



wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 25, 2017 – 08:32 PM EDT

PDB ID : 1G9M
Title : HIV-1 HXBC2 GP120 ENVELOPE GLYCOPROTEIN COMPLEXED WITH CD4 AND INDUCED NEUTRALIZING ANTIBODY 17B
Authors : Kwong, P.D.; Wyatt, R.; Majeed, S.; Robinson, J.; Sweet, R.W.; Sodroski, J.; Hendrickson, W.A.
Deposited on : unknown
Resolution : 2.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at 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	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20029824

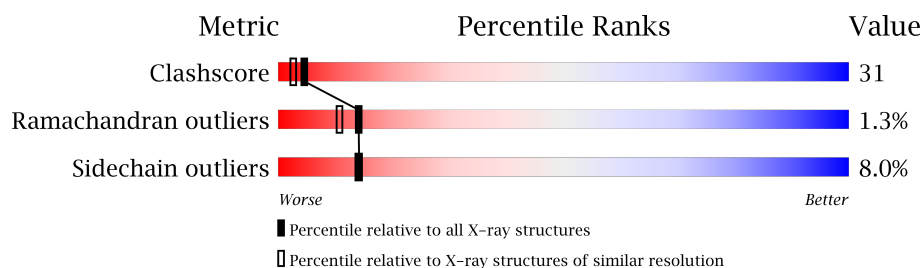
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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))
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)

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 ≥ 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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	G	321	
2	C	185	
3	L	214	
4	H	229	

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
5	NDG	G	730	-	-	X	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 8322 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	G	305	Total	C	N	O	S	0	0	0
			2361	1480	411	450	20			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	79	GLY	-	cloning artifact	UNP P04578
G	80	ALA	-	cloning artifact	UNP P04578
G	81	ARG	-	cloning artifact	UNP P04578
G	82	SER	-	cloning artifact	UNP P04578
G	128	GLY	-	SEE REMARK 999	UNP P04578
G	129	ALA	-	SEE REMARK 999	UNP P04578
G	194	GLY	-	SEE REMARK 999	UNP P04578
G	298	GLY	-	SEE REMARK 999	UNP P04578
G	299	ALA	-	SEE REMARK 999	UNP P04578
G	329	GLY	-	SEE REMARK 999	UNP P04578

- Molecule 2 is a protein called T-CELL SURFACE GLYCOPROTEIN CD4.

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

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	184	ASN	SER	ENGINEERED	UNP P01730
C	185	THR	ILE	ENGINEERED	UNP P01730

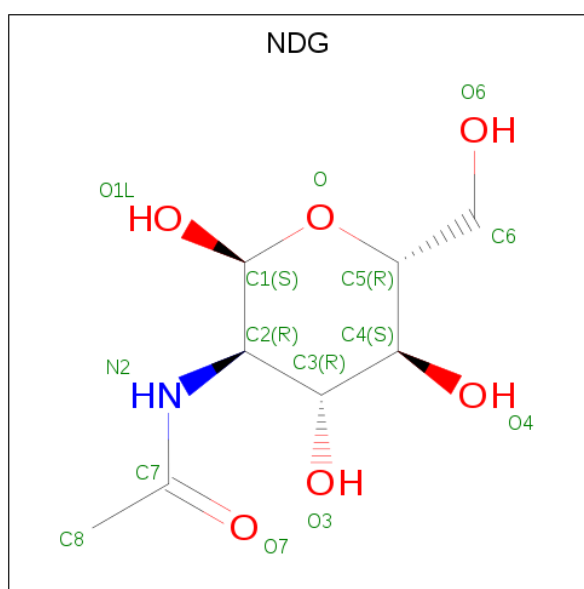
- Molecule 3 is a protein called ANTIBODY 17B, LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	214	Total	C	N	O	S	0	0	0
			1650	1032	283	331	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	229	Total	C	N	O	S	0	0	0
			1726	1090	294	337	5			

- Molecule 5 is 2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE (three-letter code: NDG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 6 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



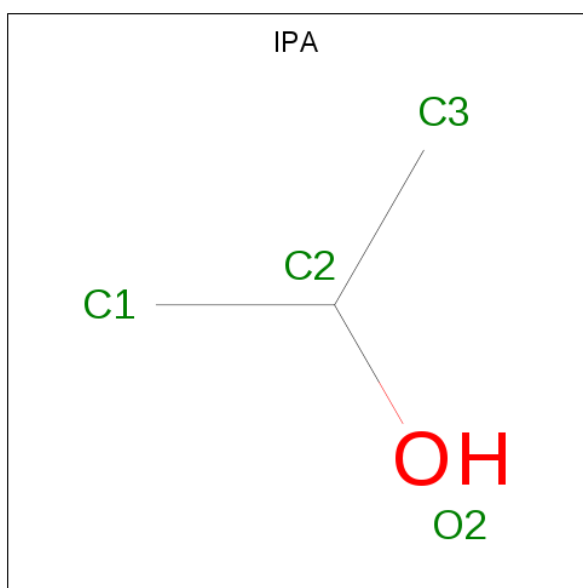
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1	Total	C	N	O	0	0
			14	8	1	5		
6	G	1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

- Molecule 7 is ALPHA-L-FUCOSE (three-letter code: FUC) (formula: C₆H₁₂O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	G	1	Total	C	O	0	0
			10	6	4		
7	G	1	Total	C	O	0	0
			10	6	4		

- Molecule 8 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	G	1	Total	C	O	0	0
			4	3	1		

- Molecule 9 is water.

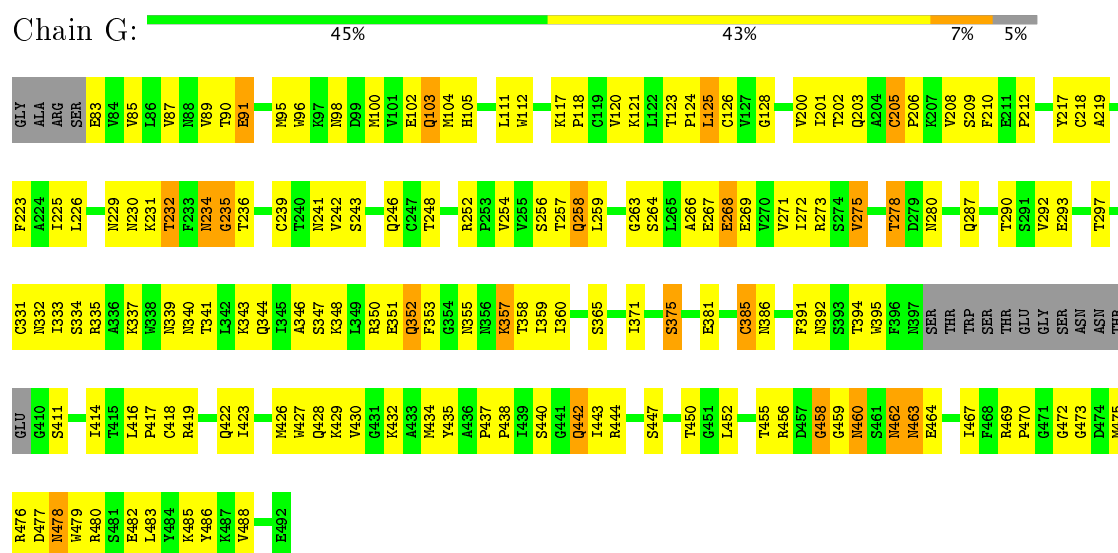
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	G	334	Total 334	O 334	0	0
9	C	157	Total 157	O 157	0	0
9	L	219	Total 219	O 219	0	0
9	H	243	Total 243	O 243	0	0

3 Residue-property plots [i](#)

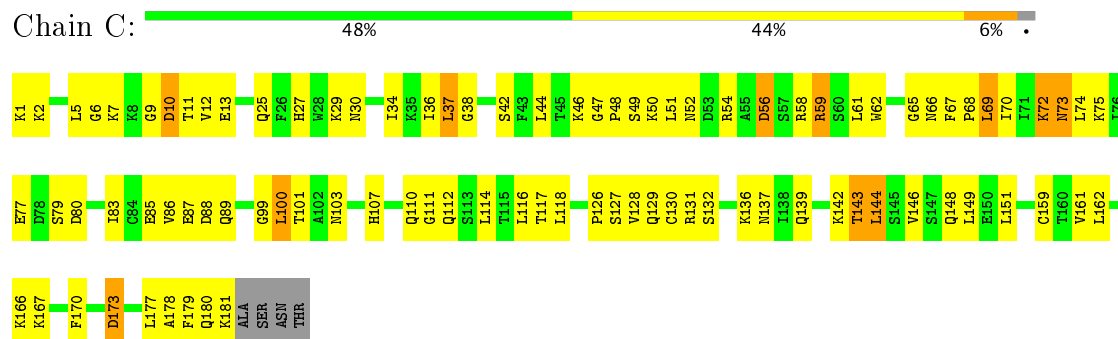
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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.

Note EDS was not executed.

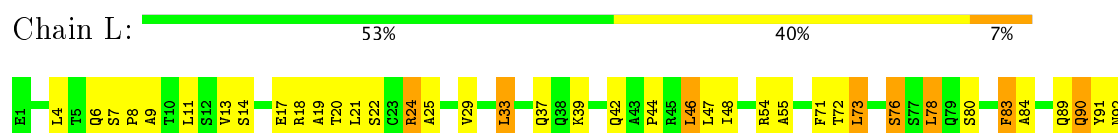
• Molecule 1: ENVELOPE GLYCOPROTEIN GP120

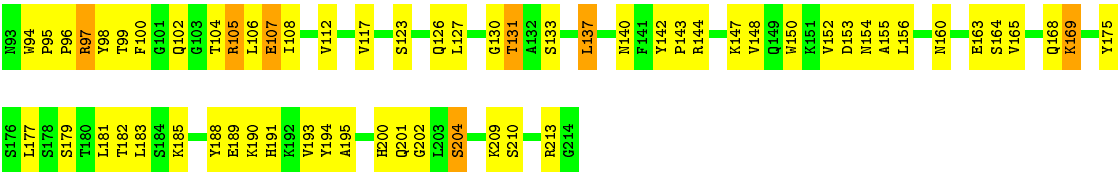


• Molecule 2: T-CELL SURFACE GLYCOPROTEIN CD4

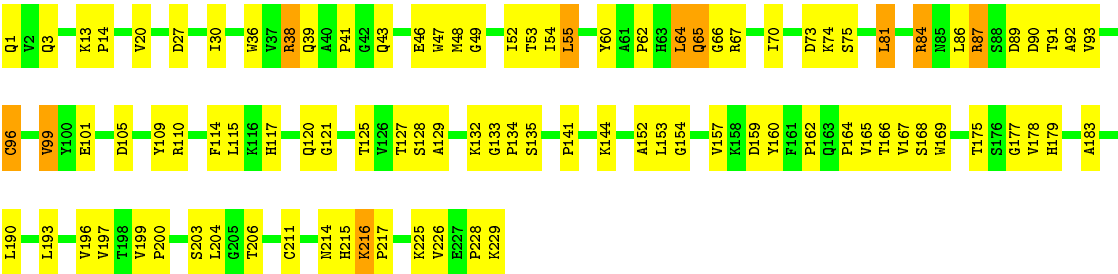


• Molecule 3: ANTIBODY 17B, LIGHT CHAIN





● Molecule 4: ANTIBODY 17B, HEAVY CHAIN



4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	71.25Å 88.11Å 196.54Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.20	Depositor
% Data completeness (in resolution range)	87.7 (20.00-2.20)	Depositor
R_{merge}	0.14	Depositor
R_{sym}	0.14	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.268 , 0.330	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8322	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

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: IPA, NAG, NDG, FUC

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	G	0.38	0/2407	0.59	0/3262
2	C	0.37	0/1432	0.56	0/1930
3	L	0.45	0/1687	0.63	0/2292
4	H	0.45	0/1766	0.65	0/2405
All	All	0.41	0/7292	0.61	0/9889

There are no bond length outliers.

There are no bond angle outliers.

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	G	2361	0	2301	163	0
2	C	1412	0	1444	92	0
3	L	1650	0	1603	99	0
4	H	1726	0	1708	102	0
5	G	42	0	39	7	0
6	G	154	0	141	17	0
7	G	20	0	20	0	0
8	G	4	0	8	0	0
9	C	157	0	0	29	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	G	334	0	0	43	0
9	H	243	0	0	25	0
9	L	219	0	0	18	0
All	All	8322	0	7264	456	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 31.

The worst 5 of 456 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:G:100:MET:HE1	1:G:486:TYR:HB2	1.47	0.95
4:H:84:ARG:HB3	9:H:281:HOH:O	1.64	0.95
3:L:6:GLN:HB2	3:L:102:GLN:HE22	1.29	0.95
1:G:347:SER:HA	1:G:350:ARG:HE	1.34	0.92
4:H:39:GLN:HB3	9:H:466:HOH:O	1.69	0.92

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	G	301/321 (94%)	261 (87%)	34 (11%)	6 (2%)	9	5
2	C	179/185 (97%)	154 (86%)	22 (12%)	3 (2%)	11	7
3	L	212/214 (99%)	192 (91%)	17 (8%)	3 (1%)	13	10
4	H	227/229 (99%)	208 (92%)	19 (8%)	0	100	100
All	All	919/949 (97%)	815 (89%)	92 (10%)	12 (1%)	14	11

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	268	GLU
3	L	76	SER
1	G	235	GLY
2	C	107	HIS
1	G	258	GLN

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	G	270/283 (95%)	247 (92%)	23 (8%)	12	12
2	C	164/167 (98%)	150 (92%)	14 (8%)	12	12
3	L	184/184 (100%)	169 (92%)	15 (8%)	13	13
4	H	193/193 (100%)	180 (93%)	13 (7%)	19	21
All	All	811/827 (98%)	746 (92%)	65 (8%)	14	14

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	72	LYS
2	C	173	ASP
4	H	96	CYS
2	C	73	ASN
2	C	143	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	73	ASN
2	C	103	ASN
3	L	149	GLN
1	G	478	ASN
2	C	40	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

17 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	$\# Z > 2$
8	IPA	G	1000	-	3,3,3	0.66	0	3,3,3	0.24	0
5	NDG	G	588	1	14,14,15	0.50	0	15,19,21	0.75	1 (6%)
6	NAG	G	697	1	14,14,15	0.55	0	15,19,21	0.72	1 (6%)
5	NDG	G	730	1	14,14,15	0.54	0	15,19,21	0.76	0
6	NAG	G	734	1	14,14,15	0.53	0	15,19,21	0.67	0
5	NDG	G	741	1	14,14,15	0.49	0	15,19,21	0.77	1 (6%)
6	NAG	G	762	1	14,14,15	0.49	0	15,19,21	0.86	0
6	NAG	G	776	1	14,14,15	0.60	0	15,19,21	0.89	1 (6%)
6	NAG	G	789	1	14,14,15	0.57	0	15,19,21	0.66	0
6	NAG	G	795	1,7	14,14,15	0.59	0	15,19,21	0.78	0
7	FUC	G	796	6	9,10,11	0.73	0	13,14,16	0.60	0
6	NAG	G	839	1	14,14,15	0.62	0	15,19,21	0.97	0
6	NAG	G	886	1	14,14,15	0.51	0	15,19,21	0.73	1 (6%)
6	NAG	G	892	1	14,14,15	0.48	0	15,19,21	0.78	1 (6%)
6	NAG	G	948	1	14,14,15	0.57	0	15,19,21	0.75	0
6	NAG	G	963	1,7	14,14,15	0.66	0	15,19,21	0.72	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
7	FUC	G	964	6	9,10,11	0.55	0	13,14,16	0.85	1 (7%)

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
8	IPA	G	1000	-	-	0/0/0/0	0/0/0/0
5	NDG	G	588	1	-	0/6/23/26	0/1/1/1
6	NAG	G	697	1	-	0/6/23/26	0/1/1/1
5	NDG	G	730	1	-	0/6/23/26	0/1/1/1
6	NAG	G	734	1	-	1/6/23/26	0/1/1/1
5	NDG	G	741	1	-	0/6/23/26	0/1/1/1
6	NAG	G	762	1	-	0/6/23/26	0/1/1/1
6	NAG	G	776	1	-	0/6/23/26	0/1/1/1
6	NAG	G	789	1	-	0/6/23/26	0/1/1/1
6	NAG	G	795	1,7	-	1/6/23/26	0/1/1/1
7	FUC	G	796	6	-	0/0/17/20	0/1/1/1
6	NAG	G	839	1	-	0/6/23/26	0/1/1/1
6	NAG	G	886	1	-	0/6/23/26	0/1/1/1
6	NAG	G	892	1	-	0/6/23/26	0/1/1/1
6	NAG	G	948	1	-	0/6/23/26	0/1/1/1
6	NAG	G	963	1,7	-	0/6/23/26	0/1/1/1
7	FUC	G	964	6	-	0/0/17/20	0/1/1/1

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	776	NAG	C2-N2-C7	-2.72	118.97	122.94
5	G	741	NDG	C2-N2-C7	-2.17	119.78	122.94
6	G	892	NAG	C2-N2-C7	-2.11	119.86	122.94
5	G	588	NDG	C2-N2-C7	-2.11	119.87	122.94
6	G	886	NAG	C2-N2-C7	-2.09	119.89	122.94

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	G	795	NAG	O7-C7-N2-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
6	G	734	NAG	O7-C7-N2-C2

There are no ring outliers.

10 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	730	NDG	7	0
6	G	734	NAG	3	0
6	G	762	NAG	3	0
6	G	776	NAG	1	0
6	G	789	NAG	3	0
6	G	795	NAG	2	0
6	G	839	NAG	1	0
6	G	886	NAG	1	0
6	G	892	NAG	2	0
6	G	963	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 ⓘ

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS was not executed - this section is therefore empty.

6.4 Ligands ⓘ

EDS was not executed - this section is therefore empty.

6.5 Other polymers ⓘ

EDS was not executed - this section is therefore empty.