



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 20, 2023 – 12:40 PM EDT

PDB ID : 2NY4  
Title : HIV-1 gp120 Envelope Glycoprotein (K231C, T257S, E268C, 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

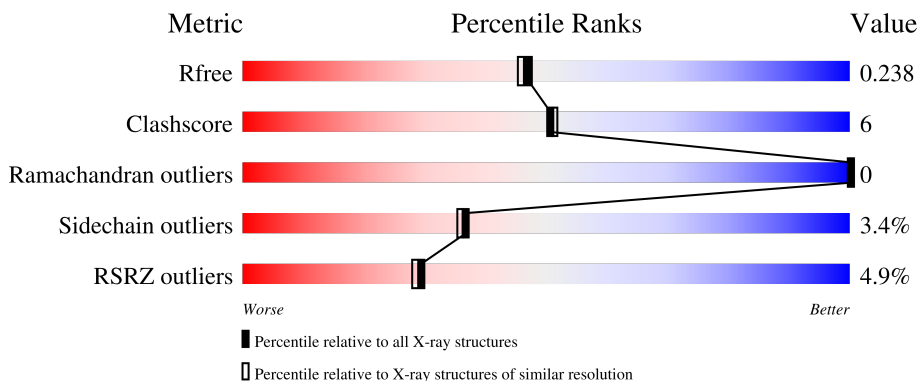
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

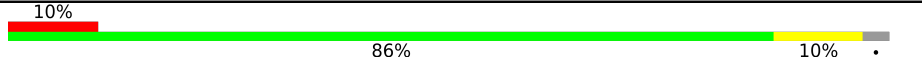

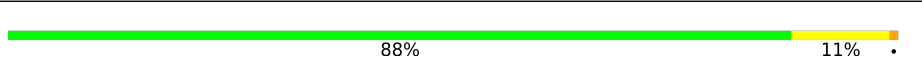
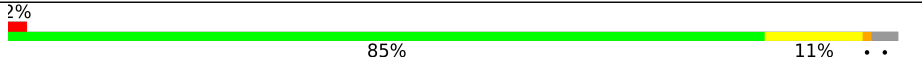
The reported resolution of this entry is 2.00 Å.

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 of 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	
2	B	184	
3	C	214	
4	D	229	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	E	2	

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 [i](#)

There are 8 unique types of molecules in this entry. The entry contains 7824 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
			Total	C	N	O	S			
1	A	306	2364	1484	412	446	22	0	0	0

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

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

There is a discrepancy between the modelled and reference sequences:

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

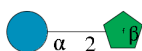
- Molecule 3 is a protein called ANTIBODY 17B, LIGHT CHAIN.

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	222	1677	1062	281	329	5	0	0	0

- 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:  $C_8H_{15}NO_6$ ).



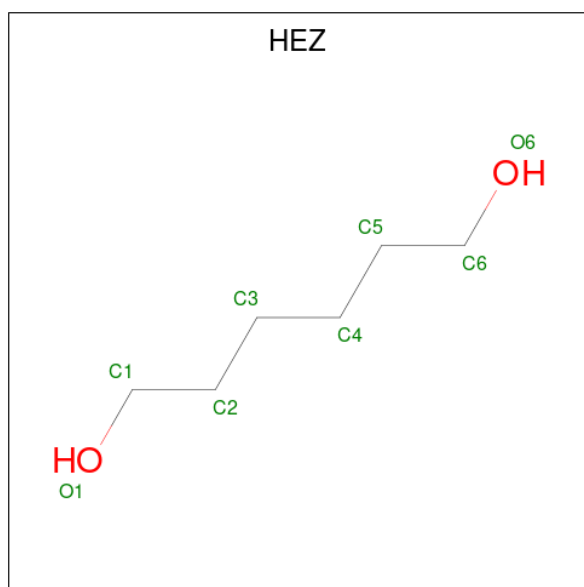
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	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 HEXANE-1,6-DIOL (three-letter code: HEZ) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total	C	O	0	0
			8	6	2		

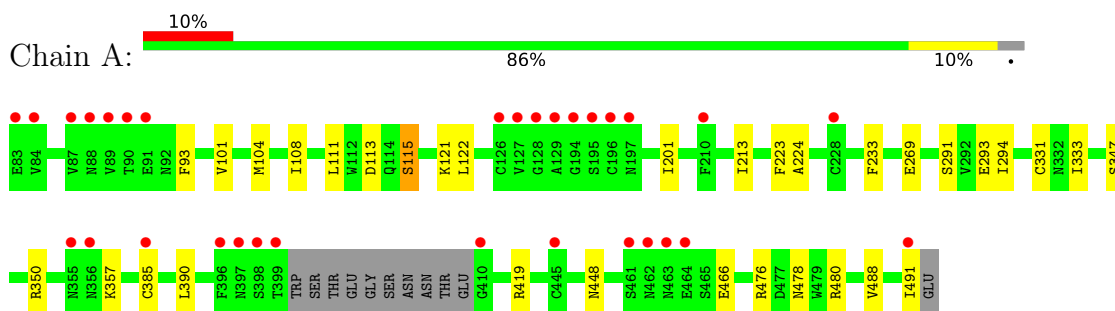
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	182	Total	O	0	0
			182	182		
8	B	78	Total	O	0	0
			78	78		
8	C	129	Total	O	0	0
			129	129		
8	D	123	Total	O	0	0
			123	123		

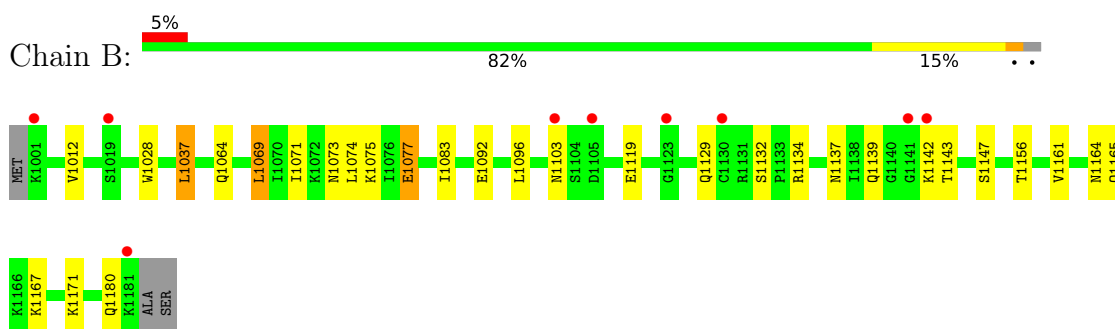
### 3 Residue-property plots [i](#)

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.

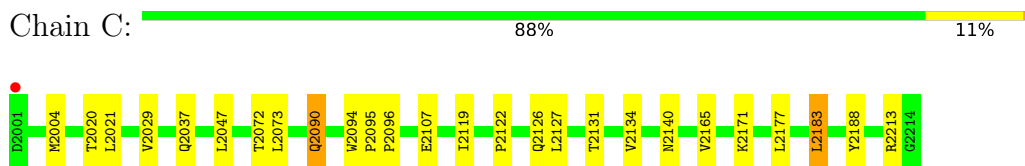
- Molecule 1: ENVELOPE GLYCOPROTEIN GP120



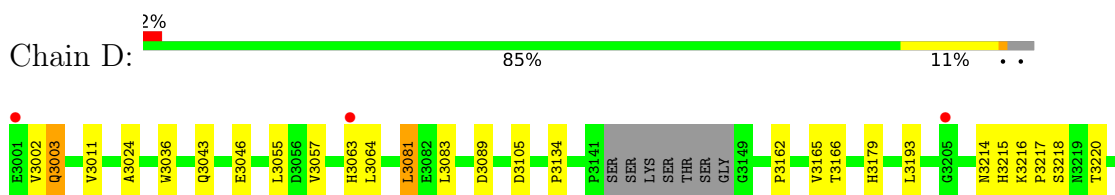
- Molecule 2: T-cell surface glycoprotein CD4



- Molecule 3: ANTIBODY 17B, LIGHT CHAIN



- Molecule 4: ANTIBODY 17B, HEAVY CHAIN



K3229

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

Chain E:



GLU1  
FRU2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.96Å 87.96Å 196.18Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.91 – 2.00 42.91 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.2 (42.91-2.00) 94.2 (42.91-2.00)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.47 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.202 , 0.241 0.201 , 0.238	Depositor DCC
$R_{free}$ test set	3989 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	25.9	Xtrriage
Anisotropy	0.027	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 42.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7824	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.33% 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: FRU, NAG, HEZ, GLC

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	0.29	0/2412	0.46	0/3273
2	B	0.29	0/1432	0.48	1/1930 (0.1%)
3	C	0.31	0/1683	0.51	0/2288
4	D	0.32	0/1716	0.51	0/2338
All	All	0.30	0/7243	0.49	1/9829 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1037	LEU	CA-CB-CG	5.65	128.28	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	2364	0	2309	27	0
2	B	1412	0	1441	14	0
3	C	1646	0	1590	17	0
4	D	1677	0	1644	24	0
5	E	23	0	21	3	0
6	A	182	0	169	2	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	8	0	14	4	0
8	A	182	0	0	8	0
8	B	78	0	0	1	0
8	C	129	0	0	0	0
8	D	123	0	0	4	0
All	All	7824	0	7188	81	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 (81) 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:3003:GLN:NE2	8:D:463:HOH:O	1.62	1.13
1:A:476:ARG:O	8:A:5120:HOH:O	1.90	0.90
4:D:3063:HIS:CE1	5:E:2:FRU:HO6	1.95	0.85
1:A:113:ASP:OD1	8:A:5091:HOH:O	1.99	0.80
1:A:101:VAL:HG11	8:A:5120:HOH:O	1.82	0.79
4:D:3215:HIS:HB3	4:D:3220:THR:CG2	2.14	0.78
4:D:3215:HIS:HB3	4:D:3220:THR:HG22	1.70	0.71
4:D:3063:HIS:CE1	5:E:2:FRU:O6	2.44	0.71
4:D:3046:GLU:OE2	4:D:3063:HIS:CE1	2.45	0.69
1:A:419:ARG:HD2	7:A:5000:HEZ:H32	1.75	0.68
4:D:3002:VAL:HG21	4:D:3024:ALA:HB1	1.83	0.61
2:B:1077:GLU:H	2:B:1077:GLU:CD	2.04	0.58
6:A:762:NAG:O4	8:A:5065:HOH:O	2.17	0.58
4:D:3011:VAL:HG21	4:D:3162:PRO:HG3	1.86	0.58
2:B:1075:LYS:HB3	2:B:1077:GLU:OE2	2.05	0.57
1:A:419:ARG:HG3	7:A:5000:HEZ:H51	1.87	0.57
1:A:333:ILE:HD12	1:A:390:LEU:HD21	1.87	0.57
3:C:2183:LEU:HD12	3:C:2183:LEU:N	2.20	0.57
3:C:2037:GLN:HB2	3:C:2047:LEU:HD11	1.88	0.56
1:A:294:ILE:HD12	1:A:333:ILE:HD11	1.88	0.56
4:D:3036:TRP:CE2	4:D:3081:LEU:HB2	2.42	0.55
3:C:2029:VAL:HG11	3:C:2090:GLN:CG	2.36	0.55
1:A:93:PHE:HB2	1:A:233:PHE:CZ	2.41	0.54
2:B:1129:GLN:HG3	2:B:1139:GLN:HG3	1.90	0.53
3:C:2004:MET:SD	3:C:2090:GLN:HB2	2.48	0.53
7:A:5000:HEZ:H42	8:A:5108:HOH:O	2.07	0.53
1:A:357:LYS:HD3	1:A:466:GLU:HG2	1.91	0.53
2:B:1071:ILE:HG23	8:B:281:HOH:O	2.09	0.53

Continued on next page...

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:2094:TRP:HA	3:C:2095:PRO:C	2.29	0.52
3:C:2029:VAL:HG11	3:C:2090:GLN:HG2	1.91	0.52
4:D:3179:HIS:HD2	8:D:347:HOH:O	1.90	0.52
4:D:3166:THR:CG2	4:D:3214:ASN:HB3	2.40	0.51
1:A:419:ARG:NH2	8:D:138:HOH:O	2.43	0.51
1:A:223:PHE:HB3	1:A:488:VAL:HG13	1.93	0.50
3:C:2165:VAL:HG22	3:C:2177:LEU:HD12	1.94	0.50
3:C:2140:ASN:HD21	4:D:3179:HIS:HE1	1.60	0.50
1:A:93:PHE:HB2	1:A:233:PHE:HZ	1.77	0.50
1:A:347:SER:HA	1:A:350:ARG:NH2	2.26	0.50
4:D:3134:PRO:HD2	4:D:3220:THR:HG21	1.93	0.49
1:A:101:VAL:CG1	8:A:5120:HOH:O	2.47	0.49
3:C:2140:ASN:HD21	4:D:3179:HIS:CE1	2.31	0.49
1:A:419:ARG:CD	7:A:5000:HEZ:H32	2.41	0.49
4:D:3229:LYS:C	4:D:3229:LYS:HD3	2.33	0.48
1:A:478:ASN:HD22	1:A:478:ASN:H	1.60	0.48
1:A:293:GLU:OE2	6:A:795:NAG:H3	2.14	0.48
1:A:331:CYS:SG	1:A:385:CYS:SG	3.12	0.47
1:A:104:MET:O	1:A:108:ILE:HG12	2.14	0.47
2:B:1180:GLN:OE1	2:B:1180:GLN:N	2.36	0.47
4:D:3002:VAL:CG2	4:D:3024:ALA:HB1	2.45	0.46
2:B:1137:ASN:HB3	2:B:1139:GLN:HE22	1.78	0.46
4:D:3043:GLN:HG3	8:D:292:HOH:O	2.15	0.46
3:C:2183:LEU:HD12	3:C:2183:LEU:H	1.81	0.46
1:A:121:LYS:HG3	1:A:201:ILE:HB	1.96	0.45
1:A:480:ARG:CZ	8:A:5120:HOH:O	2.65	0.45
3:C:2122:PRO:HD3	3:C:2134:VAL:HG22	1.98	0.45
3:C:2188:TYR:CZ	3:C:2213:ARG:HG3	2.51	0.45
4:D:3063:HIS:NE2	5:E:2:FRU:O6	2.37	0.45
2:B:1132:SER:HG	2:B:1134:ARG:HG2	1.81	0.45
3:C:2094:TRP:CE3	3:C:2096:PRO:HD3	2.51	0.45
1:A:122:LEU:HD21	4:D:3057:VAL:HG21	1.99	0.44
2:B:1028:TRP:CE2	2:B:1069:LEU:HB2	2.52	0.44
4:D:3193:LEU:C	4:D:3193:LEU:HD12	2.39	0.44
3:C:2029:VAL:HG11	3:C:2090:GLN:HG3	2.00	0.43
2:B:1083:ILE:HG12	2:B:1092:GLU:HG3	2.00	0.43
2:B:1012:VAL:HB	2:B:1074:LEU:HD11	2.00	0.43
4:D:3216:LYS:N	4:D:3217:PRO:CD	2.82	0.43
2:B:1156:THR:HG21	2:B:1171:LYS:HD3	2.01	0.43
1:A:111:LEU:O	1:A:115:SER:HB2	2.19	0.42
2:B:1129:GLN:HE21	2:B:1139:GLN:HE21	1.67	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:1161:VAL:O	2:B:1167:LYS:HA	2.20	0.42
4:D:3105:ASP:OD1	4:D:3105:ASP:N	2.45	0.42
1:A:269:GLU:HB2	8:A:5175:HOH:O	2.19	0.41
3:C:2020:THR:HG23	3:C:2072:THR:CG2	2.50	0.41
4:D:3218:SER:OG	4:D:3220:THR:HG22	2.20	0.41
1:A:224:ALA:HB2	1:A:491:ILE:CG2	2.50	0.41
1:A:291:SER:HB2	1:A:448:ASN:HB3	2.02	0.41
2:B:1164:ASN:O	2:B:1165:GLN:HB2	2.21	0.41
1:A:478:ASN:HD22	1:A:478:ASN:N	2.18	0.41
3:C:2126:GLN:HG2	3:C:2131:THR:O	2.20	0.41
3:C:2021:LEU:N	3:C:2021:LEU:HD12	2.36	0.40
4:D:3166:THR:HG23	4:D:3214:ASN:HB3	2.03	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	302/317 (95%)	289 (96%)	13 (4%)	0	100	100
2	B	179/184 (97%)	173 (97%)	6 (3%)	0	100	100
3	C	212/214 (99%)	202 (95%)	10 (5%)	0	100	100
4	D	218/229 (95%)	215 (99%)	3 (1%)	0	100	100
All	All	911/944 (96%)	879 (96%)	32 (4%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar

resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	270/280 (96%)	268 (99%)	2 (1%)	84	88
2	B	164/166 (99%)	153 (93%)	11 (7%)	16	11
3	C	184/184 (100%)	177 (96%)	7 (4%)	33	31
4	D	187/193 (97%)	180 (96%)	7 (4%)	34	32
All	All	805/823 (98%)	778 (97%)	27 (3%)	37	36

All (27) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	115	SER
1	A	213	ILE
2	B	1037	LEU
2	B	1064	GLN
2	B	1069	LEU
2	B	1073	ASN
2	B	1077	GLU
2	B	1096	LEU
2	B	1103	ASN
2	B	1119	GLU
2	B	1142	LYS
2	B	1143	THR
2	B	1147	SER
3	C	2073	LEU
3	C	2090	GLN
3	C	2107	GLU
3	C	2119	ILE
3	C	2127	LEU
3	C	2171	LYS
3	C	2183	LEU
4	D	3003	GLN
4	D	3055	LEU
4	D	3064	LEU
4	D	3081	LEU
4	D	3083	LEU
4	D	3089	ASP
4	D	3165	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	203	GLN
1	A	355	ASN
1	A	377	ASN
1	A	478	ASN
2	B	1040	GLN
2	B	1066	ASN
2	B	1073	ASN
2	B	1103	ASN
2	B	1139	GLN
2	B	1163	GLN
3	C	2140	ASN
4	D	3065	GLN
4	D	3179	HIS
4	D	3207	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

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	0.61	0	15,15,17	0.70	0
5	FRU	E	2	5	11,12,12	0.51	0	10,18,18	0.91	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	-	1/2/19/22	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.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	1	GLC	C4-C5-C6-O6

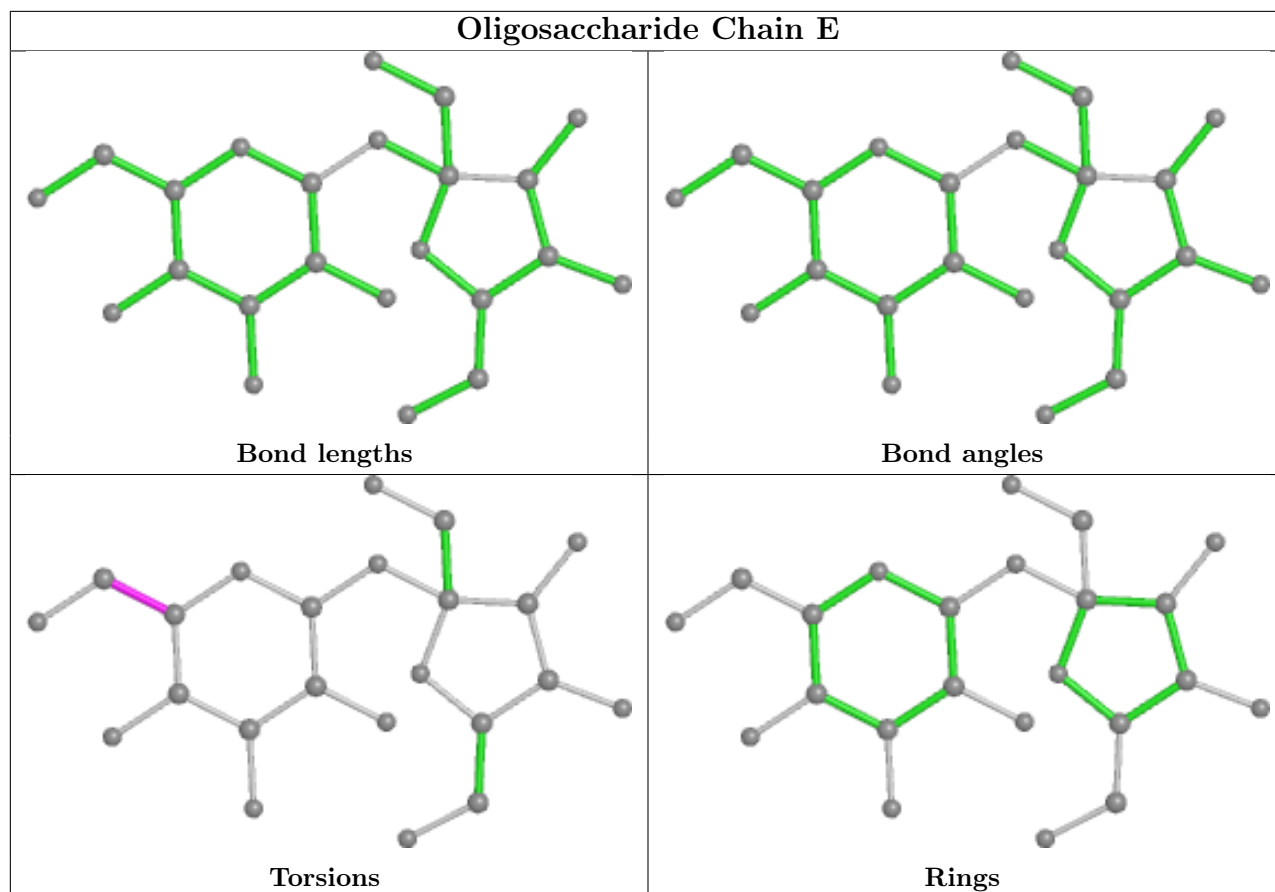
There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	E	2	FRU	3	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](#)

14 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	697	1	14,14,15	0.48	0	17,19,21	0.92	1 (5%)
6	NAG	A	588	1	14,14,15	0.52	0	17,19,21	0.76	0
7	HEZ	A	5000	-	7,7,7	0.29	0	6,6,6	0.45	0
6	NAG	A	741	1	14,14,15	0.61	0	17,19,21	1.04	1 (5%)
6	NAG	A	776	1	14,14,15	0.57	0	17,19,21	0.75	0
6	NAG	A	762	1	14,14,15	0.54	0	17,19,21	1.13	2 (11%)
6	NAG	A	886	1	14,14,15	0.51	0	17,19,21	0.69	0

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	0.60	0	17,19,21	0.67	0
6	NAG	A	839	1	14,14,15	0.59	0	17,19,21	0.81	0
6	NAG	A	730	1	14,14,15	0.54	0	17,19,21	0.97	1 (5%)
6	NAG	A	734	1	14,14,15	0.49	0	17,19,21	1.11	1 (5%)
6	NAG	A	892	1	14,14,15	0.50	0	17,19,21	0.67	0
6	NAG	A	795	1	14,14,15	0.50	0	17,19,21	0.88	1 (5%)
6	NAG	A	948	1	14,14,15	0.54	0	17,19,21	0.72	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	697	1	-	5/6/23/26	0/1/1/1
6	NAG	A	588	1	-	3/6/23/26	0/1/1/1
7	HEZ	A	5000	-	-	4/5/5/5	-
6	NAG	A	741	1	-	6/6/23/26	0/1/1/1
6	NAG	A	776	1	-	0/6/23/26	0/1/1/1
6	NAG	A	762	1	-	0/6/23/26	0/1/1/1
6	NAG	A	886	1	-	2/6/23/26	0/1/1/1
6	NAG	A	789	1	-	0/6/23/26	0/1/1/1
6	NAG	A	839	1	-	5/6/23/26	0/1/1/1
6	NAG	A	730	1	-	4/6/23/26	0/1/1/1
6	NAG	A	734	1	-	2/6/23/26	0/1/1/1
6	NAG	A	892	1	-	1/6/23/26	0/1/1/1
6	NAG	A	795	1	-	0/6/23/26	0/1/1/1
6	NAG	A	948	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	734	NAG	O5-C5-C6	3.13	112.11	107.20
6	A	762	NAG	C1-O5-C5	3.08	116.36	112.19
6	A	741	NAG	C4-C3-C2	2.90	115.26	111.02
6	A	730	NAG	O5-C5-C6	2.61	111.30	107.20
6	A	762	NAG	C6-C5-C4	-2.36	107.49	113.00
6	A	795	NAG	O5-C5-C6	2.33	110.86	107.20

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	697	NAG	O5-C5-C6	2.23	110.71	107.20

There are no chirality outliers.

All (34) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	A	697	NAG	C8-C7-N2-C2
6	A	697	NAG	O7-C7-N2-C2
6	A	730	NAG	C8-C7-N2-C2
6	A	730	NAG	O7-C7-N2-C2
6	A	741	NAG	C8-C7-N2-C2
6	A	741	NAG	O7-C7-N2-C2
6	A	948	NAG	O5-C5-C6-O6
6	A	730	NAG	O5-C5-C6-O6
6	A	734	NAG	O5-C5-C6-O6
6	A	839	NAG	C8-C7-N2-C2
6	A	730	NAG	C4-C5-C6-O6
6	A	948	NAG	C4-C5-C6-O6
6	A	839	NAG	O7-C7-N2-C2
6	A	839	NAG	O5-C5-C6-O6
6	A	839	NAG	C4-C5-C6-O6
6	A	734	NAG	C4-C5-C6-O6
6	A	886	NAG	C8-C7-N2-C2
6	A	886	NAG	O7-C7-N2-C2
6	A	588	NAG	C1-C2-N2-C7
6	A	741	NAG	C4-C5-C6-O6
6	A	839	NAG	C1-C2-N2-C7
6	A	741	NAG	O5-C5-C6-O6
7	A	5000	HEZ	C4-C5-C6-O6
7	A	5000	HEZ	C2-C3-C4-C5
7	A	5000	HEZ	O1-C1-C2-C3
6	A	741	NAG	C1-C2-N2-C7
6	A	892	NAG	O5-C5-C6-O6
6	A	697	NAG	C4-C5-C6-O6
6	A	697	NAG	C3-C2-N2-C7
6	A	588	NAG	C8-C7-N2-C2
7	A	5000	HEZ	C3-C4-C5-C6
6	A	741	NAG	C3-C2-N2-C7
6	A	588	NAG	O7-C7-N2-C2
6	A	697	NAG	O5-C5-C6-O6

There are no ring outliers.

3 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	5000	HEZ	4	0
6	A	762	NAG	1	0
6	A	795	NAG	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data

### 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	306/317 (96%)	0.52	31 (10%) <b>7</b> <b>6</b>	15, 29, 60, 90	0
2	B	181/184 (98%)	0.38	9 (4%) 28 28	16, 36, 58, 72	0
3	C	214/214 (100%)	-0.02	1 (0%) 91 90	14, 27, 51, 61	0
4	D	222/229 (96%)	0.12	4 (1%) 68 66	12, 23, 53, 73	0
All	All	923/944 (97%)	0.27	45 (4%) 29 28	12, 29, 57, 90	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	398	SER	7.4
1	A	399	THR	7.1
1	A	462	ASN	6.8
1	A	195	SER	5.9
1	A	127	VAL	5.6
1	A	129	ALA	5.2
1	A	87	VAL	5.1
2	B	1001	LYS	4.6
1	A	88	ASN	4.2
1	A	410	GLY	4.2
2	B	1181	LYS	3.8
1	A	194	GLY	3.7
4	D	3001	GLU	3.5
1	A	197	ASN	3.4
3	C	2001	ASP	3.4
1	A	128	GLY	3.4
2	B	1130	CYS	3.4
1	A	464	GLU	3.3
4	D	3229	LYS	3.1
4	D	3063	HIS	3.1
2	B	1103	ASN	3.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	356	ASN	3.0
4	D	3205	GLY	2.9
2	B	1105	ASP	2.9
1	A	396	PHE	2.8
1	A	89	VAL	2.8
1	A	196	CYS	2.7
1	A	90	THR	2.7
1	A	463	ASN	2.7
2	B	1123	GLY	2.6
1	A	210	PHE	2.6
1	A	491	ILE	2.5
1	A	397	ASN	2.5
2	B	1141	GLY	2.5
1	A	461	SER	2.5
2	B	1019	SER	2.4
1	A	84	VAL	2.3
1	A	91	GLU	2.3
1	A	83	GLU	2.2
1	A	355	ASN	2.2
1	A	445	CYS	2.2
1	A	385	CYS	2.2
2	B	1142	LYS	2.2
1	A	228	CYS	2.1
1	A	126	CYS	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.

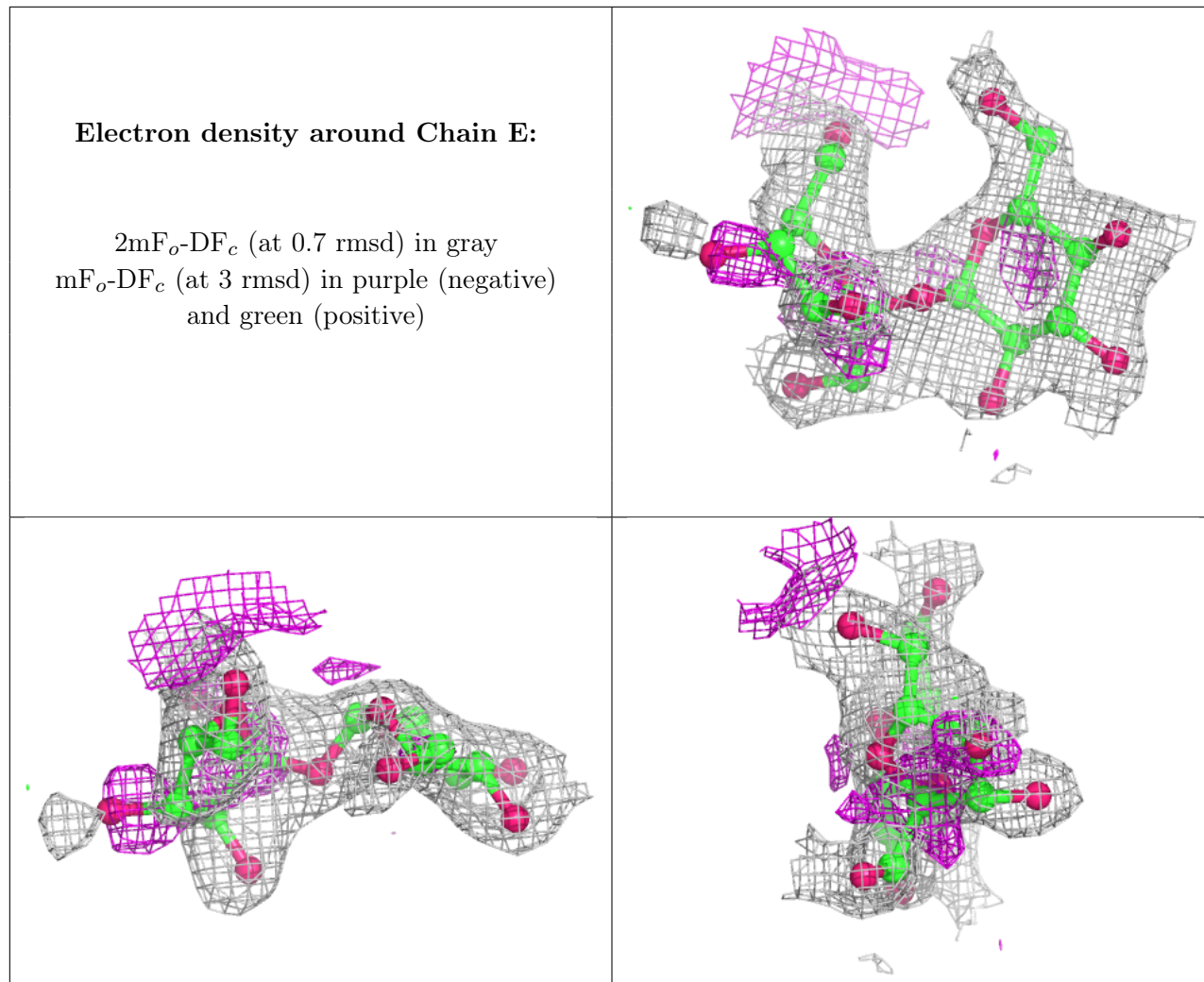
## 6.3 Carbohydrates [i](#)

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(Å <sup>2</sup> )	Q<0.9
5	FRU	E	2	12/12	0.71	0.29	44,46,47,48	0
5	GLC	E	1	11/12	0.84	0.23	45,45,46,46	0

The following is a graphical depiction of the model fit to experimental electron density for oligosac-

charide. Each fit is shown from different orientation to approximate a three-dimensional view.



## 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. 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(Å <sup>2</sup> )	Q<0.9
6	NAG	A	697	14/15	0.54	0.53	51,54,55,55	0
6	NAG	A	588	14/15	0.58	0.52	56,58,59,59	0
6	NAG	A	839	14/15	0.63	0.43	50,54,56,56	0
6	NAG	A	892	14/15	0.66	0.36	44,47,49,49	0
6	NAG	A	730	14/15	0.67	0.41	50,54,55,55	0
6	NAG	A	741	14/15	0.70	0.55	50,54,56,56	0
6	NAG	A	948	14/15	0.79	0.22	41,44,46,47	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	0.83	0.20	46,49,50,51	0
7	HEZ	A	5000	8/8	0.83	0.33	45,46,47,47	0
6	NAG	A	886	14/15	0.88	0.15	36,39,40,40	0
6	NAG	A	795	14/15	0.91	0.16	31,32,34,35	0
6	NAG	A	789	14/15	0.92	0.14	34,36,37,37	0
6	NAG	A	776	14/15	0.93	0.11	37,39,42,43	0
6	NAG	A	762	14/15	0.93	0.13	28,33,35,36	0

## 6.5 Other polymers [\(i\)](#)

There are no such residues in this entry.