



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:20 AM GMT

PDB ID : 3I26  
Title : Structure of bovine torovirus Hemagglutinin-Esterase  
Authors : Zeng, Q.H.; Huizinga, E.G.  
Deposited on : 2009-06-29  
Resolution : 1.80 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

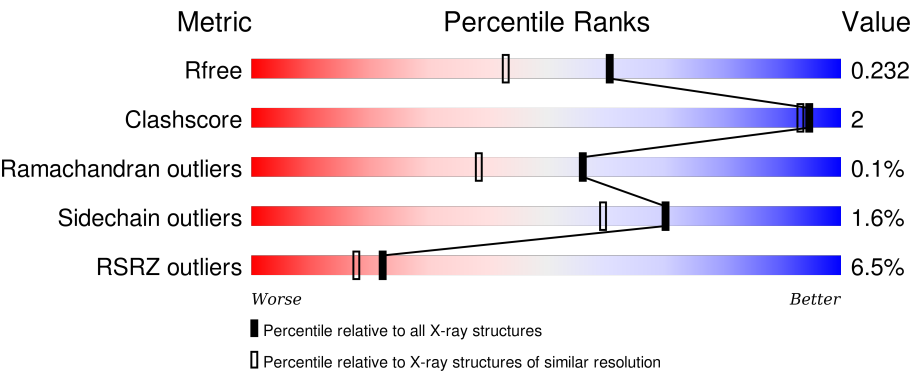
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.80 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	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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	384	<div><div>%</div><div><div></div><div>90%</div><div>5%</div><div>5%</div></div></div>
1	B	384	<div><div></div><div>91%</div><div>• • 5%</div></div>
1	C	384	<div><div>21%</div><div><div></div><div>87%</div><div>6%</div><div>7%</div></div></div>
1	D	384	<div><div>4%</div><div><div></div><div>92%</div><div>• • •</div></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	B	5156	-	-	-	X
4	NAG	A	5156	-	-	-	X
4	NAG	D	5156	-	-	-	X
6	NAG	B	5152	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 13255 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 Hemagglutinin-esterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	365	Total	C	N	O	S	0	0	0
			2862	1821	455	559	27			
1	B	364	Total	C	N	O	S	0	0	0
			2855	1819	454	555	27			
1	C	358	Total	C	N	O	S	0	0	0
			2813	1795	447	544	27			
1	D	369	Total	C	N	O	S	0	0	0
			2889	1837	460	565	27			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	393	ASP	-	EXPRESSION TAG	UNP P0C0V9
A	394	PRO	-	EXPRESSION TAG	UNP P0C0V9
A	395	LEU	-	EXPRESSION TAG	UNP P0C0V9
A	396	VAL	-	EXPRESSION TAG	UNP P0C0V9
A	397	PRO	-	EXPRESSION TAG	UNP P0C0V9
A	398	ARG	-	EXPRESSION TAG	UNP P0C0V9
B	393	ASP	-	EXPRESSION TAG	UNP P0C0V9
B	394	PRO	-	EXPRESSION TAG	UNP P0C0V9
B	395	LEU	-	EXPRESSION TAG	UNP P0C0V9
B	396	VAL	-	EXPRESSION TAG	UNP P0C0V9
B	397	PRO	-	EXPRESSION TAG	UNP P0C0V9
B	398	ARG	-	EXPRESSION TAG	UNP P0C0V9
C	393	ASP	-	EXPRESSION TAG	UNP P0C0V9
C	394	PRO	-	EXPRESSION TAG	UNP P0C0V9
C	395	LEU	-	EXPRESSION TAG	UNP P0C0V9
C	396	VAL	-	EXPRESSION TAG	UNP P0C0V9
C	397	PRO	-	EXPRESSION TAG	UNP P0C0V9
C	398	ARG	-	EXPRESSION TAG	UNP P0C0V9
D	393	ASP	-	EXPRESSION TAG	UNP P0C0V9
D	394	PRO	-	EXPRESSION TAG	UNP P0C0V9
D	395	LEU	-	EXPRESSION TAG	UNP P0C0V9

*Continued on next page...*

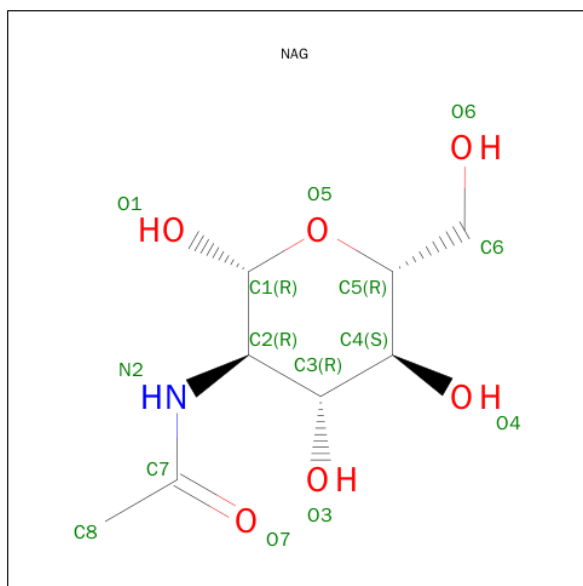
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	396	VAL	-	EXPRESSION TAG	UNP P0C0V9
D	397	PRO	-	EXPRESSION TAG	UNP P0C0V9
D	398	ARG	-	EXPRESSION TAG	UNP P0C0V9

- Molecule 2 is a polymer of unknown type called SUGAR (4-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	4	Total	C	N	O	0	0
			50	28	2	20		
2	A	4	Total	C	N	O	0	0
			50	28	2	20		
2	D	4	Total	C	N	O	0	0
			50	28	2	20		

- 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		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is a polymer of unknown type called SUGAR (2-MER).

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

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	5	Total	C	N	O	0	0
			61	34	2	25		

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

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

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

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

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	415	Total	O	0	0
			415	415		
8	B	440	Total	O	0	0
			440	440		
8	C	87	Total	O	0	0
			87	87		

*Continued on next page...*

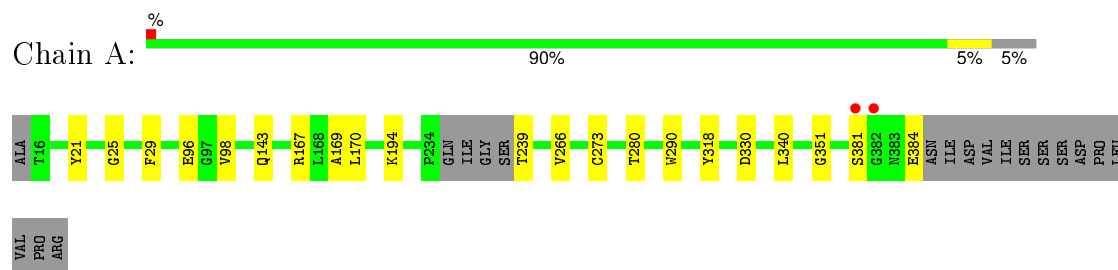
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	326	Total 326	O 326	0	0

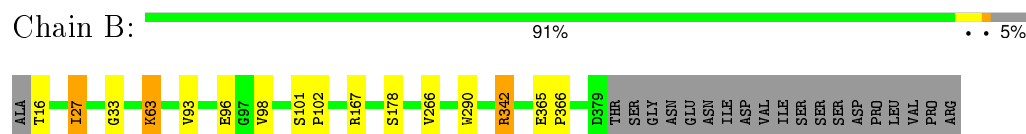
### 3 Residue-property plots

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.

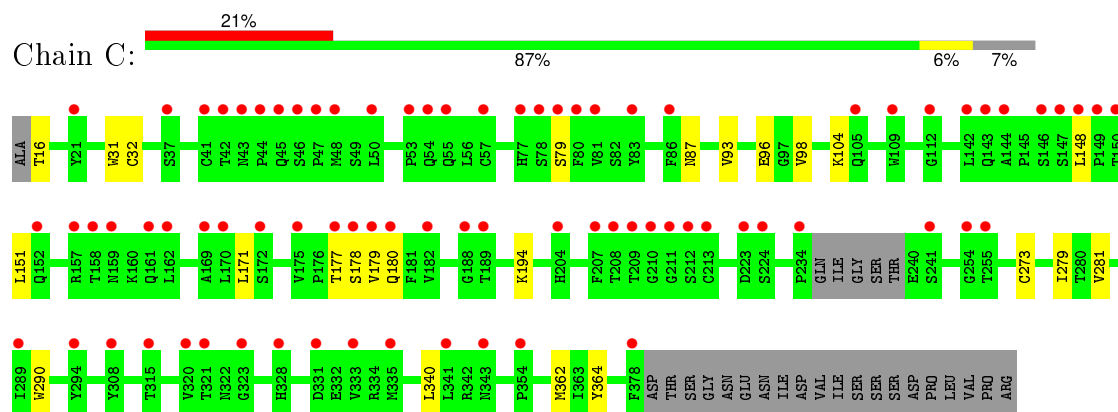
#### • Molecule 1: Hemagglutinin-esterase



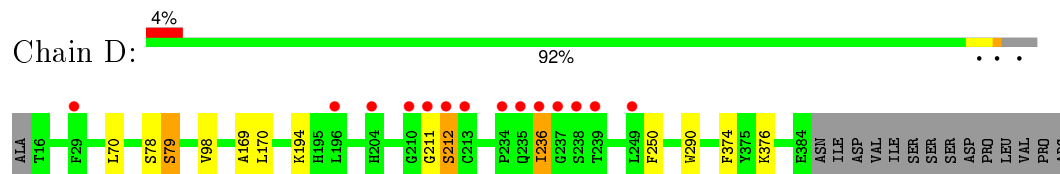
#### • Molecule 1: Hemagglutinin-esterase



#### • Molecule 1: Hemagglutinin-esterase



#### • Molecule 1: Hemagglutinin-esterase





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	67.79Å 113.43Å 273.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.16 – 1.80 48.16 – 1.80	Depositor EDS
% Data completeness (in resolution range)	95.1 (48.16-1.80) 95.2 (48.16-1.80)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, $R_{free}$	0.188 , 0.218 0.203 , 0.232	Depositor DCC
$R_{free}$ test set	9414 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	19.5	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 44.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.41$ , $\langle L^2 \rangle = 0.23$	Xtriage
Outliers	0 of 186978 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13255	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	25.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% 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, 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	0.62	0/2949	0.69	0/4021
1	B	0.66	0/2943	0.71	1/4014 (0.0%)
1	C	0.43	0/2900	0.53	0/3954
1	D	0.54	0/2977	0.65	0/4060
All	All	0.57	0/11769	0.65	1/16049 (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

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	B	27	ILE	CB-CA-C	-5.50	100.59	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	381	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	2862	0	2668	7	0
1	B	2855	0	2671	9	0
1	C	2813	0	2635	13	0
1	D	2889	0	2697	10	0
2	A	100	0	86	0	0
2	D	50	0	43	0	0
3	A	28	0	26	0	0
3	B	14	0	13	0	0
3	D	14	0	13	0	0
4	A	28	0	25	0	0
4	B	56	0	50	0	0
4	D	56	0	50	0	0
5	A	61	0	52	0	0
6	B	39	0	34	0	0
6	C	39	0	34	0	0
7	D	83	0	70	0	0
8	A	415	0	0	1	0
8	B	440	0	0	1	0
8	C	87	0	0	0	0
8	D	326	0	0	1	0
All	All	13255	0	11167	39	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 2.

All (39) 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:B:27:ILE:O	1:B:27:ILE:HG22	1.79	0.80
1:D:211:GLY:HA3	1:D:212:SER:CB	2.13	0.77
1:C:16:THR:O	1:C:16:THR:HG23	1.87	0.74
1:D:211:GLY:HA3	1:D:212:SER:HB2	1.74	0.69
1:B:27:ILE:CG2	1:B:27:ILE:O	2.46	0.64
1:C:98:VAL:HG22	1:C:290:TRP:CE2	2.41	0.55
1:C:148:LEU:HD22	1:C:151:LEU:HD12	1.90	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:280:THR:HG23	8:A:703:HOH:O	2.08	0.52
1:D:211:GLY:HA3	1:D:212:SER:HB3	1.91	0.49
1:A:25:GLY:HA2	1:A:340:LEU:HD23	1.94	0.49
1:C:16:THR:CG2	1:C:16:THR:O	2.59	0.49
1:A:21:TYR:O	1:A:351:GLY:HA3	2.11	0.49
1:D:194:LYS:NZ	8:D:997:HOH:O	2.46	0.49
1:C:32:CYS:HB2	1:C:87:ASN:ND2	2.29	0.47
1:B:167:ARG:O	1:B:266:VAL:HA	2.16	0.46
1:C:279:ILE:HD12	1:C:281:VAL:CG2	2.45	0.46
1:C:171:LEU:HD13	1:C:177:THR:HG21	1.97	0.46
1:D:211:GLY:CA	1:D:212:SER:CB	2.92	0.44
1:B:342:ARG:HD3	8:B:529:HOH:O	2.17	0.43
1:C:171:LEU:CD1	1:C:177:THR:HG21	2.48	0.43
1:A:98:VAL:HG22	1:A:290:TRP:CE2	2.54	0.43
1:C:79:SER:O	1:C:104:LYS:NZ	2.52	0.43
1:A:169:ALA:C	1:A:170:LEU:HD12	2.39	0.42
1:C:93:VAL:HG21	1:C:340:LEU:HD21	2.02	0.42
1:C:31:TRP:CZ3	1:C:340:LEU:HD22	2.54	0.42
1:D:169:ALA:C	1:D:170:LEU:HD12	2.40	0.42
1:A:318:TYR:CZ	1:A:330:ASP:HB2	2.55	0.41
1:B:63:LYS:HD3	1:B:63:LYS:HA	1.96	0.41
1:D:374:PHE:CE1	1:D:376:LYS:HB2	2.55	0.41
1:C:279:ILE:HD12	1:C:281:VAL:HG23	2.02	0.41
1:C:362:MET:HG2	1:C:364:TYR:CZ	2.55	0.41
1:B:365:GLU:HB3	1:B:366:PRO:HA	2.02	0.41
1:B:98:VAL:HG22	1:B:290:TRP:CE2	2.56	0.41
1:A:167:ARG:O	1:A:266:VAL:HA	2.21	0.41
1:B:33:GLY:HA2	1:B:93:VAL:O	2.21	0.40
1:D:78:SER:O	1:D:79:SER:CB	2.70	0.40
1:B:101:SER:HB3	1:B:102:PRO:CD	2.51	0.40
1:D:70:LEU:C	1:D:70:LEU:HD23	2.41	0.40
1:D:98:VAL:HG22	1:D:290:TRP:CE2	2.56	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	361/384 (94%)	348 (96%)	13 (4%)	0	100	100
1	B	362/384 (94%)	348 (96%)	14 (4%)	0	100	100
1	C	354/384 (92%)	338 (96%)	16 (4%)	0	100	100
1	D	367/384 (96%)	351 (96%)	14 (4%)	2 (0%)	34	17
All	All	1444/1536 (94%)	1385 (96%)	57 (4%)	2 (0%)	56	38

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	212	SER
1	D	236	ILE

### 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	327/344 (95%)	320 (98%)	7 (2%)	61	47
1	B	326/344 (95%)	321 (98%)	5 (2%)	72	62
1	C	321/344 (93%)	315 (98%)	6 (2%)	65	52
1	D	330/344 (96%)	327 (99%)	3 (1%)	84	80
All	All	1304/1376 (95%)	1283 (98%)	21 (2%)	70	59

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	29	PHE
1	A	96	GLU
1	A	143	GLN
1	A	194	LYS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	239	THR
1	A	273	CYS
1	A	384	GLU
1	B	16	THR
1	B	63	LYS
1	B	96	GLU
1	B	178	SER
1	B	342	ARG
1	C	96	GLU
1	C	178	SER
1	C	179	VAL
1	C	180	GLN
1	C	194	LYS
1	C	273	CYS
1	D	79	SER
1	D	236	ILE
1	D	250	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	54	GLN
1	B	161	GLN
1	C	152	GLN
1	C	293	GLN

### 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 ⓘ

40 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	5151	1,2	14,14,15	0.64	1 (7%)	15,19,21	1.58	1 (6%)
2	NAG	A	5152	2	14,14,15	0.69	0	15,19,21	1.01	1 (6%)
2	BMA	A	5153	2	11,11,12	0.39	0	14,15,17	0.69	0
5	NAG	A	5154	1,5	14,14,15	0.65	0	15,19,21	1.14	1 (6%)
5	NAG	A	5155	5	14,14,15	0.61	0	15,19,21	1.18	2 (13%)
4	NAG	A	5156	1,4	14,14,15	0.85	0	15,19,21	1.25	1 (6%)
2	NAG	A	5158	1,2	14,14,15	0.54	0	15,19,21	0.81	1 (6%)
2	NAG	A	5159	2	14,14,15	0.44	0	15,19,21	0.80	0
2	BMA	A	5160	2	11,11,12	0.51	0	14,15,17	0.93	1 (7%)
2	MAN	A	5161	2	11,11,12	0.57	0	14,15,17	0.70	0
5	BMA	A	5166	5	11,11,12	0.43	0	14,15,17	1.04	0
5	MAN	A	5167	5	11,11,12	0.54	0	14,15,17	1.12	1 (7%)
5	MAN	A	5168	5	11,11,12	0.60	0	14,15,17	1.03	2 (14%)
2	MAN	A	5170	2	11,11,12	0.58	0	14,15,17	1.10	1 (7%)
4	NAG	A	5174	4	14,14,15	0.43	0	15,19,21	2.33	2 (13%)
6	NAG	B	5151	1,6	14,14,15	0.59	0	15,19,21	1.53	2 (13%)
6	NAG	B	5152	6	14,14,15	0.46	0	15,19,21	1.06	0
6	BMA	B	5153	6	11,11,12	0.52	0	14,15,17	1.18	2 (14%)
4	NAG	B	5157	1,4	14,14,15	0.65	0	15,19,21	0.63	0
4	NAG	B	5158	1,4	14,14,15	0.63	0	15,19,21	1.04	1 (6%)
4	NAG	B	5159	4	14,14,15	0.50	0	15,19,21	0.98	0
4	NAG	B	5173	4	14,14,15	0.48	0	15,19,21	1.24	1 (6%)
6	NAG	C	5158	1,6	14,14,15	0.47	0	15,19,21	1.02	1 (6%)
6	NAG	C	5159	6	14,14,15	0.56	0	15,19,21	0.82	0
6	BMA	C	5160	6	11,11,12	0.49	0	14,15,17	0.93	1 (7%)
2	NAG	D	5151	1,2	14,14,15	0.54	0	15,19,21	0.94	0
2	NAG	D	5152	2	14,14,15	0.69	0	15,19,21	1.00	1 (6%)
2	BMA	D	5153	2	11,11,12	0.51	0	14,15,17	2.67	4 (28%)
4	NAG	D	5154	1,4	14,14,15	0.54	0	15,19,21	1.19	1 (6%)
4	NAG	D	5155	4	14,14,15	0.48	0	15,19,21	1.47	2 (13%)
4	NAG	D	5156	1,4	14,14,15	0.40	0	15,19,21	1.49	1 (6%)
7	NAG	D	5158	1,7	14,14,15	0.53	0	15,19,21	0.71	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	NAG	D	5159	7	14,14,15	0.67	0	15,19,21	1.12	0
7	BMA	D	5160	7	11,11,12	0.36	0	14,15,17	1.10	0
7	MAN	D	5161	7	11,11,12	0.68	0	14,15,17	0.95	0
7	MAN	D	5162	7	11,11,12	0.57	0	14,15,17	0.97	0
7	MAN	D	5163	7	11,11,12	0.51	0	14,15,17	0.89	1 (7%)
7	MAN	D	5164	7	11,11,12	0.56	0	14,15,17	0.75	0
2	MAN	D	5170	2	11,11,12	0.50	0	14,15,17	0.83	1 (7%)
4	NAG	D	5174	4	14,14,15	0.45	0	15,19,21	1.43	3 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	5151	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	5152	2	-	0/6/23/26	0/1/1/1
2	BMA	A	5153	2	-	0/2/19/22	0/1/1/1
5	NAG	A	5154	1,5	-	0/6/23/26	0/1/1/1
5	NAG	A	5155	5	-	0/6/23/26	0/1/1/1
4	NAG	A	5156	1,4	-	0/6/23/26	0/1/1/1
2	NAG	A	5158	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	5159	2	-	0/6/23/26	0/1/1/1
2	BMA	A	5160	2	-	0/2/19/22	0/1/1/1
2	MAN	A	5161	2	-	0/2/19/22	0/1/1/1
5	BMA	A	5166	5	-	0/2/19/22	0/1/1/1
5	MAN	A	5167	5	-	0/2/19/22	0/1/1/1
5	MAN	A	5168	5	-	0/2/19/22	0/1/1/1
2	MAN	A	5170	2	-	0/2/19/22	0/1/1/1
4	NAG	A	5174	4	-	0/6/23/26	0/1/1/1
6	NAG	B	5151	1,6	-	0/6/23/26	0/1/1/1
6	NAG	B	5152	6	-	0/6/23/26	0/1/1/1
6	BMA	B	5153	6	-	0/2/19/22	0/1/1/1
4	NAG	B	5157	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	5158	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	5159	4	-	0/6/23/26	0/1/1/1
4	NAG	B	5173	4	-	0/6/23/26	0/1/1/1
6	NAG	C	5158	1,6	-	0/6/23/26	0/1/1/1
6	NAG	C	5159	6	-	0/6/23/26	0/1/1/1
6	BMA	C	5160	6	-	0/2/19/22	0/1/1/1
2	NAG	D	5151	1,2	-	0/6/23/26	0/1/1/1
2	NAG	D	5152	2	-	0/6/23/26	0/1/1/1

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BMA	D	5153	2	-	0/2/19/22	0/1/1/1
4	NAG	D	5154	1,4	-	0/6/23/26	0/1/1/1
4	NAG	D	5155	4	-	0/6/23/26	0/1/1/1
4	NAG	D	5156	1,4	-	0/6/23/26	0/1/1/1
7	NAG	D	5158	1,7	-	0/6/23/26	0/1/1/1
7	NAG	D	5159	7	-	0/6/23/26	0/1/1/1
7	BMA	D	5160	7	-	0/2/19/22	0/1/1/1
7	MAN	D	5161	7	-	0/2/19/22	0/1/1/1
7	MAN	D	5162	7	-	0/2/19/22	0/1/1/1
7	MAN	D	5163	7	-	0/2/19/22	0/1/1/1
7	MAN	D	5164	7	-	0/2/19/22	1/1/1/1
2	MAN	D	5170	2	-	0/2/19/22	0/1/1/1
4	NAG	D	5174	4	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	5151	NAG	O5-C1	-2.04	1.40	1.43

All (36) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	5174	NAG	C4-C3-C2	-2.78	106.91	111.23
4	B	5158	NAG	C2-N2-C7	-2.57	119.74	123.04
4	A	5156	NAG	O7-C7-C8	-2.49	117.49	122.06
6	B	5151	NAG	O4-C4-C5	-2.30	103.15	109.24
4	D	5155	NAG	C3-C4-C5	-2.24	106.29	110.20
4	D	5174	NAG	O4-C4-C3	-2.18	105.42	110.34
2	D	5152	NAG	O4-C4-C3	-2.12	105.57	110.34
4	D	5174	NAG	C3-C2-N2	-2.11	105.51	110.56
5	A	5155	NAG	O4-C4-C3	-2.07	105.67	110.34
4	D	5174	NAG	O3-C3-C4	-2.06	105.69	110.34
5	A	5154	NAG	O7-C7-C8	-2.03	118.33	122.06
6	B	5153	BMA	C2-C3-C4	-2.01	107.63	111.04
2	A	5170	MAN	C3-C4-C5	2.08	113.83	110.20
5	A	5155	NAG	C2-N2-C7	2.22	125.89	123.04
2	A	5158	NAG	C1-O5-C5	2.23	115.08	112.25
5	A	5168	MAN	C1-C2-C3	2.24	112.19	109.54
5	A	5168	MAN	C1-O5-C5	2.30	115.17	112.25
6	C	5160	BMA	C1-O5-C5	2.30	115.17	112.25
6	B	5153	BMA	O5-C5-C6	2.32	112.36	107.35
7	D	5163	MAN	C1-O5-C5	2.32	115.20	112.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	5170	MAN	C1-O5-C5	2.35	115.23	112.25
2	A	5160	BMA	C1-C2-C3	2.38	112.35	109.54
2	A	5152	NAG	O5-C5-C6	2.42	112.59	107.35
2	D	5153	BMA	C1-C2-C3	2.60	112.62	109.54
4	D	5154	NAG	C2-N2-C7	2.62	126.40	123.04
5	A	5167	MAN	C1-O5-C5	2.92	115.95	112.25
4	D	5155	NAG	C1-O5-C5	2.97	116.01	112.25
2	D	5153	BMA	O5-C1-C2	3.06	115.81	110.86
2	D	5153	BMA	C3-C4-C5	3.48	116.27	110.20
6	C	5158	NAG	C1-O5-C5	3.52	116.71	112.25
6	B	5151	NAG	C1-O5-C5	3.73	116.98	112.25
4	B	5173	NAG	C1-O5-C5	3.98	117.30	112.25
4	D	5156	NAG	C1-O5-C5	4.45	117.89	112.25
2	A	5151	NAG	C1-O5-C5	5.11	118.73	112.25
2	D	5153	BMA	C1-O5-C5	7.90	122.27	112.25
4	A	5174	NAG	C1-O5-C5	8.16	122.60	112.25

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	5164	MAN	C1-C2-C3-C4-C5-O5

No monomer is involved in short contacts.

## 5.6 Ligand geometry [i](#)

4 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	$\# Z  > 2$
3	NAG	A	5157	1	14,14,15	0.45	0	15,19,21	0.90	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	A	5180	1	14,14,15	0.54	0	15,19,21	1.15	1 (6%)
3	NAG	B	5156	1	14,14,15	0.40	0	15,19,21	0.86	1 (6%)
3	NAG	D	5157	1	14,14,15	0.50	0	15,19,21	1.01	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	5157	1	-	0/6/23/26	0/1/1/1
3	NAG	A	5180	1	-	0/6/23/26	0/1/1/1
3	NAG	B	5156	1	-	0/6/23/26	0/1/1/1
3	NAG	D	5157	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	5156	NAG	C1-O5-C5	2.04	114.84	112.25
3	A	5180	NAG	C4-C3-C2	3.07	116.00	111.23

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	365/384 (95%)	-0.07	2 (0%) 91 90	13, 18, 26, 43	0
1	B	364/384 (94%)	-0.12	0 100 100	13, 18, 26, 31	0
1	C	358/384 (93%)	1.10	79 (22%) 1 1	20, 33, 44, 49	0
1	D	369/384 (96%)	0.19	14 (3%) 44 38	17, 23, 32, 48	0
All	All	1456/1536 (94%)	0.27	95 (6%) 22 18	13, 21, 39, 49	0

All (95) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	236	ILE	8.0
1	D	238	SER	7.8
1	D	237	GLY	7.1
1	D	239	THR	6.6
1	C	78	SER	6.3
1	D	210	GLY	4.8
1	D	235	GLN	4.5
1	D	211	GLY	4.5
1	C	148	LEU	4.4
1	C	189	THR	4.3
1	C	41	CYS	4.3
1	C	209	THR	4.2
1	C	147	SER	4.2
1	D	212	SER	4.2
1	C	112	GLY	4.2
1	C	321	THR	4.0
1	C	223	ASP	3.9
1	C	177	THR	3.9
1	C	46	SER	3.8
1	C	42	THR	3.8
1	C	159	ASN	3.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	146	SER	3.6
1	C	80	PHE	3.6
1	C	182	VAL	3.5
1	C	143	GLN	3.4
1	C	180	GLN	3.4
1	C	47	PRO	3.3
1	C	341	LEU	3.3
1	C	204	HIS	3.3
1	C	255	THR	3.3
1	C	57	CYS	3.3
1	C	44	PRO	3.2
1	C	170	LEU	3.2
1	C	54	GLN	3.1
1	C	48	MET	3.1
1	C	158	THR	3.1
1	C	315	THR	3.0
1	C	224	SER	3.0
1	C	241	SER	3.0
1	C	157	ARG	2.9
1	C	162	LEU	2.9
1	D	234	PRO	2.9
1	C	105	GLN	2.8
1	C	175	VAL	2.8
1	C	178	SER	2.8
1	C	172	SER	2.8
1	C	323	GLY	2.7
1	C	333	VAL	2.7
1	C	294	TYR	2.7
1	D	29	PHE	2.7
1	C	43	ASN	2.6
1	C	210	GLY	2.6
1	C	50	LEU	2.6
1	C	208	THR	2.6
1	C	207	PHE	2.6
1	C	21	TYR	2.6
1	C	161	GLN	2.6
1	C	149	PRO	2.5
1	D	196	LEU	2.5
1	C	86	PHE	2.5
1	C	320	VAL	2.5
1	C	335	MET	2.5
1	C	77	HIS	2.5

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	289	ILE	2.4
1	C	254	GLY	2.4
1	C	53	PRO	2.4
1	C	378	PHE	2.4
1	C	150	THR	2.3
1	C	179	VAL	2.3
1	C	83	TYR	2.3
1	D	249	LEU	2.3
1	C	81	VAL	2.2
1	A	382	GLY	2.2
1	C	109	TRP	2.2
1	C	211	GLY	2.2
1	C	354	PRO	2.2
1	D	213	CYS	2.2
1	C	55	GLN	2.2
1	C	144	ALA	2.2
1	C	142	LEU	2.2
1	C	328	HIS	2.2
1	C	169	ALA	2.1
1	C	331	ASP	2.1
1	C	37	SER	2.1
1	C	234	PRO	2.1
1	C	45	GLN	2.1
1	C	343	ASN	2.1
1	C	79	SER	2.1
1	C	308	TYR	2.1
1	C	212	SER	2.0
1	C	188	GLY	2.0
1	C	152	GLN	2.0
1	C	213	CYS	2.0
1	A	381	SER	2.0
1	D	204	HIS	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	NAG	A	5156	14/15	0.91	0.20	2.91	30,33,42,42	0
6	NAG	B	5152	14/15	0.87	0.15	2.69	40,43,47,48	0
4	NAG	D	5156	14/15	0.87	0.21	2.54	44,49,54,57	0
6	NAG	C	5158	14/15	0.79	0.21	1.88	45,48,52,56	0
4	NAG	D	5154	14/15	0.94	0.17	0.38	36,39,43,43	0
7	NAG	D	5158	14/15	0.92	0.10	-0.08	33,36,38,40	0
6	NAG	B	5151	14/15	0.93	0.09	-0.51	27,29,32,36	0
5	NAG	A	5154	14/15	0.97	0.08	-0.65	24,28,34,35	0
2	NAG	A	5158	14/15	0.95	0.07	-1.13	25,28,30,37	0
2	NAG	D	5151	14/15	0.97	0.11	-1.34	20,23,25,27	0
2	NAG	A	5151	14/15	0.96	0.08	-1.58	23,26,28,30	0
4	NAG	B	5158	14/15	0.96	0.08	-2.42	23,26,28,31	0
5	MAN	A	5167	11/12	0.79	0.32	-	72,73,74,74	0
4	NAG	D	5155	14/15	0.88	0.20	-	46,48,52,52	0
6	BMA	B	5153	11/12	0.73	0.23	-	52,54,55,56	0
7	MAN	D	5163	11/12	0.85	0.25	-	59,60,60,61	0
2	MAN	A	5170	11/12	0.86	0.24	-	57,59,61,61	0
4	NAG	B	5159	14/15	0.93	0.11	-	35,37,41,45	0
2	MAN	D	5170	11/12	0.84	0.19	-	49,50,51,53	0
2	NAG	A	5152	14/15	0.93	0.09	-	30,32,37,42	0
2	BMA	D	5153	11/12	0.78	0.21	-	44,48,51,53	0
7	NAG	D	5159	14/15	0.91	0.15	-	36,42,48,49	0
2	BMA	A	5153	11/12	0.80	0.26	-	47,52,54,56	0
6	NAG	C	5159	14/15	0.77	0.27	-	61,63,67,71	0
2	MAN	A	5161	11/12	0.71	0.33	-	67,68,69,70	0
4	NAG	D	5174	14/15	0.72	0.39	-	60,62,63,64	0
7	MAN	D	5162	11/12	0.85	0.18	-	42,46,48,50	0
5	NAG	A	5155	14/15	0.90	0.13	-	38,43,49,56	0
5	MAN	A	5168	11/12	0.75	0.41	-	72,74,75,76	0
5	BMA	A	5166	11/12	0.64	0.23	-	62,67,70,70	0
7	MAN	D	5164	11/12	0.72	0.39	-	66,69,70,70	0
7	MAN	D	5161	11/12	0.84	0.17	-	47,49,52,57	0
6	BMA	C	5160	11/12	0.45	0.30	-	74,75,75,75	0
4	NAG	A	5174	14/15	0.79	0.33	-	46,49,51,51	0
4	NAG	B	5173	14/15	0.74	0.48	-	61,63,66,67	0
7	BMA	D	5160	11/12	0.81	0.21	-	52,55,58,62	0
2	NAG	A	5159	14/15	0.88	0.20	-	40,44,49,55	0
2	BMA	A	5160	11/12	0.78	0.36	-	58,61,64,65	0
4	NAG	B	5157	14/15	0.81	0.31	-	45,51,54,58	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
2	NAG	D	5152	14/15	0.94	0.09	-	26,29,33,39	0

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	5156	14/15	0.80	0.25	11.62	42,48,51,52	0
3	NAG	A	5180	14/15	0.64	0.40	-	56,62,64,65	0
3	NAG	D	5157	14/15	0.88	0.23	-	44,50,53,53	0
3	NAG	A	5157	14/15	0.87	0.33	-	50,56,58,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.