



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 09:49 AM BST

PDB ID : 2NY3  
Title : HIV-1 gp120 Envelope Glycoprotein (K231C, T257S, E267C, S334A, S375W)  
Complexed with CD4 and Antibody 17b  
Authors : Zhou, T.; Xu, L.; Dey, B.; Hessel, A.J.; Van Ryk, D.; Xiang, S.H.; Yang, X.;  
Zhang, M.Y.; Zwick, M.B.; Arthos, J.; Burton, D.R.; Dimitrov, D.S.; Sodroski,  
J.; Wyatt, R.; Nabel, G.J.; Kwong, P.D.  
Deposited on : 2006-11-20  
Resolution : 2.00 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

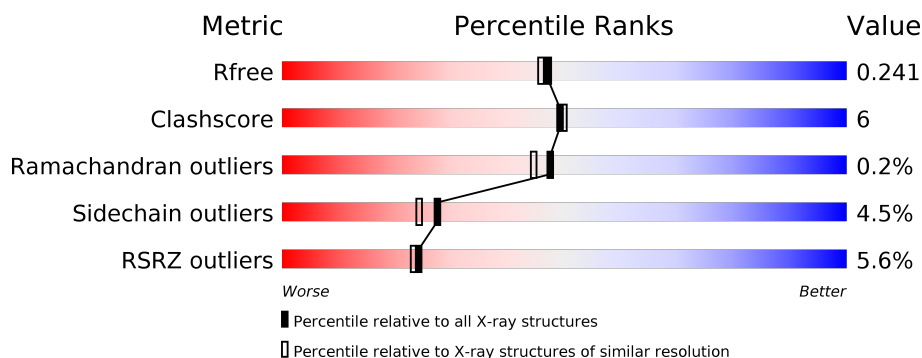
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 respectively. 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	317	<div> <div>9%</div> <div> <div></div> <div>86%</div> <div>9%</div> <div>• •</div> </div> </div>
2	B	184	<div> <div>5%</div> <div> <div></div> <div>78%</div> <div>18%</div> <div>• •</div> </div> </div>
3	C	214	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div>•</div> </div> </div>
4	D	229	<div> <div>3%</div> <div> <div></div> <div>86%</div> <div>10%</div> <div>• •</div> </div> </div>
5	E	2	<div> <div></div> <div> <div>50%</div> <div>50%</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
6	NAG	A	588	-	-	-	X
6	NAG	A	697	-	-	-	X
6	NAG	A	730	-	-	-	X
6	NAG	A	741	-	-	-	X
6	NAG	A	839	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7974 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 ENVELOPE GLYCOPROTEIN GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	305	Total	C	N	O	S	0	0	0
			2357	1480	411	444	22			

- Molecule 2 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	181	Total	C	N	O	S	0	0	0
			1412	885	247	276	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1000	MET	-	insertion	UNP P01730

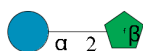
- Molecule 3 is a protein called ANTIBODY 17B, LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	214	Total	C	N	O	S	0	0	0
			1646	1028	282	331	5			

- Molecule 4 is a protein called ANTIBODY 17B, HEAVY CHAIN.

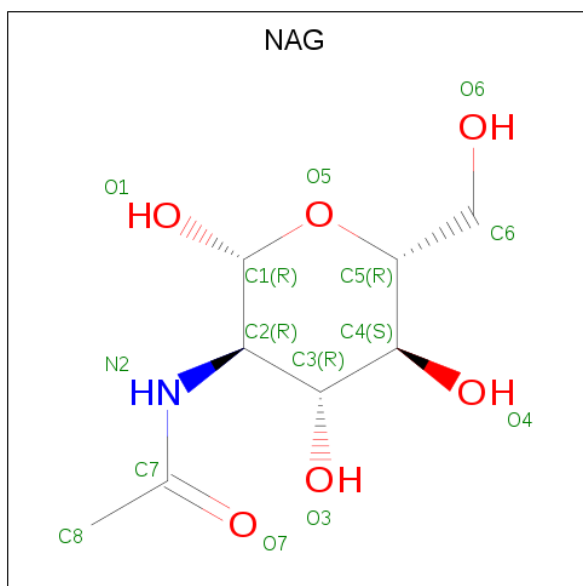
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	223	Total	C	N	O	S	0	2	0
			1689	1068	285	331	5			

- Molecule 5 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
5	E	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	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
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		

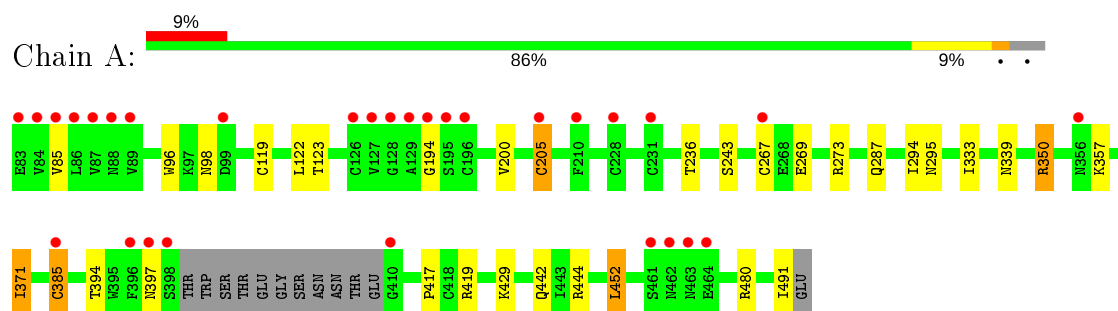
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	215	Total	O	0	0
			215	215		
7	B	99	Total	O	0	0
			99	99		
7	C	162	Total	O	0	0
			162	162		
7	D	189	Total	O	0	0
			189	189		

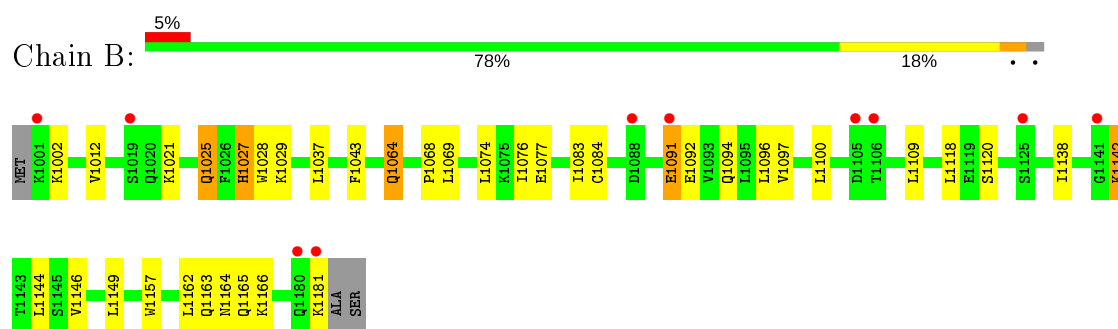
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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.

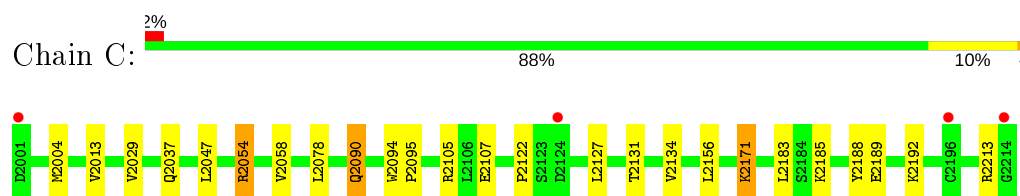
- Molecule 1: ENVELOPE GLYCOPROTEIN GP120



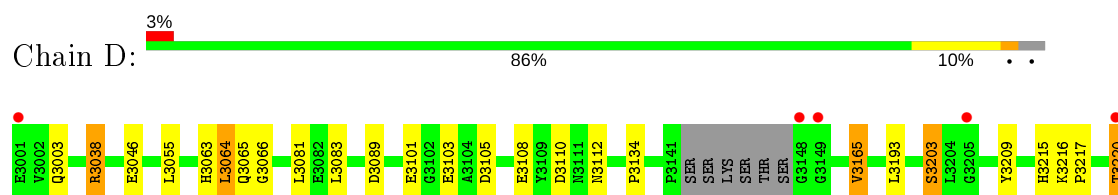
- Molecule 2: T-cell surface glycoprotein CD4



- Molecule 3: ANTIBODY 17B, LIGHT CHAIN



- Molecule 4: ANTIBODY 17B, HEAVY CHAIN





- Molecule 5: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose

Chain E:  50% 50%





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.74Å 88.20Å 197.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.11 – 2.00 44.10 – 2.00	Depositor EDS
% Data completeness (in resolution range)	91.6 (44.11-2.00) 91.6 (44.10-2.00)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.59 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.203 , 0.243 0.202 , 0.241	Depositor DCC
$R_{free}$ test set	3926 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.0	Xtriage
Anisotropy	0.069	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 47.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7974	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: GLC, NAG, FRU

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	0.33	1/2405 (0.0%)	0.48	1/3263 (0.0%)
2	B	0.29	0/1432	0.48	0/1930
3	C	0.35	1/1683 (0.1%)	0.51	0/2288
4	D	0.32	0/1740	0.50	0/2370
All	All	0.32	2/7260 (0.0%)	0.49	1/9851 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	2192	LYS	CE-NZ	6.20	1.64	1.49
1	A	491	ILE	C-O	5.04	1.32	1.23

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	452	LEU	CA-CB-CG	5.49	127.93	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	2357	0	2296	20	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	1412	0	1441	31	0
3	C	1646	0	1590	12	0
4	D	1689	0	1650	23	0
5	E	23	0	21	1	0
6	A	182	0	169	3	0
7	A	215	0	0	7	0
7	B	99	0	0	12	0
7	C	162	0	0	1	0
7	D	189	0	0	6	0
All	All	7974	0	7167	86	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 6.

All (86) 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:D:3215:HIS:HB3	4:D:3220:THR:HG22	1.33	1.05
2:B:1068:PRO:HA	7:B:316:HOH:O	1.57	1.04
1:A:350:ARG:HH11	1:A:350:ARG:HG3	1.33	0.93
2:B:1025:GLN:HG2	7:B:153:HOH:O	1.69	0.91
1:A:419:ARG:NH2	4:D:3103:GLU:OE2	2.06	0.89
4:D:3003:GLN:OE1	7:D:636:HOH:O	1.94	0.85
4:D:3065:GLN:NE2	7:D:36:HOH:O	2.11	0.84
1:A:350:ARG:HH11	1:A:350:ARG:CG	1.95	0.79
4:D:3046:GLU:OE2	4:D:3063[A]:HIS:NE2	2.15	0.79
2:B:1076:ILE:HD12	7:B:173:HOH:O	1.84	0.76
4:D:3112:ASN:ND2	7:D:506:HOH:O	2.20	0.75
4:D:3215:HIS:HB3	4:D:3220:THR:CG2	2.15	0.75
1:A:480:ARG:NH2	7:A:1093:HOH:O	2.21	0.73
2:B:1097:VAL:O	7:B:173:HOH:O	2.06	0.72
2:B:1027:HIS:CD2	7:B:501:HOH:O	2.41	0.71
4:D:3101:GLU:HA	7:D:506:HOH:O	1.90	0.70
2:B:1028:TRP:HZ2	7:B:316:HOH:O	1.76	0.69
6:A:734:NAG:O7	7:A:1163:HOH:O	2.08	0.69
4:D:3134:PRO:HD2	4:D:3220:THR:HG21	1.76	0.67
1:A:123:THR:HG21	1:A:429:LYS:HE3	1.76	0.67
4:D:3110:ASP:HB2	7:D:506:HOH:O	1.95	0.66
2:B:1029:LYS:HG2	7:B:501:HOH:O	1.95	0.65
1:A:119:CYS:N	1:A:205:CYS:SG	2.70	0.65
2:B:1164:ASN:O	2:B:1166:LYS:HD2	1.99	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:1029:LYS:CG	7:B:501:HOH:O	2.46	0.62
2:B:1120:SER:HA	7:B:173:HOH:O	2.01	0.61
4:D:3063[A]:HIS:CD2	4:D:3064:LEU:HD13	2.37	0.59
4:D:3063[A]:HIS:HD2	4:D:3064:LEU:HD13	1.68	0.59
1:A:294:ILE:HD12	1:A:333:ILE:HD11	1.85	0.58
4:D:3063[A]:HIS:CE1	5:E:2:FRU:O6	2.58	0.56
3:C:2188:TYR:CZ	3:C:2213:ARG:HG3	2.40	0.56
2:B:1083:ILE:HG13	2:B:1092:GLU:HG3	1.87	0.56
6:A:762:NAG:O4	7:A:1141:HOH:O	2.07	0.56
2:B:1157:TRP:HB3	7:B:527:HOH:O	2.05	0.56
1:A:295:ASN:HD22	6:A:795:NAG:H83	1.72	0.55
1:A:122:LEU:HD22	1:A:200:VAL:HG22	1.88	0.55
3:C:2171:LYS:HA	3:C:2171:LYS:HE3	1.89	0.54
2:B:1142:LYS:HA	2:B:1142:LYS:NZ	2.24	0.53
3:C:2029:VAL:HG11	3:C:2090:GLN:HG2	1.91	0.52
4:D:3066:GLY:HA3	7:D:254:HOH:O	2.09	0.52
4:D:3203:SER:HG	4:D:3209:TYR:HH	1.56	0.52
3:C:2094:TRP:HA	3:C:2095:PRO:C	2.29	0.52
3:C:2185:LYS:HE2	3:C:2189:GLU:OE2	2.10	0.52
2:B:1084:CYS:HB3	2:B:1091:GLU:HG3	1.93	0.51
2:B:1142:LYS:HA	2:B:1142:LYS:HZ2	1.75	0.51
2:B:1120:SER:OG	2:B:1142:LYS:HE2	2.10	0.51
2:B:1138:ILE:HG21	2:B:1146:VAL:HG22	1.93	0.50
3:C:2029:VAL:HG11	3:C:2090:GLN:CG	2.41	0.50
4:D:3038:ARG:HD2	4:D:3046:GLU:OE1	2.12	0.50
3:C:2037:GLN:HB2	3:C:2047:LEU:HD11	1.93	0.49
3:C:2054:ARG:HG2	3:C:2058:VAL:HB	1.94	0.49
3:C:2004:MET:SD	3:C:2090:GLN:HB2	2.53	0.49
1:A:394:THR:HG22	1:A:397:ASN:HB2	1.95	0.48
2:B:1029:LYS:HG3	7:B:501:HOH:O	2.13	0.47
1:A:385:CYS:SG	1:A:417:PRO:O	2.73	0.47
4:D:3081:LEU:C	4:D:3081:LEU:HD13	2.35	0.47
3:C:2122:PRO:HD3	3:C:2134:VAL:HG22	1.98	0.46
3:C:2105:ARG:NH1	7:C:168:HOH:O	2.49	0.45
2:B:1012:VAL:HB	2:B:1074:LEU:HD11	1.97	0.45
2:B:1094:GLN:HE21	2:B:1096:LEU:HD21	1.82	0.45
1:A:273:ARG:NH2	1:A:287:GLN:OE1	2.49	0.45
2:B:1021:LYS:HE3	2:B:1064:GLN:OE1	2.17	0.44
1:A:350:ARG:NH1	1:A:350:ARG:CG	2.62	0.44
2:B:1181:LYS:HB3	2:B:1181:LYS:HE3	1.84	0.44
4:D:3038:ARG:HG3	4:D:3046:GLU:HB3	2.00	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:D:3105:ASP:OD1	4:D:3105:ASP:N	2.51	0.44
1:A:371:ILE:HG12	2:B:1043:PHE:HB3	2.00	0.43
1:A:480:ARG:NE	7:A:1155:HOH:O	2.44	0.43
4:D:3165:VAL:CG1	4:D:3193:LEU:HD21	2.49	0.42
2:B:1162:LEU:HD23	2:B:1163:GLN:N	2.34	0.42
4:D:3216:LYS:N	4:D:3217:PRO:CD	2.83	0.42
1:A:269:GLU:HB2	7:A:1161:HOH:O	2.20	0.42
1:A:442:GLN:HG2	1:A:444:ARG:HH12	1.85	0.42
2:B:1027:HIS:HD2	7:B:501:HOH:O	1.93	0.41
2:B:1138:ILE:HD12	2:B:1144:LEU:HB3	2.01	0.41
1:A:96:TRP:CD1	1:A:236:THR:HG21	2.56	0.41
1:A:444:ARG:HG3	7:A:1074:HOH:O	2.20	0.41
2:B:1118:LEU:O	2:B:1142:LYS:HE3	2.21	0.41
1:A:85:VAL:HA	1:A:243:SER:HB3	2.03	0.41
2:B:1077:GLU:CD	2:B:1077:GLU:H	2.24	0.40
2:B:1164:ASN:O	2:B:1165:GLN:HB2	2.20	0.40
2:B:1118:LEU:HD23	2:B:1142:LYS:NZ	2.37	0.40
3:C:2013:VAL:CG1	3:C:2078:LEU:HD12	2.51	0.40
4:D:3193:LEU:HD12	4:D:3193:LEU:C	2.42	0.40
2:B:1021:LYS:NZ	2:B:1064:GLN:HG3	2.36	0.40
7:A:1106:HOH:O	4:D:3108:GLU:HB2	2.21	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	301/317 (95%)	283 (94%)	17 (6%)	1 (0%)	41	37
2	B	179/184 (97%)	174 (97%)	5 (3%)	0	100	100
3	C	212/214 (99%)	205 (97%)	7 (3%)	0	100	100
4	D	221/229 (96%)	216 (98%)	4 (2%)	1 (0%)	29	23

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	913/944 (97%)	878 (96%)	33 (4%)	2 (0%)	47	44

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	194	GLY
4	D	3228	PRO

### 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	269/280 (96%)	260 (97%)	9 (3%)	38	37
2	B	164/166 (99%)	153 (93%)	11 (7%)	16	11
3	C	184/184 (100%)	176 (96%)	8 (4%)	29	26
4	D	189/193 (98%)	181 (96%)	8 (4%)	30	27
All	All	806/823 (98%)	770 (96%)	36 (4%)	27	24

All (36) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	98	ASN
1	A	205	CYS
1	A	267	CYS
1	A	339	ASN
1	A	350	ARG
1	A	357	LYS
1	A	371	ILE
1	A	385	CYS
1	A	452	LEU
2	B	1002	LYS
2	B	1025	GLN
2	B	1027	HIS
2	B	1037	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	1064	GLN
2	B	1069	LEU
2	B	1091	GLU
2	B	1100	LEU
2	B	1109	LEU
2	B	1142	LYS
2	B	1149	LEU
3	C	2054	ARG
3	C	2090	GLN
3	C	2107	GLU
3	C	2127	LEU
3	C	2131	THR
3	C	2156	LEU
3	C	2171	LYS
3	C	2183	LEU
4	D	3038	ARG
4	D	3055	LEU
4	D	3064	LEU
4	D	3083	LEU
4	D	3089	ASP
4	D	3165	VAL
4	D	3203	SER
4	D	3220	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	98	ASN
1	A	355	ASN
2	B	1094	GLN
3	C	2102	GLN
4	D	3003	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

2 monosaccharides 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	GLC	E	1	5	11,11,12	0.60	0	15,15,17	0.70	0
5	FRU	E	2	5	11,12,12	0.55	0	10,18,18	0.85	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	GLC	E	1	5	-	0/2/19/22	0/1/1/1
5	FRU	E	2	5	-	0/5/24/24	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

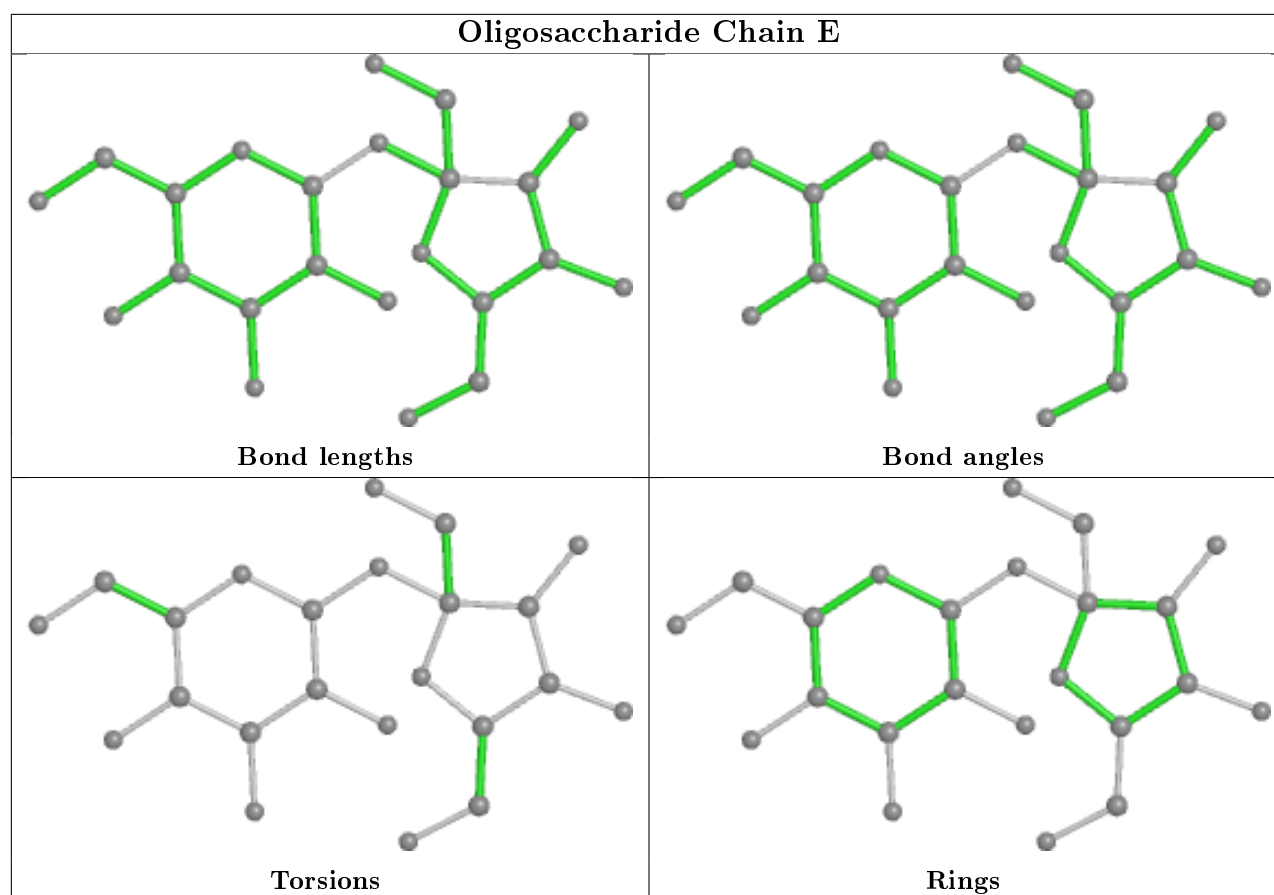
There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	2	FRU	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.





## 5.6 Ligand geometry [i](#)

13 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	$\# Z  > 2$
6	NAG	A	789	1	14,14,15	0.56	0	17,19,21	0.77	0
6	NAG	A	892	1	14,14,15	0.58	0	17,19,21	0.88	1 (5%)
6	NAG	A	741	1	14,14,15	0.56	0	17,19,21	0.96	1 (5%)
6	NAG	A	762	1	14,14,15	0.56	0	17,19,21	0.93	1 (5%)
6	NAG	A	697	1	14,14,15	0.45	0	17,19,21	0.88	0
6	NAG	A	795	1	14,14,15	0.52	0	17,19,21	0.81	0
6	NAG	A	839	1	14,14,15	0.73	0	17,19,21	1.32	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	A	730	1	14,14,15	0.62	0	17,19,21	0.98	1 (5%)
6	NAG	A	588	1	14,14,15	0.52	0	17,19,21	0.75	0
6	NAG	A	776	1	14,14,15	0.51	0	17,19,21	0.88	0
6	NAG	A	734	1	14,14,15	0.52	0	17,19,21	1.97	4 (23%)
6	NAG	A	886	1	14,14,15	0.50	0	17,19,21	0.62	0
6	NAG	A	948	1	14,14,15	0.49	0	17,19,21	0.82	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
6	NAG	A	789	1	-	1/6/23/26	0/1/1/1
6	NAG	A	892	1	-	2/6/23/26	0/1/1/1
6	NAG	A	741	1	-	4/6/23/26	0/1/1/1
6	NAG	A	762	1	-	0/6/23/26	0/1/1/1
6	NAG	A	697	1	-	4/6/23/26	0/1/1/1
6	NAG	A	795	1	-	2/6/23/26	0/1/1/1
6	NAG	A	839	1	-	3/6/23/26	0/1/1/1
6	NAG	A	730	1	-	4/6/23/26	0/1/1/1
6	NAG	A	588	1	-	5/6/23/26	0/1/1/1
6	NAG	A	776	1	-	0/6/23/26	0/1/1/1
6	NAG	A	734	1	-	5/6/23/26	0/1/1/1
6	NAG	A	886	1	-	0/6/23/26	0/1/1/1
6	NAG	A	948	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (9) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	734	NAG	C1-O5-C5	5.98	120.30	112.19
6	A	839	NAG	C4-C3-C2	4.20	117.18	111.02
6	A	734	NAG	C4-C3-C2	-3.30	106.19	111.02
6	A	734	NAG	O5-C1-C2	2.99	116.01	111.29
6	A	730	NAG	C4-C3-C2	2.62	114.86	111.02
6	A	734	NAG	C3-C4-C5	-2.38	106.00	110.24
6	A	741	NAG	C4-C3-C2	2.27	114.34	111.02
6	A	762	NAG	O5-C5-C6	2.18	110.62	107.20
6	A	892	NAG	O5-C5-C6	2.12	110.52	107.20

There are no chirality outliers.

All (30) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	741	NAG	C8-C7-N2-C2
6	A	741	NAG	O7-C7-N2-C2
6	A	697	NAG	C8-C7-N2-C2
6	A	697	NAG	O7-C7-N2-C2
6	A	839	NAG	C3-C2-N2-C7
6	A	839	NAG	C8-C7-N2-C2
6	A	839	NAG	O7-C7-N2-C2
6	A	730	NAG	C8-C7-N2-C2
6	A	730	NAG	O7-C7-N2-C2
6	A	734	NAG	C8-C7-N2-C2
6	A	734	NAG	O7-C7-N2-C2
6	A	697	NAG	O5-C5-C6-O6
6	A	730	NAG	O5-C5-C6-O6
6	A	697	NAG	C4-C5-C6-O6
6	A	730	NAG	C4-C5-C6-O6
6	A	892	NAG	O5-C5-C6-O6
6	A	734	NAG	O5-C5-C6-O6
6	A	741	NAG	O5-C5-C6-O6
6	A	795	NAG	C8-C7-N2-C2
6	A	795	NAG	O7-C7-N2-C2
6	A	588	NAG	C8-C7-N2-C2
6	A	892	NAG	C4-C5-C6-O6
6	A	734	NAG	C1-C2-N2-C7
6	A	741	NAG	C4-C5-C6-O6
6	A	734	NAG	C4-C5-C6-O6
6	A	588	NAG	O5-C5-C6-O6
6	A	588	NAG	C4-C5-C6-O6
6	A	588	NAG	O7-C7-N2-C2
6	A	789	NAG	O5-C5-C6-O6
6	A	588	NAG	C1-C2-N2-C7

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	762	NAG	1	0
6	A	795	NAG	1	0
6	A	734	NAG	1	0

## 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	305/317 (96%)	0.47	30 (9%) 7 7	14, 30, 65, 95	0
2	B	181/184 (98%)	0.27	10 (5%) 25 24	16, 35, 59, 70	0
3	C	214/214 (100%)	-0.09	4 (1%) 66 65	12, 27, 53, 64	0
4	D	223/229 (97%)	0.09	8 (3%) 42 42	13, 23, 55, 77	0
All	All	923/944 (97%)	0.21	52 (5%) 24 23	12, 29, 59, 95	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	410	GLY	7.1
1	A	194	GLY	5.9
1	A	128	GLY	5.2
1	A	462	ASN	5.0
1	A	129	ALA	4.9
1	A	398	SER	4.9
1	A	385	CYS	4.4
2	B	1105	ASP	4.3
4	D	3148	GLY	4.0
1	A	88	ASN	3.7
2	B	1001	LYS	3.5
1	A	126	CYS	3.5
1	A	84	VAL	3.3
1	A	195	SER	3.2
1	A	87	VAL	3.2
4	D	3229	LYS	3.2
1	A	205	CYS	3.1
1	A	463	ASN	3.0
1	A	196	CYS	3.0
1	A	228	CYS	3.0
1	A	83	GLU	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	461	SER	2.9
1	A	89	VAL	2.8
3	C	2196	CYS	2.8
1	A	127	VAL	2.8
4	D	3228	PRO	2.7
1	A	85	VAL	2.6
2	B	1181	LYS	2.6
2	B	1141	GLY	2.6
1	A	231	CYS	2.6
1	A	86	LEU	2.6
3	C	2001	ASP	2.5
1	A	356	ASN	2.5
2	B	1106	THR	2.5
3	C	2214	GLY	2.5
2	B	1125	SER	2.4
1	A	464	GLU	2.4
2	B	1088	ASP	2.3
4	D	3001	GLU	2.3
2	B	1180	GLN	2.3
2	B	1091	GLU	2.3
1	A	210	PHE	2.2
4	D	3149	GLY	2.2
4	D	3227	GLU	2.2
1	A	396	PHE	2.2
4	D	3205	GLY	2.2
2	B	1019	SER	2.1
3	C	2124	ASP	2.1
1	A	99	ASP	2.0
1	A	267	CYS	2.0
4	D	3220	THR	2.0
1	A	397	ASN	2.0

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

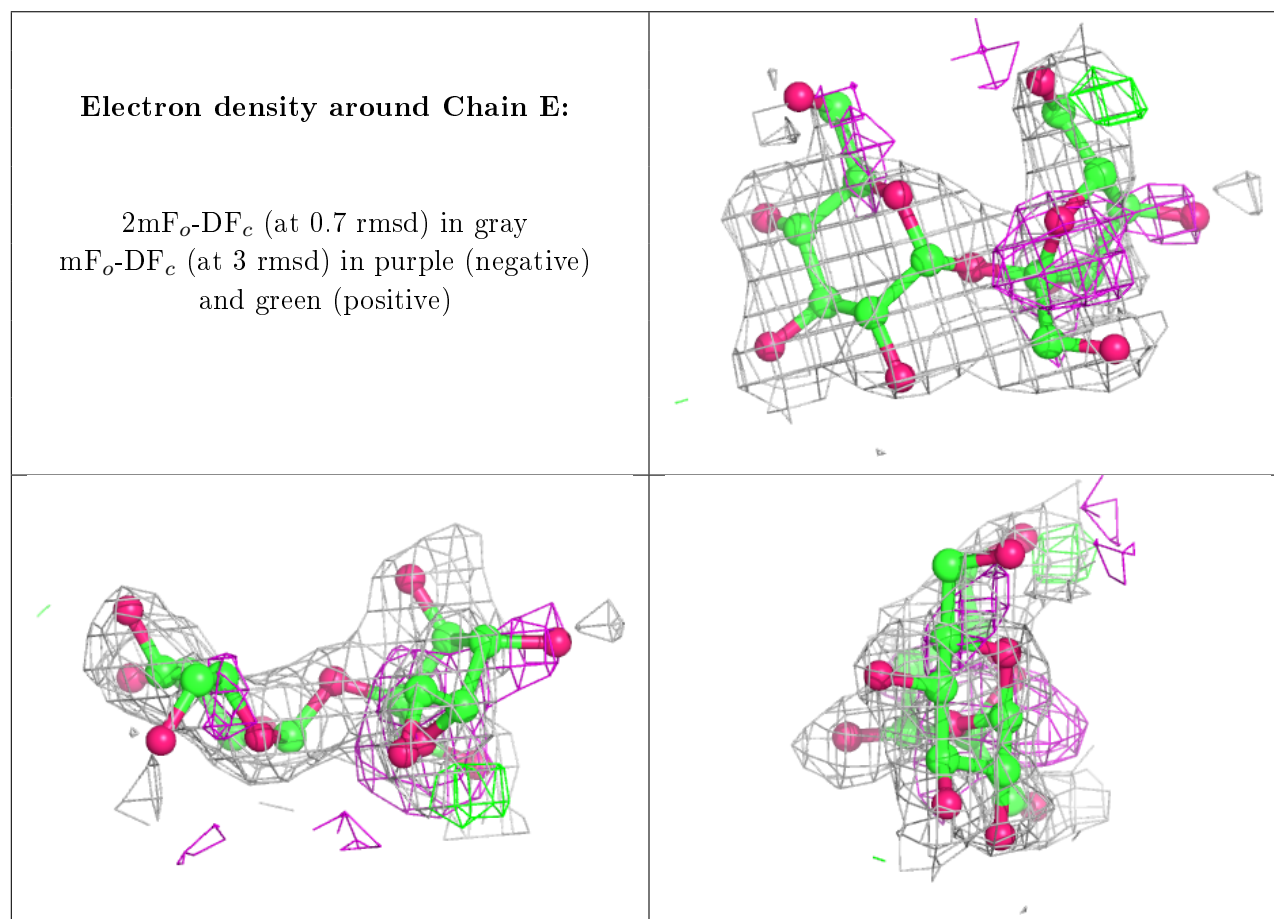
There are no non-standard protein/DNA/RNA residues in this entry.

## 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
5	FRU	E	2	12/12	0.68	0.39	54,55,55,56	0
5	GLC	E	1	11/12	0.84	0.28	54,54,54,55	0

The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



## 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. 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	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NAG	A	697	14/15	0.55	0.52	53,55,55,56	0
6	NAG	A	588	14/15	0.55	0.43	60,61,62,62	0
6	NAG	A	741	14/15	0.62	0.44	52,53,54,54	0
6	NAG	A	730	14/15	0.65	0.43	54,56,56,57	0
6	NAG	A	839	14/15	0.66	0.43	48,50,51,51	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NAG	A	734	14/15	0.71	0.33	49,50,51,51	0
6	NAG	A	892	14/15	0.79	0.34	42,44,45,45	0
6	NAG	A	948	14/15	0.81	0.21	40,42,43,43	0
6	NAG	A	795	14/15	0.89	0.15	31,33,34,34	0
6	NAG	A	886	14/15	0.90	0.12	37,38,39,39	0
6	NAG	A	762	14/15	0.91	0.15	28,32,34,35	0
6	NAG	A	776	14/15	0.92	0.12	36,38,39,39	0
6	NAG	A	789	14/15	0.92	0.19	38,39,40,40	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.