	1.55	0.53
1:C:277:ASP:OD1	1:C:340:HIS:HE1	1.92	0.53
1:A:109:SER:HA	1:A:113:ILE:HD12	1.91	0.52
1:C:430:TRP:HE1	1:C:432:GLN:HE21	1.58	0.50
1:D:277:ASP:OD1	1:D:340:HIS:HE1	1.95	0.50
1:B:220:PRO:HA	1:B:263:ASP:CG	2.33	0.49
1:C:430:TRP:HE1	1:C:432:GLN:NE2	2.11	0.49
1:C:262:LEU:HG	1:C:296:LEU:HD11	1.94	0.49
1:B:182:ASN:HB2	7:B:2086:HOH:O	2.13	0.48
1:C:250:VAL:HG11	1:C:296:LEU:HD22	1.95	0.47
1:D:157:LYS:HG3	7:D:2060:HOH:O	2.14	0.47
5:C:603:BGC:H3	7:C:2084:HOH:O	2.15	0.47
1:C:220:PRO:HA	1:C:263:ASP:CG	2.35	0.47
1:A:220:PRO:HA	1:A:263:ASP:CG	2.36	0.46
1:A:430:TRP:HE1	1:A:432:GLN:HE21	1.62	0.46
1:A:430:TRP:HE1	1:A:432:GLN:NE2	2.14	0.45
1:B:93:PHE:CE2	1:B:215:LEU:HD13	2.52	0.45
1:D:220:PRO:HA	1:D:263:ASP:CG	2.38	0.44
1:B:220:PRO:HA	1:B:263:ASP:CB	2.48	0.44
1:D:277:ASP:O	1:D:281:GLN:HG3	2.17	0.44
1:A:262:LEU:HG	1:A:296:LEU:HD11	1.99	0.44
1:D:298:GLY:HA3	1:D:346:PHE:O	2.17	0.44
1:A:112:ALA:HA	1:A:432:GLN:HE22	1.83	0.44
1:C:109:SER:HA	1:C:113:ILE:HD12	2.00	0.43
1:D:250:VAL:CG1	1:D:293:PRO:HG2	2.48	0.43
1:D:294:ARG:HD2	7:D:2136:HOH:O	2.17	0.43
1:B:97:THR:O	1:B:162:TYR:HA	2.18	0.43

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:298:GLY:HA3	1:A:346:PHE:O	2.19	0.43
4:A:600:4MU:HM4	4:A:600:4MU:H5	1.81	0.42
1:C:311:ILE:HG21	1:C:357:GLN:HG3	2.01	0.42
1:C:220:PRO:HA	1:C:263:ASP:CB	2.49	0.42
1:A:135:TRP:CE3	5:A:604:BGC:H5	2.55	0.42
1:D:220:PRO:HA	1:D:263:ASP:CB	2.50	0.42
1:B:157:LYS:C	1:B:159:GLY:H	2.23	0.41
1:C:112:ALA:HA	1:C:432:GLN:HE22	1.84	0.41
4:B:600:4MU:H5	4:B:600:4MU:HM4	1.92	0.41
1:B:298:GLY:HA3	1:B:346:PHE:O	2.20	0.41
1:A:220:PRO:HA	1:A:263:ASP:CB	2.50	0.41
1:D:250:VAL:HG12	1:D:293:PRO:HG2	2.04	0.41
1:D:262:LEU:HG	1:D:296:LEU:HD11	2.02	0.40
1:D:87:THR:HG22	1:D:256:PRO:HB2	2.03	0.40
1:A:181:SER:OG	5:A:604:BGC:O2	2.21	0.40
1:D:298:GLY:HA2	1:D:342:TRP:CZ2	2.57	0.40
1:D:374:GLY:HA2	1:D:438:LEU:HA	2.02	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	363/363 (100%)	353 (97%)	10 (3%)	0	100	100
1	B	359/363 (99%)	343 (96%)	15 (4%)	1 (0%)	50	27
1	C	364/363 (100%)	351 (96%)	13 (4%)	0	100	100
1	D	359/363 (99%)	345 (96%)	14 (4%)	0	100	100
All	All	1445/1452 (100%)	1392 (96%)	52 (4%)	1 (0%)	59	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	158	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	285/283 (101%)	282 (99%)	3 (1%)	84	72
1	B	283/283 (100%)	280 (99%)	3 (1%)	84	72
1	C	285/283 (101%)	280 (98%)	5 (2%)	71	53
1	D	283/283 (100%)	277 (98%)	6 (2%)	66	45
All	All	1136/1132 (100%)	1119 (98%)	17 (2%)	76	60

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	181	SER
1	A	305	ASN
1	A	431	PHE
1	B	182	ASN
1	B	294	ARG
1	B	431	PHE
1	C	213	ARG
1	C	294	ARG
1	C	305	ASN
1	C	319	GLN
1	C	431	PHE
1	D	171	LEU
1	D	182	ASN
1	D	294	ARG
1	D	319	GLN
1	D	362	GLN
1	D	431	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (28) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	182	ASN
1	A	281	GLN
1	A	305	ASN
1	A	426	GLN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	A	432	GLN
1	A	441	ASN
1	B	158	ASN
1	B	285	ASN
1	B	305	ASN
1	B	340	HIS
1	B	382	ASN
1	C	161	ASN
1	C	182	ASN
1	C	204	GLN
1	C	285	ASN
1	C	305	ASN
1	C	319	GLN
1	C	363	GLN
1	C	426	GLN
1	C	432	GLN
1	C	441	ASN
1	D	158	ASN
1	D	285	ASN
1	D	305	ASN
1	D	319	GLN
1	D	340	HIS
1	D	426	GLN
1	D	441	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 ⓘ

12 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	$\# Z  > 2$
5	BGC	A	601	5,4	10,11,12	0.55	0	11,15,17	0.78	0
5	BGC	A	602	5	10,11,12	1.05	1 (10%)	11,15,17	0.68	0
5	BGC	A	603	5	10,11,12	0.81	0	11,15,17	1.77	3 (27%)
5	BGC	A	604	5	10,11,12	0.81	0	11,15,17	1.35	1 (9%)
6	BGC	B	601	4,6	10,11,12	0.49	0	11,15,17	1.16	0
6	BGC	B	602	6	10,11,12	0.49	0	11,15,17	0.52	0
5	BGC	C	601	5,4	10,11,12	0.68	0	11,15,17	0.85	0
5	BGC	C	602	5	10,11,12	1.17	1 (10%)	11,15,17	0.84	0
5	BGC	C	603	5	10,11,12	0.55	0	11,15,17	1.87	2 (18%)
5	BGC	C	604	5	10,11,12	0.38	0	11,15,17	0.94	0
6	BGC	D	601	4,6	10,11,12	0.59	0	11,15,17	1.01	1 (9%)
6	BGC	D	602	6	10,11,12	0.62	0	11,15,17	0.60	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
5	BGC	A	601	5,4	-	0/2/19/22	0/1/1/1
5	BGC	A	602	5	-	0/2/19/22	0/1/1/1
5	BGC	A	603	5	-	0/2/19/22	1/1/1/1
5	BGC	A	604	5	-	0/2/19/22	0/1/1/1
6	BGC	B	601	4,6	-	0/2/19/22	0/1/1/1
6	BGC	B	602	6	-	0/2/19/22	0/1/1/1
5	BGC	C	601	5,4	-	0/2/19/22	0/1/1/1
5	BGC	C	602	5	-	0/2/19/22	0/1/1/1
5	BGC	C	603	5	-	0/2/19/22	0/1/1/1
5	BGC	C	604	5	-	0/2/19/22	0/1/1/1
6	BGC	D	601	4,6	-	0/2/19/22	0/1/1/1
6	BGC	D	602	6	-	0/2/19/22	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	602	BGC	O4-C4	2.51	1.49	1.43
5	C	602	BGC	O4-C4	2.39	1.48	1.43

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	603	BGC	O5-C5-C6	4.51	111.71	106.98
5	A	603	BGC	O5-C5-C6	4.03	111.21	106.98
5	A	604	BGC	O5-C5-C6	3.51	110.66	106.98
6	D	601	BGC	O5-C5-C6	-2.68	104.17	106.98
5	C	603	BGC	O4-C4-C3	-2.60	104.53	110.35
5	A	603	BGC	O2-C2-C3	-2.10	105.64	110.18
5	A	603	BGC	O5-C5-C4	2.02	113.22	110.65

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	603	BGC	C1-C2-C3-C4-C5-O5

## 5.6 Ligand geometry ⓘ

39 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	NAG	A	501	1	12,14,15	0.66	0	15,19,21	0.99	0
2	NAG	A	502	1	12,14,15	0.61	0	15,19,21	0.70	0
3	MAN	A	503	1	10,11,12	0.81	1 (10%)	11,15,17	1.51	2 (18%)
3	MAN	A	504	1	10,11,12	0.72	0	11,15,17	1.34	3 (27%)
3	MAN	A	505	1	10,11,12	0.76	0	11,15,17	0.90	1 (9%)
3	MAN	A	506	1	10,11,12	0.73	0	11,15,17	0.59	0
3	MAN	A	507	1	10,11,12	0.67	0	11,15,17	0.77	0
3	MAN	A	508	1	10,11,12	0.67	0	11,15,17	0.59	0
3	MAN	A	509	1	10,11,12	0.72	0	11,15,17	1.64	2 (18%)
4	4MU	A	600	5	14,14,14	2.68	2 (14%)	18,20,20	1.23	1 (5%)
2	NAG	B	501	1	12,14,15	0.55	0	15,19,21	1.00	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	B	502	1	12,14,15	0.67	0	15,19,21	0.70	0
3	MAN	B	503	1	10,11,12	0.74	0	11,15,17	0.76	0
3	MAN	B	504	1	10,11,12	0.79	0	11,15,17	0.96	1 (9%)
3	MAN	B	505	1	10,11,12	0.72	0	11,15,17	1.09	1 (9%)
3	MAN	B	506	1	10,11,12	0.69	0	11,15,17	0.66	0
3	MAN	B	507	1	10,11,12	0.68	0	11,15,17	1.08	1 (9%)
3	MAN	B	508	1	10,11,12	0.68	0	11,15,17	0.86	1 (9%)
3	MAN	B	509	1	10,11,12	0.65	0	11,15,17	0.97	1 (9%)
4	4MU	B	600	6	14,14,14	2.24	1 (7%)	18,20,20	1.53	4 (22%)
2	NAG	C	501	1	12,14,15	0.71	0	15,19,21	1.34	2 (13%)
2	NAG	C	502	1	12,14,15	0.60	0	15,19,21	0.83	0
3	MAN	C	504	1	10,11,12	0.73	0	11,15,17	1.32	1 (9%)
3	MAN	C	505	1	10,11,12	0.76	0	11,15,17	0.44	0
3	MAN	C	506	1	10,11,12	0.76	0	11,15,17	0.93	0
3	MAN	C	507	1	10,11,12	0.69	0	11,15,17	0.66	0
3	MAN	C	508	1	10,11,12	0.76	0	11,15,17	1.72	1 (9%)
3	MAN	C	509	1	10,11,12	0.73	0	11,15,17	1.18	2 (18%)
4	4MU	C	600	5	14,14,14	3.03	2 (14%)	18,20,20	1.38	2 (11%)
2	NAG	D	501	1	12,14,15	0.60	0	15,19,21	0.81	1 (6%)
2	NAG	D	502	1	12,14,15	0.54	0	15,19,21	1.37	2 (13%)
3	MAN	D	503	1	10,11,12	0.73	0	11,15,17	0.80	0
3	MAN	D	504	1	10,11,12	0.70	0	11,15,17	0.75	0
3	MAN	D	505	1	10,11,12	0.72	0	11,15,17	0.93	0
3	MAN	D	506	1	10,11,12	0.66	0	11,15,17	0.59	0
3	MAN	D	507	1	10,11,12	0.72	0	11,15,17	0.84	0
3	MAN	D	508	1	10,11,12	0.71	0	11,15,17	0.88	1 (9%)
3	MAN	D	509	1	10,11,12	0.66	0	11,15,17	1.26	1 (9%)
4	4MU	D	600	6	14,14,14	2.33	1 (7%)	18,20,20	1.66	3 (16%)

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	501	1	-	0/6/23/26	0/1/1/1
2	NAG	A	502	1	-	0/6/23/26	0/1/1/1
3	MAN	A	503	1	-	0/2/19/22	0/1/1/1
3	MAN	A	504	1	-	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	505	1	-	0/2/19/22	0/1/1/1
3	MAN	A	506	1	-	0/2/19/22	0/1/1/1
3	MAN	A	507	1	-	0/2/19/22	0/1/1/1
3	MAN	A	508	1	-	0/2/19/22	0/1/1/1
3	MAN	A	509	1	-	0/2/19/22	0/1/1/1
4	4MU	A	600	5	-	0/0/0/0	0/0/2/2
2	NAG	B	501	1	-	0/6/23/26	0/1/1/1
2	NAG	B	502	1	-	0/6/23/26	0/1/1/1
3	MAN	B	503	1	-	0/2/19/22	0/1/1/1
3	MAN	B	504	1	-	0/2/19/22	0/1/1/1
3	MAN	B	505	1	-	0/2/19/22	0/1/1/1
3	MAN	B	506	1	-	0/2/19/22	0/1/1/1
3	MAN	B	507	1	-	0/2/19/22	0/1/1/1
3	MAN	B	508	1	-	0/2/19/22	0/1/1/1
3	MAN	B	509	1	-	0/2/19/22	0/1/1/1
4	4MU	B	600	6	-	0/0/0/0	0/0/2/2
2	NAG	C	501	1	-	0/6/23/26	0/1/1/1
2	NAG	C	502	1	-	0/6/23/26	0/1/1/1
3	MAN	C	504	1	-	0/2/19/22	0/1/1/1
3	MAN	C	505	1	-	0/2/19/22	0/1/1/1
3	MAN	C	506	1	-	0/2/19/22	0/1/1/1
3	MAN	C	507	1	-	0/2/19/22	0/1/1/1
3	MAN	C	508	1	-	0/2/19/22	0/1/1/1
3	MAN	C	509	1	-	0/2/19/22	0/1/1/1
4	4MU	C	600	5	-	0/0/0/0	0/0/2/2
2	NAG	D	501	1	-	0/6/23/26	0/1/1/1
2	NAG	D	502	1	1/1/5/7	0/6/23/26	0/1/1/1
3	MAN	D	503	1	-	0/2/19/22	0/1/1/1
3	MAN	D	504	1	-	0/2/19/22	0/1/1/1
3	MAN	D	505	1	-	0/2/19/22	0/1/1/1
3	MAN	D	506	1	-	0/2/19/22	0/1/1/1
3	MAN	D	507	1	-	0/2/19/22	0/1/1/1
3	MAN	D	508	1	-	0/2/19/22	0/1/1/1
3	MAN	D	509	1	-	0/2/19/22	0/1/1/1
4	4MU	D	600	6	-	0/0/0/0	0/0/2/2

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	600	4MU	O1-C2	9.54	1.37	1.33
4	D	600	4MU	O1-C2	8.07	1.37	1.33
4	A	600	4MU	O1-C2	8.05	1.37	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	600	4MU	O1-C2	7.59	1.37	1.33
4	C	600	4MU	CM4-C4	-5.27	1.40	1.51
4	A	600	4MU	CM4-C4	-5.18	1.40	1.51
3	A	503	MAN	O5-C5	-2.03	1.41	1.45

All (35) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	508	MAN	O5-C5-C6	5.43	112.68	106.98
3	C	504	MAN	O5-C5-C6	3.95	111.12	106.98
4	D	600	4MU	C2-O1-C8A	-3.94	116.96	122.33
3	A	509	MAN	C4-C3-C2	3.82	115.64	110.50
2	D	502	NAG	O5-C5-C6	3.59	110.75	106.98
3	D	509	MAN	O5-C5-C6	3.50	110.65	106.98
3	A	503	MAN	O5-C5-C6	3.46	110.61	106.98
3	A	509	MAN	O5-C5-C6	3.44	110.59	106.98
4	B	600	4MU	C2-O1-C8A	-3.35	117.76	122.33
3	B	507	MAN	O5-C5-C6	3.33	110.48	106.98
3	B	505	MAN	O5-C5-C6	3.12	110.25	106.98
4	C	600	4MU	O1-C8A-C8	2.96	119.71	115.90
4	D	600	4MU	O1-C2-C3	2.89	122.80	119.41
3	A	503	MAN	C4-C3-C2	2.80	114.27	110.50
3	C	509	MAN	O5-C5-C6	2.76	109.88	106.98
4	D	600	4MU	C5-C4A-C8A	2.69	119.59	116.30
3	A	504	MAN	O5-C5-C6	2.61	109.72	106.98
2	B	501	NAG	O5-C5-C6	2.60	109.71	106.98
3	A	505	MAN	O5-C5-C6	2.55	109.66	106.98
2	C	501	NAG	C2-N2-C7	2.49	127.27	123.09
3	D	508	MAN	O5-C5-C6	2.41	109.51	106.98
4	B	600	4MU	O1-C2-C3	2.36	122.18	119.41
3	B	508	MAN	O5-C5-C6	2.36	109.45	106.98
3	C	509	MAN	C4-C3-C2	2.31	113.61	110.50
3	B	504	MAN	O5-C5-C6	2.27	109.36	106.98
4	B	600	4MU	C5-C4A-C8A	2.26	119.07	116.30
3	A	504	MAN	C4-C3-C2	2.18	113.43	110.50
4	A	600	4MU	O1-C8A-C8	2.16	118.68	115.90
3	B	509	MAN	O5-C5-C6	2.13	109.22	106.98
4	C	600	4MU	C8-C8A-C4A	-2.11	120.52	123.10
2	D	502	NAG	C3-C2-N2	-2.11	108.55	111.76
2	C	501	NAG	O5-C5-C6	2.10	109.18	106.98
3	A	504	MAN	C3-C4-C5	2.07	113.91	110.20
4	B	600	4MU	C8-C8A-C4A	-2.06	120.58	123.10

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	501	NAG	O5-C5-C6	2.05	109.13	106.98

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	D	502	NAG	C1

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, 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	363/363 (100%)	-0.01	6 (1%) 67 72	7, 14, 23, 29	1 (0%)
1	B	360/363 (99%)	0.11	11 (3%) 47 52	7, 15, 27, 43	0
1	C	363/363 (100%)	0.05	7 (1%) 64 69	9, 16, 24, 37	0
1	D	360/363 (99%)	0.19	12 (3%) 44 49	10, 18, 27, 44	1 (0%)
All	All	1446/1452 (99%)	0.09	36 (2%) 54 60	7, 15, 26, 44	2 (0%)

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	159	GLY	7.8
1	D	160	GLY	5.5
1	B	447	LEU	5.4
1	D	409	PRO	5.0
1	D	158	ASN	4.9
1	D	178	ALA	4.3
1	C	85	THR	4.2
1	D	85	THR	3.9
1	B	160	GLY	3.2
1	B	408	ALA	3.2
1	B	158	ASN	3.1
1	A	320	GLY	3.1
1	C	447	LEU	2.8
1	A	319	GLN	2.6
1	C	319	GLN	2.6
1	D	161	ASN	2.6
1	A	447	LEU	2.6
1	D	133	PHE	2.5
1	C	303	VAL	2.4
1	C	117	THR	2.4
1	C	119	ALA	2.3

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	217	VAL	2.3
1	B	159	GLY	2.3
1	B	407	SER	2.3
1	A	408	ALA	2.3
1	B	409	PRO	2.3
1	A	303	VAL	2.2
1	D	410	ARG	2.2
1	D	157	LYS	2.2
1	B	413	SER	2.2
1	C	320	GLY	2.2
1	B	167	VAL	2.1
1	B	406	SER	2.1
1	D	394	VAL	2.1
1	D	447	LEU	2.1
1	B	217	VAL	2.0

## 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, 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	BGC	C	603	11/12	0.41	11.60	18,19,20,20	10
5	BGC	A	603	11/12	0.34	8.95	17,18,19,21	10
5	BGC	C	604	11/12	0.13	2.04	16,20,21,21	0
5	BGC	A	604	11/12	0.15	1.69	17,21,22,22	0
5	BGC	A	602	11/12	0.12	0.35	14,15,16,17	0
5	BGC	A	601	11/12	0.08	0.20	14,15,16,17	0
5	BGC	C	602	11/12	0.11	0.16	19,20,21,21	0
5	BGC	C	601	11/12	0.08	-0.18	17,19,20,21	0
6	BGC	D	602	11/12	0.09	-0.30	13,15,15,15	0
6	BGC	D	601	11/12	0.08	-0.60	15,16,18,20	0
6	BGC	B	602	11/12	0.08	-0.71	12,13,13,13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	BGC	B	601	11/12	0.08	-0.95	14,14,16,16	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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	B	505	11/12	0.12	8.87	22,23,24,25	0
3	MAN	B	507	11/12	0.18	8.49	27,29,30,32	0
2	NAG	B	501	14/15	0.32	7.57	39,42,43,44	0
3	MAN	A	505	11/12	0.15	7.57	22,24,27,27	0
4	4MU	C	600	13/13	0.22	7.53	24,29,31,32	0
3	MAN	A	509	11/12	0.24	6.46	30,32,33,33	0
3	MAN	D	505	11/12	0.12	5.54	23,26,27,28	0
3	MAN	A	508	11/12	0.11	5.33	24,25,27,28	0
2	NAG	C	501	14/15	0.32	5.21	33,36,41,41	0
3	MAN	B	503	11/12	0.29	5.06	33,38,39,41	0
3	MAN	A	503	11/12	0.25	4.18	31,35,38,39	0
3	MAN	D	509	11/12	0.23	3.87	26,30,32,33	0
3	MAN	C	505	11/12	0.12	2.87	21,23,24,25	0
2	NAG	D	501	14/15	0.25	2.71	38,40,41,43	0
3	MAN	C	504	11/12	0.20	2.62	28,31,32,34	0
3	MAN	A	504	11/12	0.16	2.51	22,26,29,30	0
3	MAN	B	509	11/12	0.31	2.34	37,39,41,42	0
3	MAN	B	504	11/12	0.22	2.32	26,29,32,32	0
4	4MU	B	600	13/13	0.17	2.15	18,22,25,26	0
3	MAN	A	507	11/12	0.14	1.89	25,28,30,30	0
3	MAN	D	503	11/12	0.27	1.77	36,39,40,41	0
3	MAN	A	506	11/12	0.11	1.52	20,22,24,24	0
3	MAN	B	506	11/12	0.12	1.41	21,22,23,25	0
2	NAG	D	502	14/15	0.11	1.35	18,20,22,24	0
3	MAN	C	507	11/12	0.14	1.34	25,27,29,29	0
3	MAN	D	504	11/12	0.23	1.13	32,35,37,37	0
3	MAN	D	508	11/12	0.08	1.09	18,19,20,20	0
3	MAN	D	506	11/12	0.09	0.98	18,20,21,23	0
2	NAG	C	502	14/15	0.11	0.70	19,21,22,24	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	MAN	D	507	11/12	0.18	0.63	26,30,31,32	0
3	MAN	C	506	11/12	0.11	0.51	22,23,24,27	0
4	4MU	A	600	13/13	0.15	0.47	17,21,22,23	0
4	4MU	D	600	13/13	0.15	0.34	19,22,24,24	0
2	NAG	A	502	14/15	0.12	0.23	18,19,21,21	0
3	MAN	C	509	11/12	0.17	0.08	33,34,36,36	0
2	NAG	A	501	14/15	0.11	0.05	28,29,32,33	0
3	MAN	C	508	11/12	0.08	-0.06	21,21,22,23	0
2	NAG	B	502	14/15	0.09	-0.27	16,18,19,20	0
3	MAN	B	508	11/12	0.10	-0.43	29,31,32,32	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.