



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2020 – 09:35 AM BST

PDB ID : 6V80  
Title : Crystal structure of human CD1d presenting alpha-Galactosylceramide in complex with NKT12 TCR and VHH nanobody 1D12  
Authors : Shahine, A.; Rossjohn, J.  
Deposited on : 2019-12-10  
Resolution : 3.53 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](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:

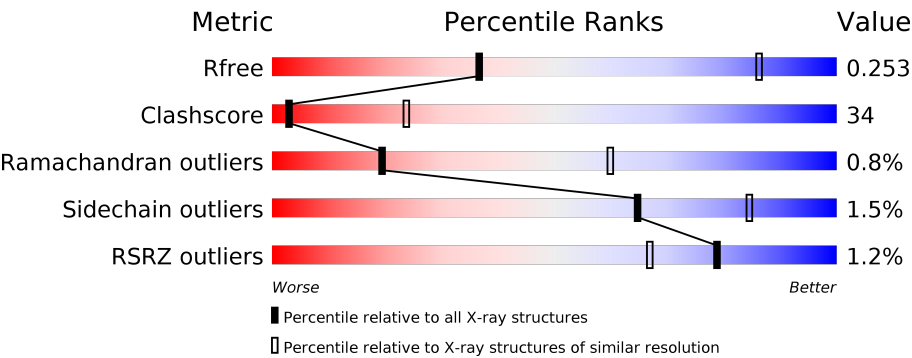
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.6  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.14.6

# 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 3.53 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	1028 (3.60-3.48)
Clashscore	141614	1109 (3.60-3.48)
Ramachandran outliers	138981	1073 (3.60-3.48)
Sidechain outliers	138945	1074 (3.60-3.48)
RSRZ outliers	127900	1079 (3.62-3.46)

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	347	<div><div></div><div>44%32%21%</div></div>
1	F	347	<div><div>10%</div><div>36%36%22%</div></div>
2	B	100	<div><div></div><div>57%40%</div></div>
2	G	100	<div><div>10%</div><div>53%37%6%</div></div>
3	C	210	<div><div></div><div>67%27%6%</div></div>
3	H	210	<div><div></div><div>56%30%14%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	D	241	
4	I	241	
5	E	119	
5	J	119	
6	K	2	
6	L	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	L	1	-	-	X	-

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 14208 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 Antigen-presenting glycoprotein CD1d.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2130	1362	372	389	7			
1	F	270	Total	C	N	O	S	0	0	0
			2075	1327	366	376	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	4	MET	-	initiating methionine	UNP P15813
F	4	MET	-	initiating methionine	UNP P15813

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	98	Total	C	N	O	S	0	0	0
			743	477	116	148	2			
2	G	94	Total	C	N	O	S	0	0	0
			745	473	126	144	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	MET	-	initiating methionine	UNP P61769
G	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called T cell receptor alpha variable 10, nkt tcr alpha chain fusion.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	198	Total	C	N	O	S	0	0	0
			1514	937	256	312	9			
3	H	181	Total	C	N	O	S	0	0	0
			1376	844	237	287	8			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	HIS	-	expression tag	UNP A0A0B4J240
C	2	MET	-	expression tag	UNP A0A0B4J240
H	1	HIS	-	expression tag	UNP A0A0B4J240
H	2	MET	-	expression tag	UNP A0A0B4J240

- Molecule 4 is a protein called nkt tcr beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	240	Total	C	N	O	S	0	0	0
			1848	1166	317	357	8			
4	I	237	Total	C	N	O	S	0	0	0
			1888	1190	327	363	8			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	92	THR	SER	linker	UNP K7N5M4
D	93	SER	GLU	linker	UNP K7N5M4
D	94	ARG	ASN	linker	UNP K7N5M4
D	95	ARG	SER	linker	UNP K7N5M4
D	97	SER	THR	linker	UNP K7N5M4
D	?	-	GLY	deletion	UNP K7N5M4
D	?	-	ARG	deletion	UNP K7N5M4
D	?	-	ILE	deletion	UNP K7N5M4
I	92	THR	SER	linker	UNP K7N5M4
I	93	SER	GLU	linker	UNP K7N5M4
I	94	ARG	ASN	linker	UNP K7N5M4
I	95	ARG	SER	linker	UNP K7N5M4
I	97	SER	THR	linker	UNP K7N5M4
I	?	-	GLY	deletion	UNP K7N5M4
I	?	-	ARG	deletion	UNP K7N5M4
I	?	-	ILE	deletion	UNP K7N5M4

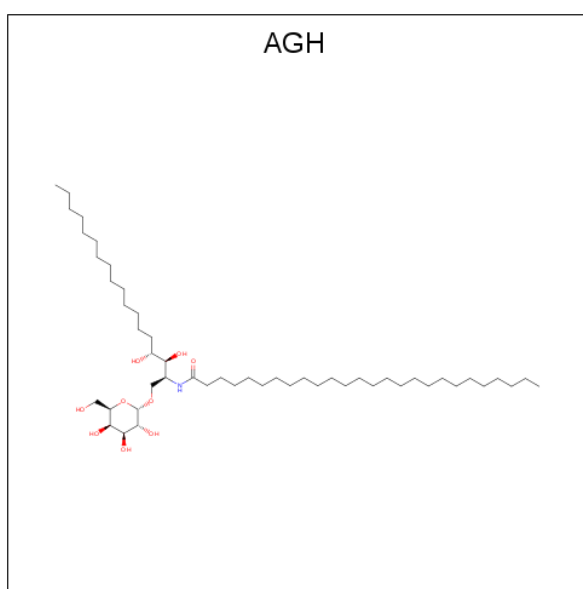
- Molecule 5 is a protein called Nanobody VHH ID12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	117	Total	C	N	O	S	0	0	0
			851	529	144	174	4			
5	J	118	Total	C	N	O	S	0	1	0
			834	522	140	168	4			

- Molecule 6 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
6	K	2	Total	C	N	O	0	0	0
			28	16	2	10			
6	L	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 7 is N-{(1S,2R,3S)-1-[(ALPHA-D-GALACTOPYRANOSYLOXY)METHYL]-2,3-DIHYDROXYHEPTADECYL}HEXACOSANAMIDE (three-letter code: AGH) (formula: C<sub>50</sub>H<sub>99</sub>NO<sub>9</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
7	A	1	Total	C	N	O	7	0
			60	50	1	9		
7	F	1	Total	C	N	O	16	0
			60	50	1	9		

- Molecule 8 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).

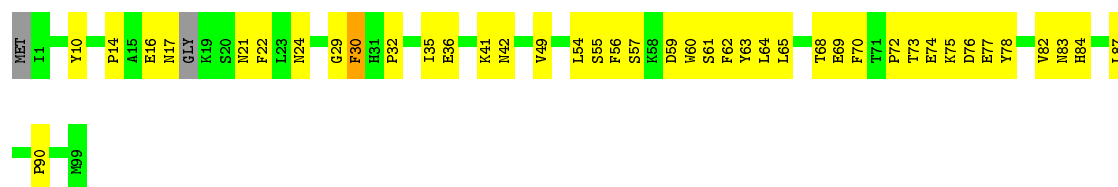


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	F	1	Total	C	N	O	0	0
			14	8	1	5		
8	F	1	Total	C	N	O	0	0
			14	8	1	5		



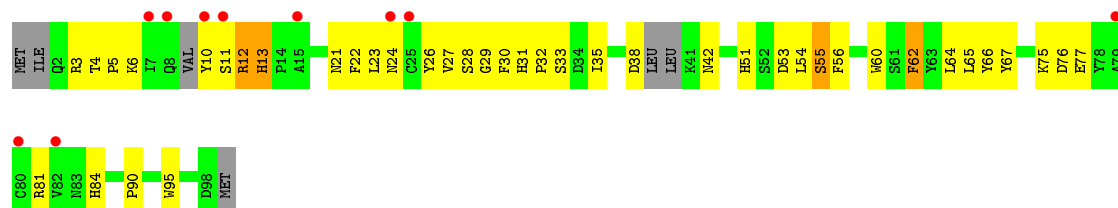


Chain B: 



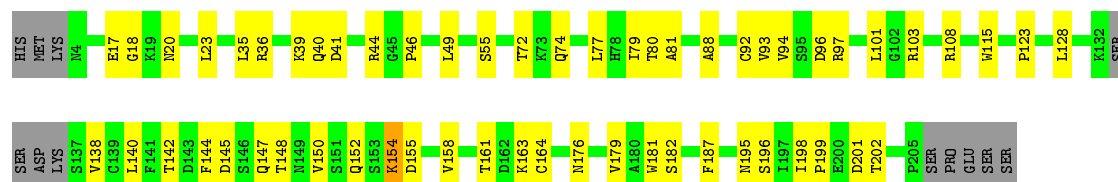
• Molecule 2: Beta-2-microglobulin

Chain G: 



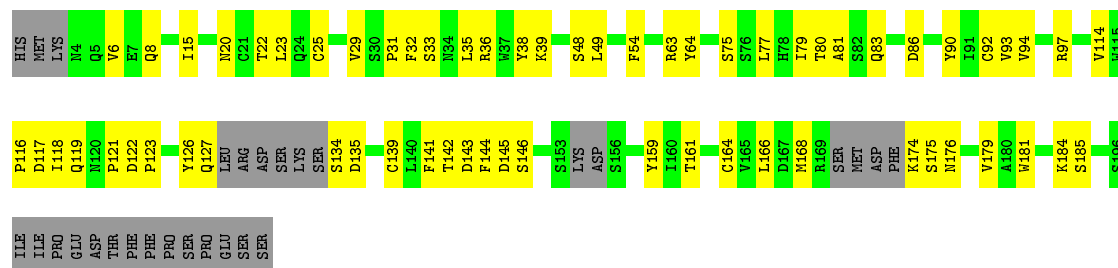
• Molecule 3: T cell receptor alpha variable 10, nkt tcr alpha chain fusion

Chain C: 



• Molecule 3: T cell receptor alpha variable 10, nkt tcr alpha chain fusion

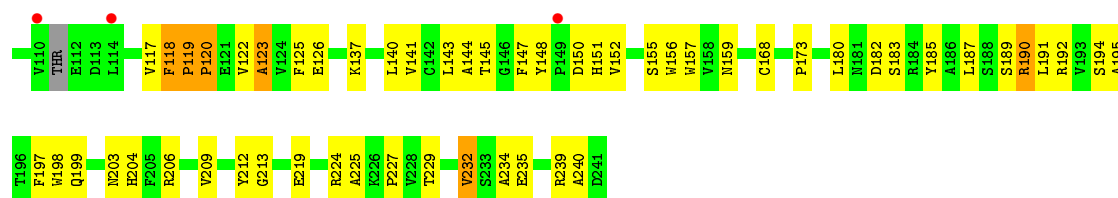
Chain H: 



• Molecule 4: nkt tcr beta chain

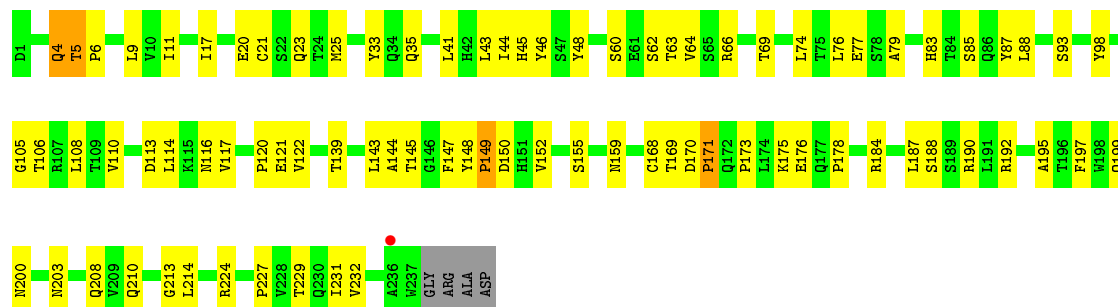
Chain D: 





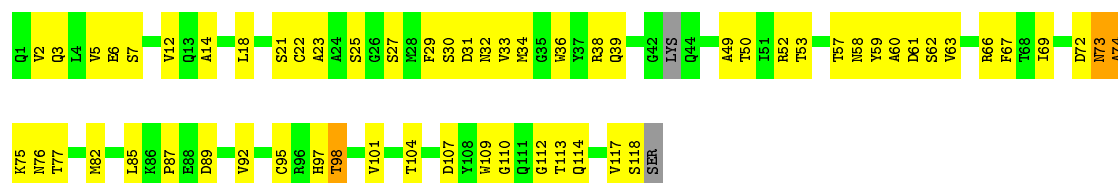
• Molecule 4: nkt tcr beta chain

Chain I: 64% 33% ..



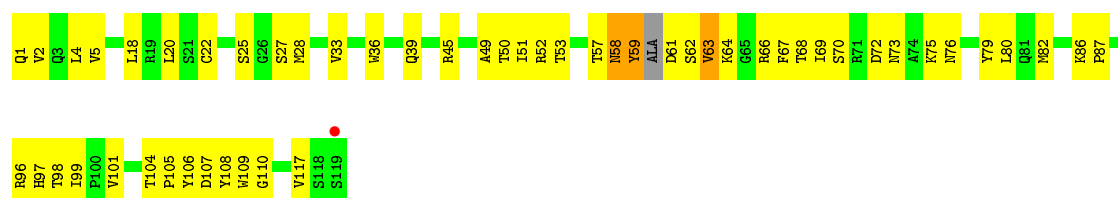
• Molecule 5: Nanobody VHH ID12

Chain E: 48% 48% ..



• Molecule 5: Nanobody VHH ID12

Chain J: 55% 42% ..



• Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain K: 50% 50%



- Molecule 6: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain L:  100%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	210.67Å 165.25Å 84.50Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	82.63 – 3.53 82.62 – 3.53	Depositor EDS
% Data completeness (in resolution range)	100.0 (82.63-3.53) 100.0 (82.62-3.53)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 3.49Å)	Xtriage
Refinement program	PHENIX 1.14 _3260	Depositor
R, $R_{free}$	0.195 , 0.254 0.210 , 0.253	Depositor DCC
$R_{free}$ test set	1873 reflections (5.27%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	104.5	Xtriage
Anisotropy	0.327	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 96.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.36$ , $\langle L^2 \rangle = 0.19$	Xtriage
Estimated twinning fraction	0.297 for -h,-k,l	Xtriage
Reported twinning fraction	0.390 for -h,-k,l	Depositor
Outliers	0 of 35558 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	14208	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	120.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.70% 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

### 5.1 Standard geometry

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

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	0.55	2/2194 (0.1%)	0.76	5/2996 (0.2%)
1	F	0.79	6/2135 (0.3%)	1.00	11/2917 (0.4%)
2	B	0.36	0/763	0.74	0/1049
2	G	0.60	2/765 (0.3%)	0.90	3/1038 (0.3%)
3	C	0.32	0/1542	0.55	0/2097
3	H	0.36	0/1398	0.65	0/1901
4	D	0.52	2/1901 (0.1%)	0.73	6/2598 (0.2%)
4	I	0.62	4/1942 (0.2%)	0.71	4/2648 (0.2%)
5	E	0.47	0/868	0.79	1/1185 (0.1%)
5	J	0.47	0/854	0.95	2/1168 (0.2%)
All	All	0.55	16/14362 (0.1%)	0.78	32/19597 (0.2%)

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	F	0	2
5	J	0	1
All	All	0	3

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	237	PRO	N-CD	-20.49	1.19	1.47
1	F	138	PRO	N-CD	-15.73	1.25	1.47
4	D	119	PRO	N-CA	12.84	1.69	1.47
4	I	149	PRO	N-CA	12.09	1.67	1.47
1	A	232	PRO	N-CD	-10.44	1.33	1.47

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	237	PRO	CA-N-CD	11.30	127.53	111.70
4	D	225	ALA	N-CA-CB	8.73	122.33	110.10
1	F	213	PRO	N-CA-CB	-8.45	93.16	103.30
2	G	32	PRO	CA-N-CD	-8.22	99.99	111.50
1	A	232	PRO	CA-N-CD	8.20	123.18	111.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	F	212	TYR	Peptide
1	F	41	SER	Mainchain
5	J	58	ASN	Mainchain

## 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	2130	0	2007	168	0
1	F	2075	0	1936	254	0
2	B	743	0	633	57	0
2	G	745	0	662	61	0
3	C	1514	0	1424	47	0
3	H	1376	0	1281	63	0
4	D	1848	0	1681	110	0
4	I	1888	0	1781	80	0
5	E	851	0	780	69	0
5	J	834	0	755	75	0
6	K	28	0	25	1	0
6	L	28	0	25	11	0
7	A	60	0	99	3	0
7	F	60	0	99	3	0
8	F	28	0	26	4	0
All	All	14208	0	13214	941	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 34.

The worst 5 of 941 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:F:166:CYS:SG	1:F:167:PRO:HD3	1.11	1.61
1:F:213:PRO:CG	1:F:265:HIS:HE1	1.28	1.46
1:F:213:PRO:HG2	1:F:265:HIS:CE1	1.47	1.45
1:F:185:VAL:CG1	1:F:212:TYR:H	1.27	1.45
1:F:166:CYS:SG	1:F:167:PRO:CD	2.04	1.45

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	272/347 (78%)	249 (92%)	19 (7%)	4 (2%)	10	47
1	F	266/347 (77%)	234 (88%)	24 (9%)	8 (3%)	4	33
2	B	94/100 (94%)	87 (93%)	7 (7%)	0	100	100
2	G	88/100 (88%)	83 (94%)	5 (6%)	0	100	100
3	C	194/210 (92%)	178 (92%)	16 (8%)	0	100	100
3	H	173/210 (82%)	159 (92%)	14 (8%)	0	100	100
4	D	236/241 (98%)	209 (89%)	27 (11%)	0	100	100
4	I	235/241 (98%)	222 (94%)	12 (5%)	1 (0%)	34	71
5	E	113/119 (95%)	98 (87%)	14 (12%)	1 (1%)	17	58
5	J	115/119 (97%)	101 (88%)	13 (11%)	1 (1%)	17	58
All	All	1786/2034 (88%)	1620 (91%)	151 (8%)	15 (1%)	19	60

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	212	TYR
1	A	232	PRO
1	F	19	ALA
5	E	74	ALA
1	F	21	SER

### 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	222/302 (74%)	218 (98%)	4 (2%)	59	81
1	F	212/302 (70%)	205 (97%)	7 (3%)	38	69
2	B	76/95 (80%)	75 (99%)	1 (1%)	69	87
2	G	79/95 (83%)	77 (98%)	2 (2%)	47	76
3	C	173/192 (90%)	172 (99%)	1 (1%)	86	94
3	H	157/192 (82%)	157 (100%)	0	100	100
4	D	193/214 (90%)	191 (99%)	2 (1%)	76	89
4	I	207/214 (97%)	206 (100%)	1 (0%)	88	95
5	E	88/98 (90%)	85 (97%)	3 (3%)	37	69
5	J	83/98 (85%)	82 (99%)	1 (1%)	71	87
All	All	1490/1802 (83%)	1468 (98%)	22 (2%)	65	84

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	E	73	ASN
1	F	32	LEU
4	I	4	GLN
5	E	98	THR
1	F	20	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:



Mol	Chain	Res	Type
2	G	13	HIS
2	G	24	ASN
5	J	32	ASN
1	F	55	GLN
4	I	151	HIS

### 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 ⓘ

4 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	$\# Z  > 2$
6	NAG	K	1	1,6	14,14,15	0.47	0	17,19,21	0.69	0
6	NAG	K	2	6	14,14,15	0.25	0	17,19,21	0.28	0
6	NAG	L	1	1,6	14,14,15	1.48	4 (28%)	17,19,21	3.88	8 (47%)
6	NAG	L	2	6	14,14,15	0.69	0	17,19,21	1.99	6 (35%)

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	K	1	1,6	-	2/6/23/26	0/1/1/1
6	NAG	K	2	6	-	1/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	L	1	1,6	-	3/6/23/26	0/1/1/1
6	NAG	L	2	6	-	4/6/23/26	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	L	1	NAG	O5-C1	-2.64	1.39	1.43
6	L	1	NAG	C2-N2	-2.35	1.42	1.46
6	L	1	NAG	C4-C5	-2.04	1.48	1.53
6	L	1	NAG	C1-C2	-2.02	1.49	1.52

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	L	1	NAG	C2-N2-C7	-11.12	107.06	122.90
6	L	1	NAG	C6-C5-C4	-6.43	97.95	113.00
6	L	1	NAG	O3-C3-C4	-5.10	98.56	110.35
6	L	2	NAG	O5-C5-C6	4.86	114.83	107.20
6	L	1	NAG	O4-C4-C3	4.19	120.05	110.35

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

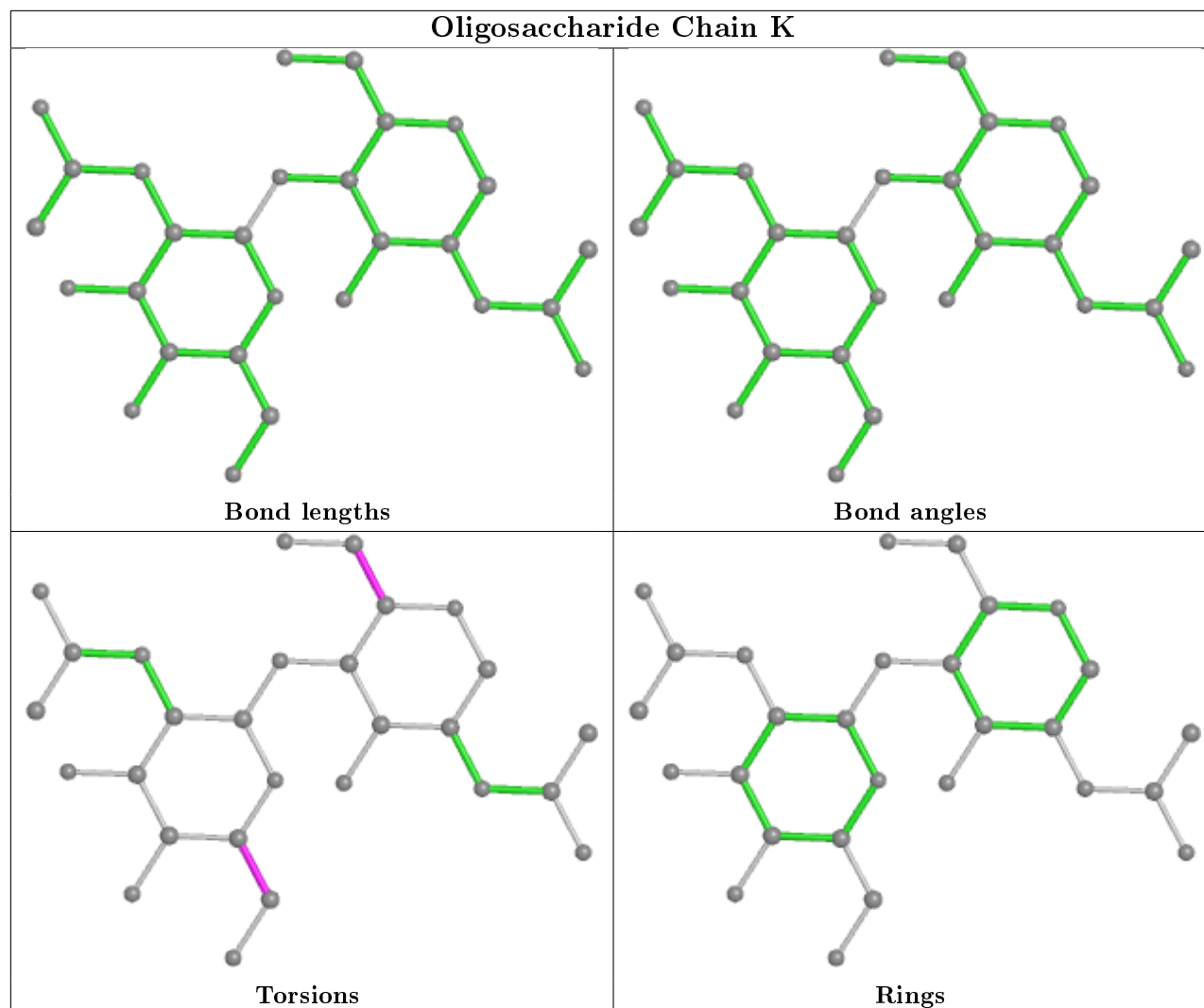
Mol	Chain	Res	Type	Atoms
6	L	1	NAG	C3-C2-N2-C7
6	L	1	NAG	C8-C7-N2-C2
6	L	1	NAG	O7-C7-N2-C2
6	L	2	NAG	C8-C7-N2-C2
6	K	1	NAG	O5-C5-C6-O6

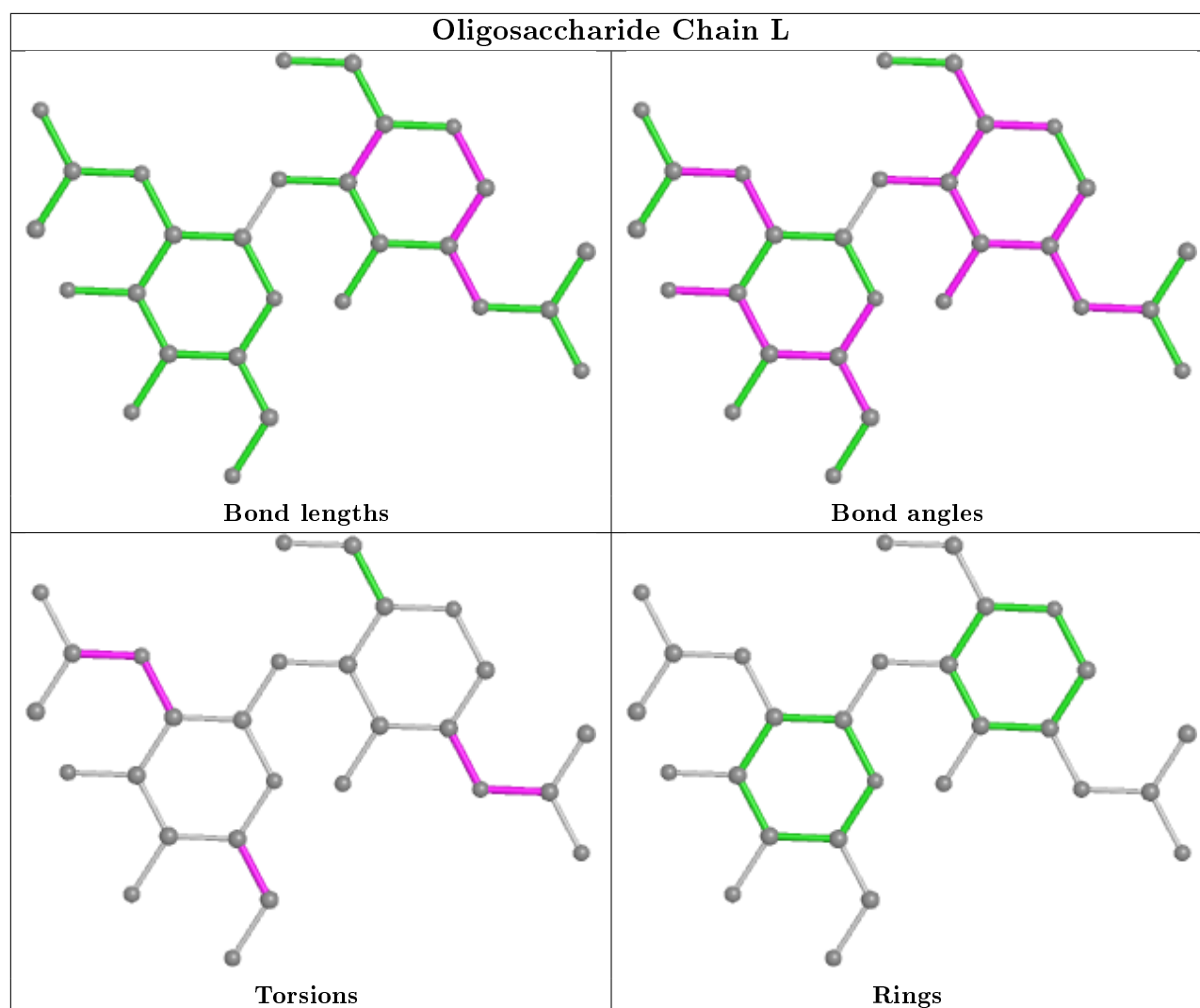
There are no ring outliers.

3 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	L	2	NAG	3	0
6	L	1	NAG	8	0
6	K	1	NAG	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](#)

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
8	NAG	F	404	1	14,14,15	0.76	0	17,19,21	1.55	4 (23%)
7	AGH	F	401	-	60,60,60	1.38	5 (8%)	65,69,69	1.15	4 (6%)
7	AGH	A	401	-	60,60,60	1.40	5 (8%)	65,69,69	1.02	4 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	NAG	F	405	1	14,14,15	1.10	2 (14%)	17,19,21	3.02	6 (35%)

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	NAG	F	404	1	-	4/6/23/26	0/1/1/1
7	AGH	F	401	-	-	22/58/78/78	0/1/1/1
7	AGH	A	401	-	-	15/58/78/78	0/1/1/1
8	NAG	F	405	1	-	5/6/23/26	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	A	401	AGH	CAA-N2	7.02	1.49	1.34
7	F	401	AGH	CAA-N2	6.96	1.48	1.34
7	F	401	AGH	O1A-C1	-3.08	1.38	1.43
7	A	401	AGH	O1A-C1	-2.98	1.38	1.43
7	A	401	AGH	CAB-CAA	2.78	1.56	1.51

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	F	405	NAG	C1-O5-C5	7.61	122.50	112.19
8	F	405	NAG	O5-C1-C2	-6.09	101.66	111.29
8	F	405	NAG	C1-C2-N2	-4.26	103.20	110.49
8	F	405	NAG	C2-N2-C7	-3.98	117.23	122.90
7	F	401	AGH	C6-C5-C4	-3.67	108.15	114.18

There are no chirality outliers.

5 of 46 torsion outliers are listed below:

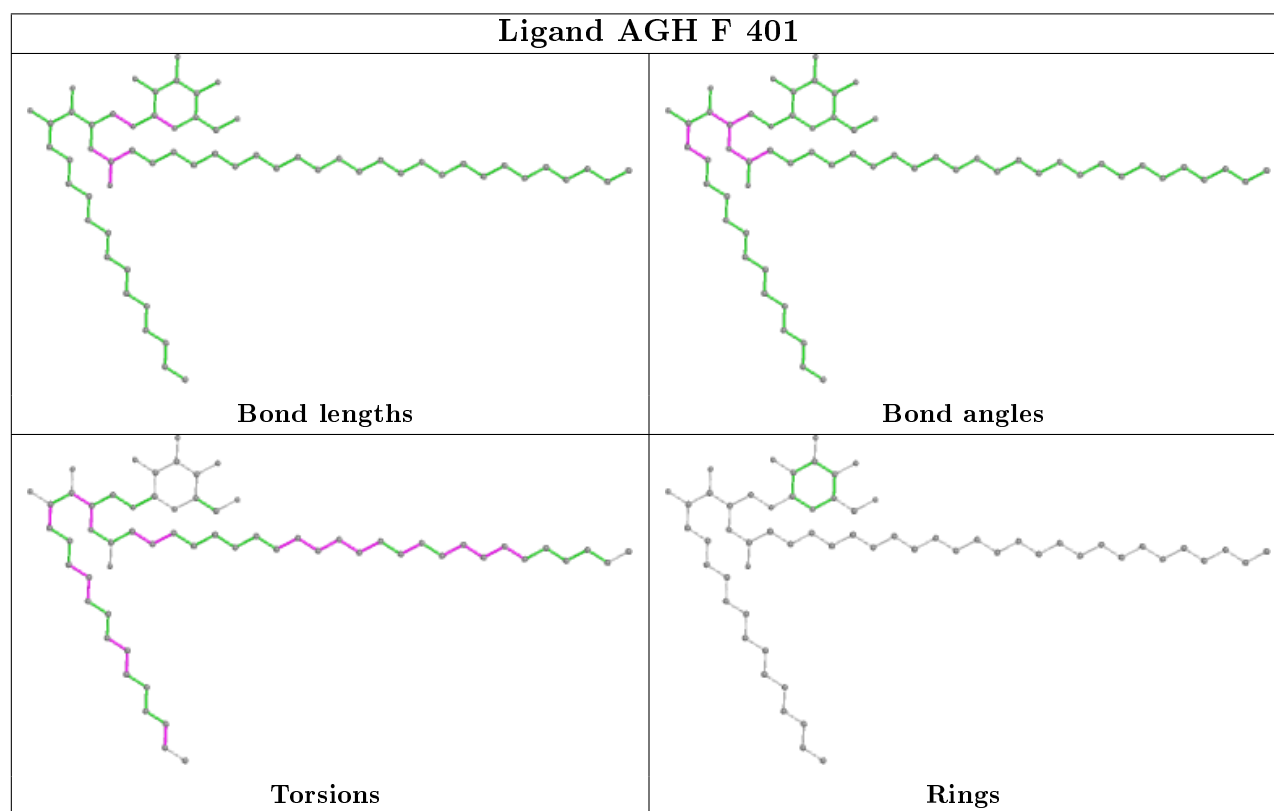
Mol	Chain	Res	Type	Atoms
8	F	404	NAG	C8-C7-N2-C2
8	F	404	NAG	O7-C7-N2-C2
7	F	401	AGH	N2-C2-C3-O3
7	A	401	AGH	O1A-C1-C2-C3
8	F	405	NAG	C3-C2-N2-C7

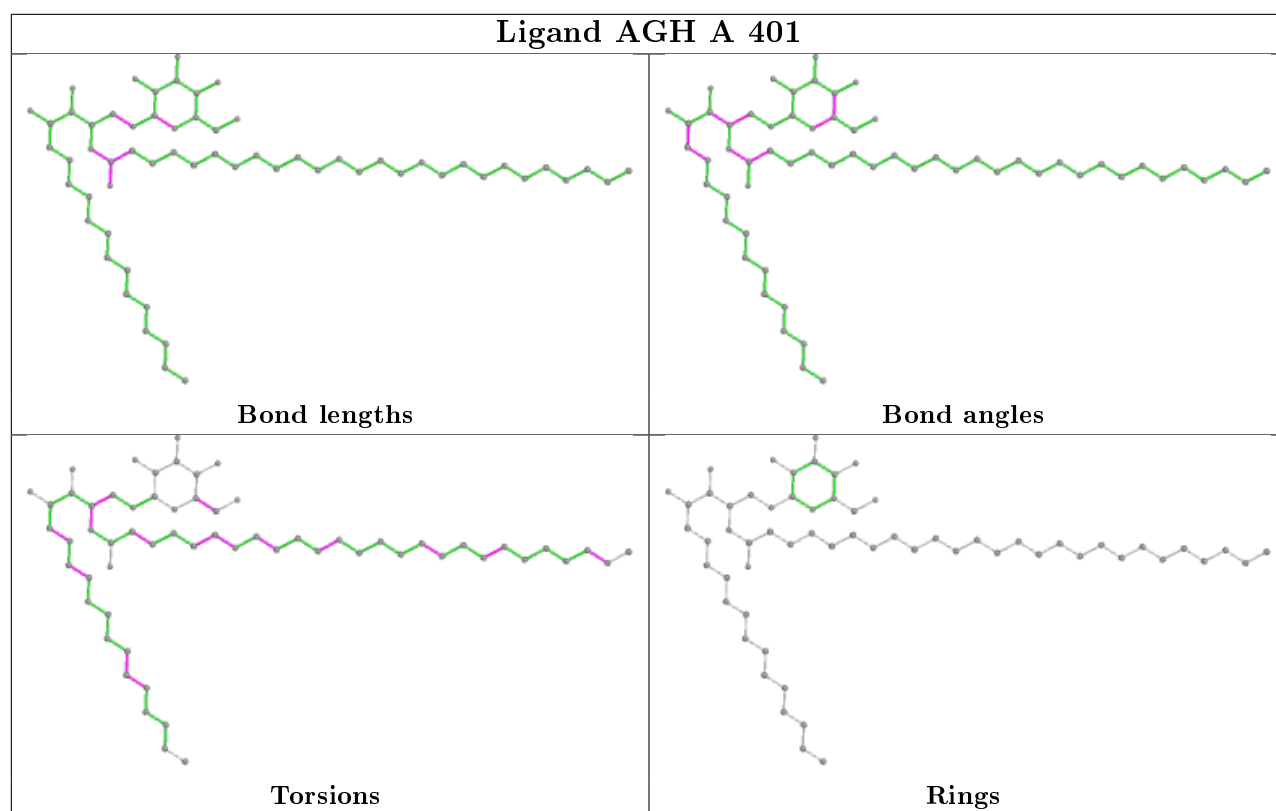
There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	F	401	AGH	3	0
7	A	401	AGH	3	0
8	F	405	NAG	4	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 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	274/347 (78%)	-0.44	0 100 100	30, 100, 147, 217	0
1	F	270/347 (77%)	-0.28	5 (1%) 66 53	30, 121, 230, 352	0
2	B	98/100 (98%)	-0.36	0 100 100	63, 114, 175, 229	0
2	G	94/100 (94%)	0.07	10 (10%) 6 6	62, 159, 247, 309	0
3	C	198/210 (94%)	-0.37	0 100 100	59, 113, 164, 211	0
3	H	181/210 (86%)	-0.26	0 100 100	64, 114, 190, 235	0
4	D	240/241 (99%)	-0.28	5 (2%) 63 50	64, 117, 197, 300	0
4	I	237/241 (98%)	-0.35	1 (0%) 92 87	51, 112, 170, 248	0
5	E	117/119 (98%)	-0.23	0 100 100	72, 126, 178, 236	0
5	J	118/119 (99%)	-0.36	1 (0%) 86 75	72, 127, 189, 233	0
All	All	1827/2034 (89%)	-0.31	22 (1%) 79 67	30, 116, 197, 352	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	80	CYS	3.7
1	F	208	VAL	3.7
2	G	24	ASN	3.5
2	G	11	SER	3.3
2	G	79	ALA	3.3

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

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

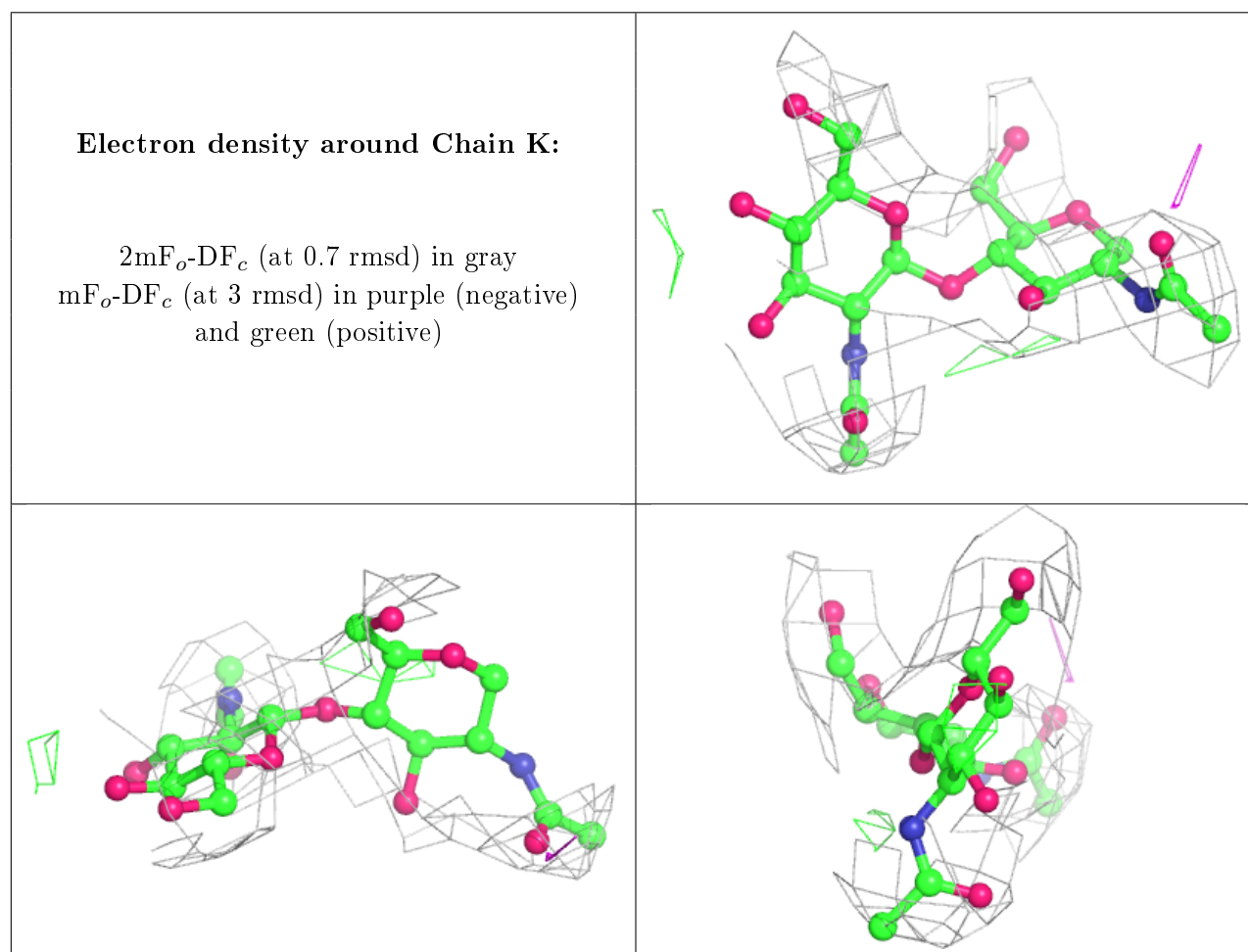


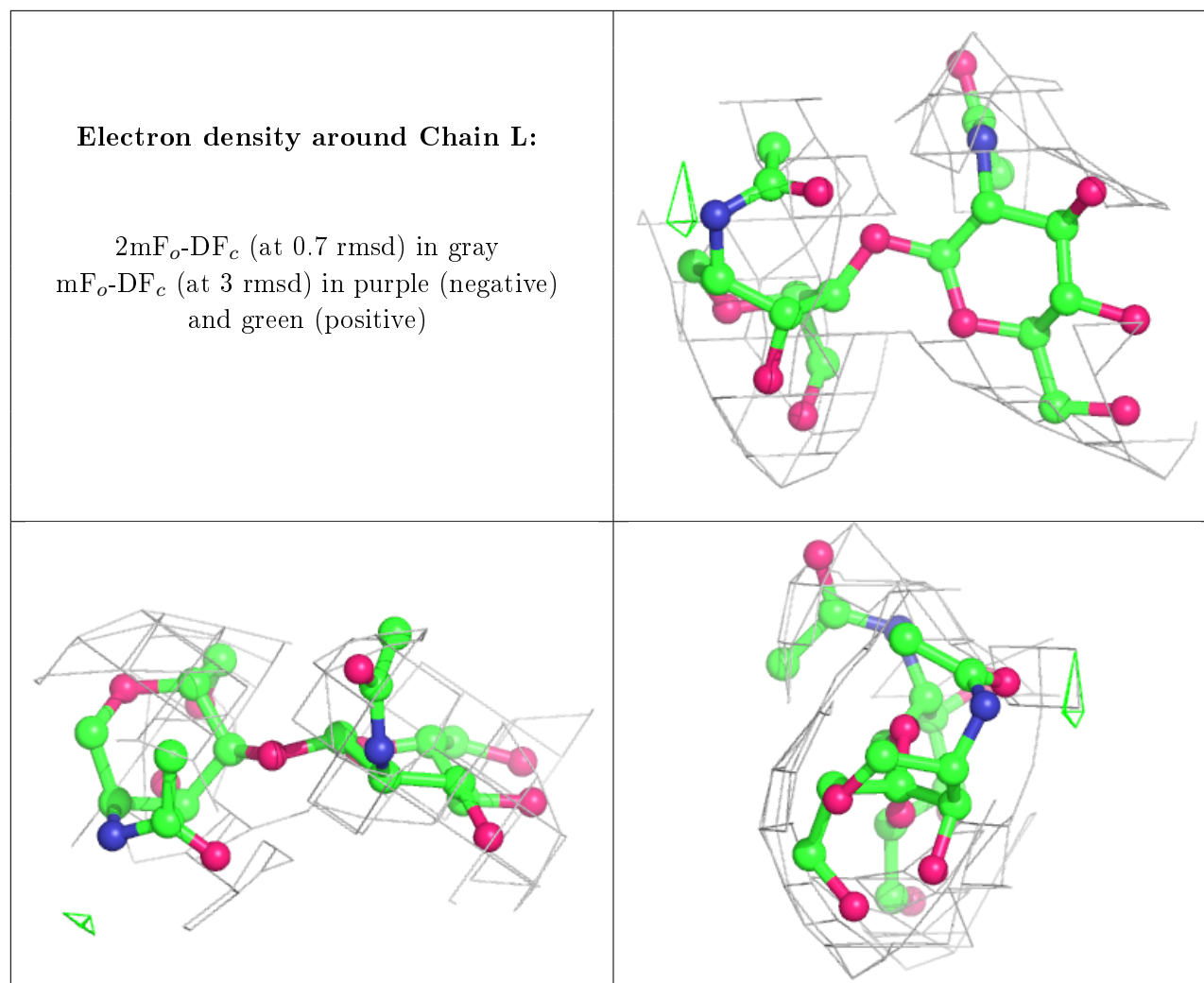
### 6.3 Carbohydrates ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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	K	2	14/15	0.87	0.19	93,102,109,111	0
6	NAG	L	2	14/15	0.88	0.20	136,142,154,156	0
6	NAG	K	1	14/15	0.91	0.17	85,94,100,102	0
6	NAG	L	1	14/15	0.96	0.12	87,107,117,129	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 ⓘ

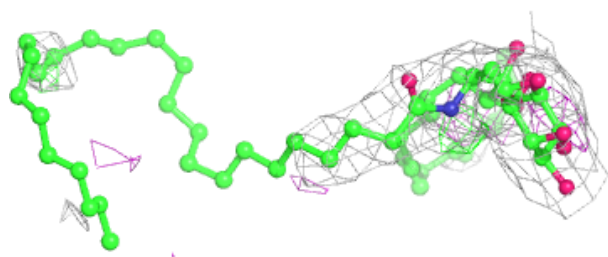
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
8	NAG	F	404	14/15	0.80	0.21	138,156,179,180	0
7	AGH	F	401	60/60	0.94	0.53	52,95,146,155	16
7	AGH	A	401	60/60	0.95	0.62	71,115,135,141	7
8	NAG	F	405	14/15	0.96	0.09	30,30,30,30	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

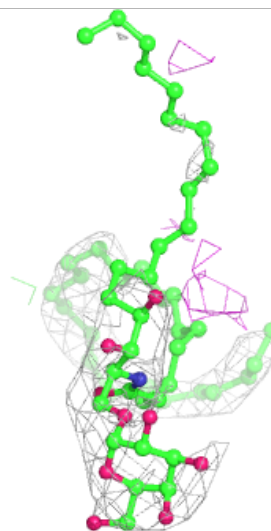
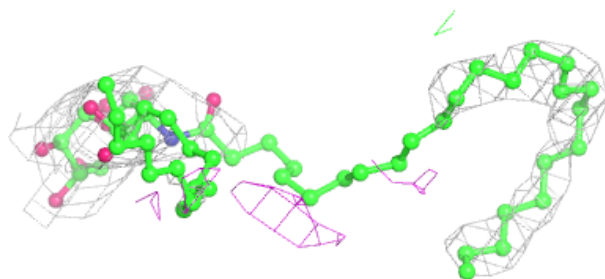
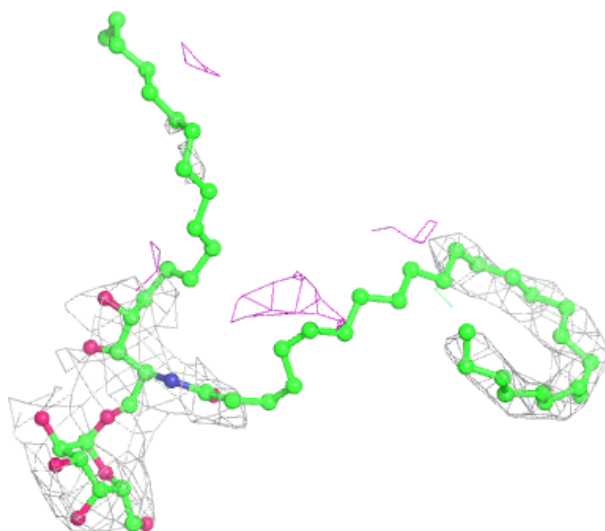
**Electron density around AGH F 401:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around AGH A 401:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



## 6.5 Other polymers ⓘ

There are no such residues in this entry.