



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 7, 2024 – 04:28 PM EST

PDB ID : 9CZA  
Title : Crystal structure of integrin avb6 headpiece in complex with compound 18  
Authors : Monroy, M.F.; Qiao, Q.; Lin, F.Y.  
Deposited on : 2024-08-05  
Resolution : 2.49 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 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 3.0  
buster-report : 1.1.7 (2018)  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.003 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.39

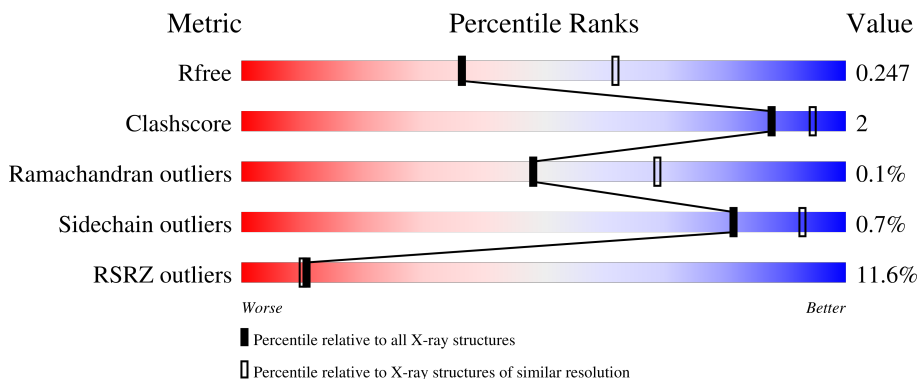
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.49 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))
$R_{free}$	164625	5504 (2.50-2.50)
Clashscore	180529	6282 (2.50-2.50)
Ramachandran outliers	177936	6191 (2.50-2.50)
Sidechain outliers	177891	6193 (2.50-2.50)
RSRZ outliers	164620	5504 (2.50-2.50)

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	605	 12% (red), 95% (green), 5% (yellow), 5% (grey)
2	B	481	 15% (red), 92% (green), 5% (yellow), 5% (grey)
3	C	214	 7% (red), 90% (green), 9% (yellow), 5% (grey)
4	D	218	 7% (red), 94% (green), 5% (yellow), 5% (grey)
5	G	3	 67% (yellow), 33% (orange)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	I	3	 67% 33%
6	E	6	 50% 50%
7	F	5	 60% 40%
8	L	2	 100%

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
15	ACT	D	302	-	-	X	-

## 2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 23181 atoms, of which 11290 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 Integrin alpha-V heavy chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	595	9093	2930	4472	783	887	21	0	2	0

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	400	CYS	MET	conflict	UNP P06756
A	596	GLY	-	expression tag	UNP P06756
A	597	GLY	-	expression tag	UNP P06756
A	598	SER	-	expression tag	UNP P06756
A	599	LEU	-	expression tag	UNP P06756
A	600	GLU	-	expression tag	UNP P06756
A	601	VAL	-	expression tag	UNP P06756
A	602	LEU	-	expression tag	UNP P06756
A	603	PHE	-	expression tag	UNP P06756
A	604	GLN	-	expression tag	UNP P06756
A	605	GLY	-	expression tag	UNP P06756

- Molecule 2 is a protein called Integrin beta-6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	466	7075	2234	3506	612	692	31	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	270	CYS	ILE	conflict	UNP P18564
B	475	SER	-	expression tag	UNP P18564
B	476	GLY	-	expression tag	UNP P18564
B	477	HIS	-	expression tag	UNP P18564
B	478	SER	-	expression tag	UNP P18564

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	479	LEU	-	expression tag	UNP P18564
B	480	GLU	-	expression tag	UNP P18564
B	481	VAL	-	expression tag	UNP P18564
B	482	LEU	-	expression tag	UNP P18564
B	483	PHE	-	expression tag	UNP P18564
B	484	GLN	-	expression tag	UNP P18564
B	485	GLY	-	expression tag	UNP P18564

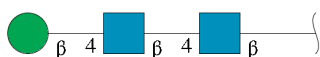
- Molecule 3 is a protein called 17E6 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
3	C	214	3248	1035	1582	280	344	7	0	0	0

- Molecule 4 is a protein called 17E6 Fab heavy chain.

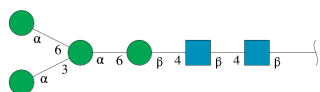
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
4	D	213	3121	1004	1531	258	317	11	0	0	0

- Molecule 5 is an oligosaccharide called beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



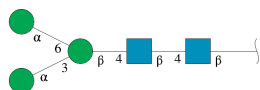
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
5	I	3	63	22	24	2	15	0	0	0
5	G	3	63	22	24	2	15	0	0	0

- Molecule 6 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
6	E	6	117	40	45	2	30	0	0	0

- Molecule 7 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
7	F	5	99	34	38	2	25	0	0	0

- Molecule 8 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

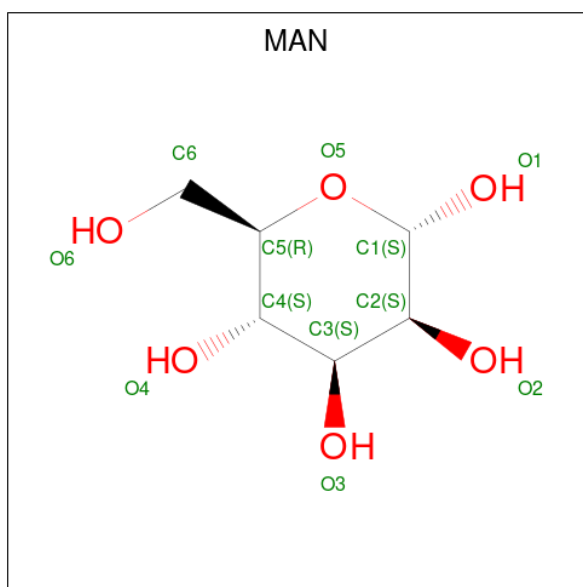


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
8	L	2	47	16	19	2	10	0	0	0

- Molecule 9 is CALCIUM ION (three-letter code: CA) (formula: Ca).

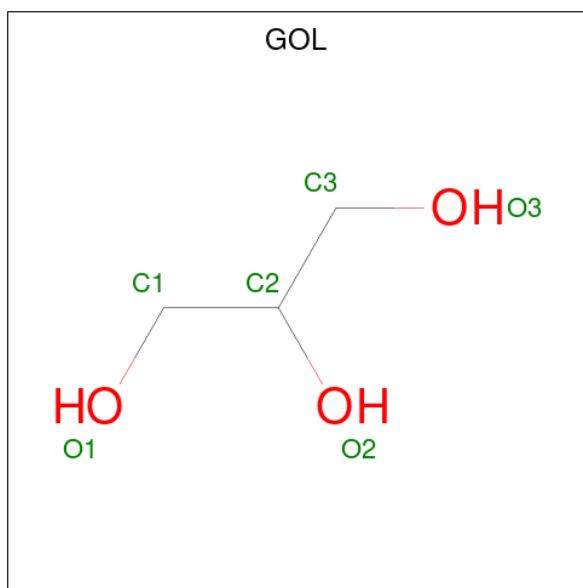
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	4	Total	Ca	0	0
			4	4		
9	B	2	Total	Ca	0	0
			2	2		

- Molecule 10 is alpha-D-mannopyranose (three-letter code: MAN) (formula: C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
10	A	1	19	6	8	5	0	0

- Molecule 11 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
11	A	1	9	3	3	3	0	0
11	A	1	9	3	3	3	0	0
11	A	1	9	3	3	3	0	0

*Continued on next page...*

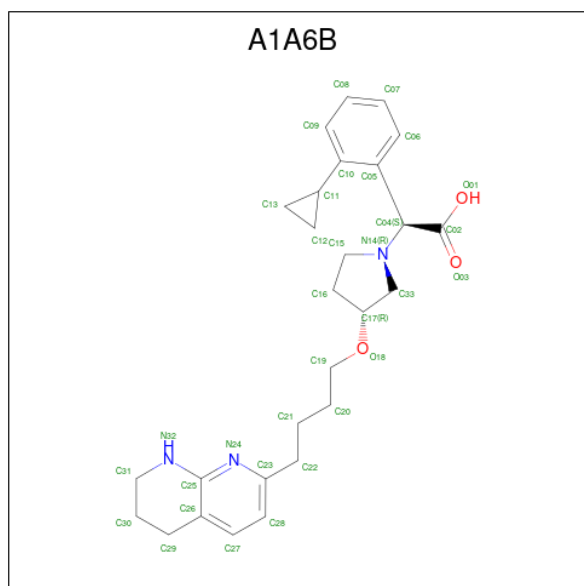
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	B	1	Total	C	H	O	0	0
			9	3	3	3		
11	B	1	Total	C	H	O	0	0
			9	3	3	3		
11	C	1	Total	C	H	O	0	0
			9	3	3	3		
11	D	1	Total	C	H	O	0	0
			9	3	3	3		

- Molecule 12 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	B	1	Total	Mg	0	0
			1	1		

- Molecule 13 is (2S)-(2-cyclopropylphenyl){(3R)-3-[4-(5,6,7,8-tetrahydro-1,8-naphthyridin-2-yl)butoxy]pyrrolidin-1-yl}acetic acid (three-letter code: A1A6B) (formula: C<sub>27</sub>H<sub>35</sub>N<sub>3</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
13	B	1	Total	C	N	O	0	0
			33	27	3	3		

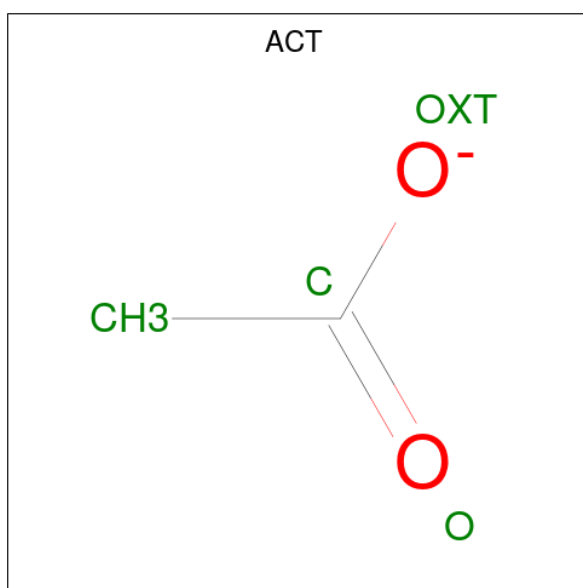
- Molecule 14 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf		
			Total	C	H	N			O	
14	B	1	Total	24	8	10	1	5	0	0
14	B	1	Total	24	8	10	1	5	0	0

- Molecule 15 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	C O		
15	B	1	Total	4 2 2	0	0
15	D	1	Total	4 2 2	0	0

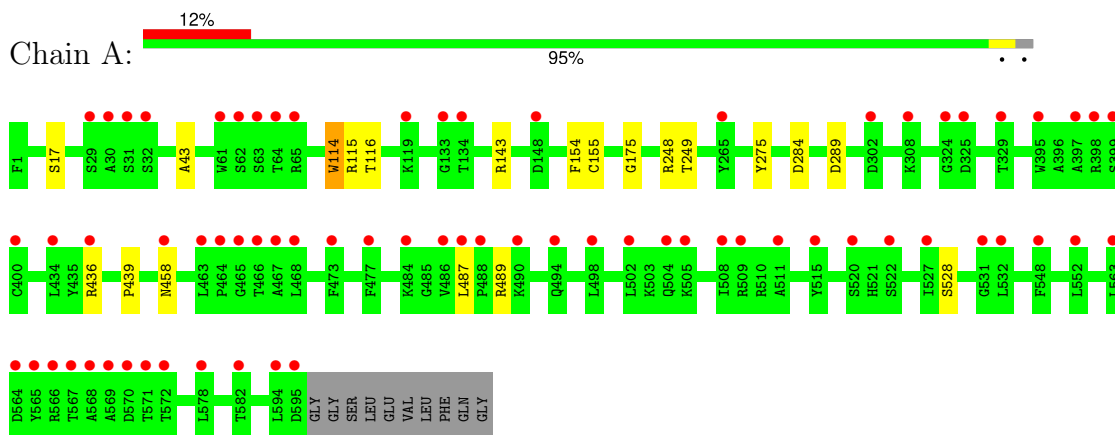
- Molecule 16 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
16	A	49	Total 49	O 49	0	0
16	B	26	Total 26	O 26	0	0
16	D	2	Total 2	O 2	0	0

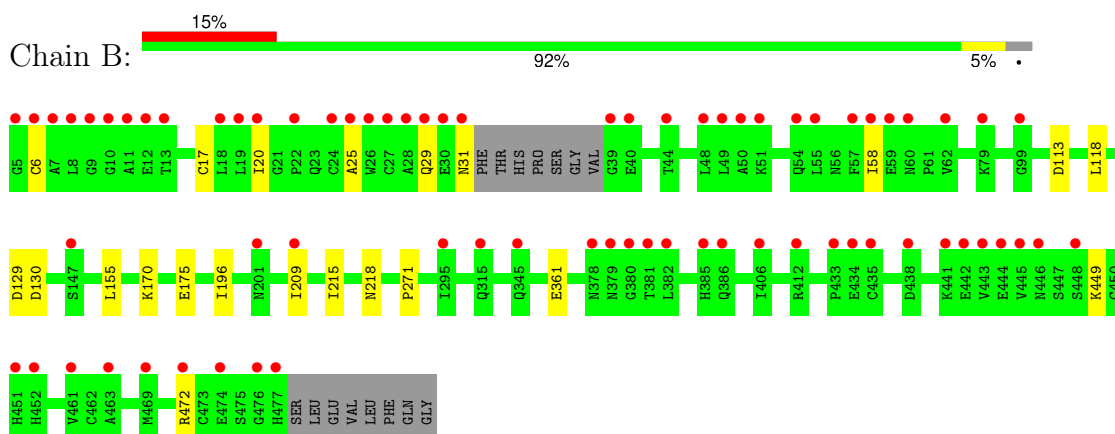
### 3 Residue-property plots [i](#)

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.

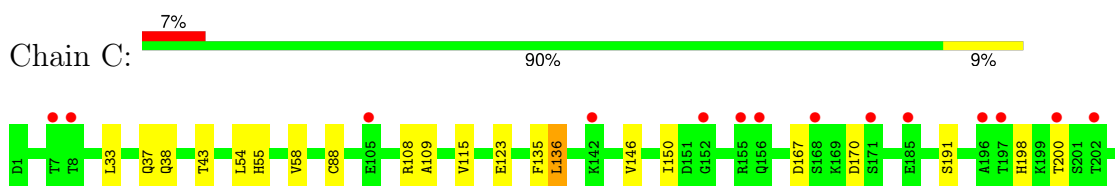
- Molecule 1: Integrin alpha-V heavy chain



- Molecule 2: Integrin beta-6

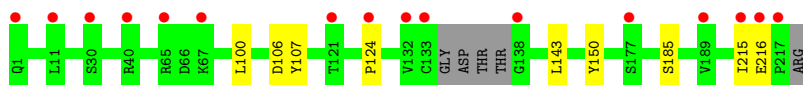


- Molecule 3: 17E6 Fab light chain





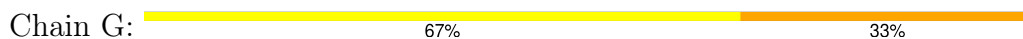
- Molecule 4: 17E6 Fab heavy chain



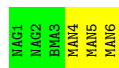
- Molecule 5: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 5: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 6: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 7: alpha-D-mannopyranose-(1-3)-[alpha-D-mannopyranose-(1-6)]beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 8: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



PAGE 13

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	96.25Å 132.06Å 168.61Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.80 – 2.49 41.80 – 2.49	Depositor EDS
% Data completeness (in resolution range)	99.7 (41.80-2.49) 99.7 (41.80-2.49)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.18 (at 2.48Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.223 , 0.244 0.227 , 0.247	Depositor DCC
$R_{free}$ test set	3774 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.6	Xtrriage
Anisotropy	0.092	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	23181	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	75.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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: MG, NAG, GOL, MAN, A1A6B, CA, BMA, ACT

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	0.27	0/4734	0.48	0/6410
2	B	0.27	0/3635	0.47	0/4922
3	C	0.26	0/1702	0.47	0/2309
4	D	0.27	0/1631	0.47	0/2226
All	All	0.27	0/11702	0.48	0/15867

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	A	4621	4472	4463	10	0
2	B	3569	3506	3500	13	0
3	C	1666	1582	1582	15	0
4	D	1590	1531	1531	6	0
5	G	39	24	34	1	0
5	I	39	24	34	0	0
6	E	72	45	61	0	0
7	F	61	38	52	0	0
8	L	28	19	25	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	A	4	0	0	0	0
9	B	2	0	0	0	0
10	A	11	8	10	0	0
11	A	18	9	24	1	0
11	B	12	6	16	0	0
11	C	6	3	8	0	0
11	D	6	3	8	0	0
12	B	1	0	0	0	0
13	B	33	0	0	0	0
14	B	28	20	26	1	0
15	B	4	0	3	1	0
15	D	4	0	3	3	0
16	A	49	0	0	0	0
16	B	26	0	0	0	0
16	D	2	0	0	0	0
All	All	11891	11290	11380	45	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

The worst 5 of 45 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)
3:C:54:LEU:HD11	3:C:58:VAL:HG23	1.65	0.78
1:A:439:PRO:HB2	1:A:487:LEU:HD21	1.69	0.75
5:G:1:NAG:O3	5:G:1:NAG:O7	2.12	0.66
3:C:136:LEU:HD12	3:C:136:LEU:N	2.15	0.62
2:B:113:ASP:OD2	15:B:2009:ACT:H2	2.00	0.61

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	595/605 (98%)	577 (97%)	18 (3%)	0	100	100
2	B	462/481 (96%)	438 (95%)	24 (5%)	0	100	100
3	C	212/214 (99%)	199 (94%)	13 (6%)	0	100	100
4	D	209/218 (96%)	207 (99%)	1 (0%)	1 (0%)	25	44
All	All	1478/1518 (97%)	1421 (96%)	56 (4%)	1 (0%)	48	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	216	GLU

### 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	490/495 (99%)	485 (99%)	5 (1%)	73	88
2	B	406/419 (97%)	404 (100%)	2 (0%)	86	95
3	C	192/192 (100%)	190 (99%)	2 (1%)	73	88
4	D	178/182 (98%)	178 (100%)	0	100	100
All	All	1266/1288 (98%)	1257 (99%)	9 (1%)	81	93

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	37	GLN
3	C	136	LEU
1	A	436	ARG
1	A	458	ASN
2	B	218	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

19 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	E	1	1,6	14,14,15	0.28	0	17,19,21	0.60	0
6	NAG	E	2	6	14,14,15	0.30	0	17,19,21	0.41	0
6	BMA	E	3	6	11,11,12	0.69	0	15,15,17	0.70	0
6	MAN	E	4	6	11,11,12	0.69	0	15,15,17	1.11	2 (13%)
6	MAN	E	5	6	11,11,12	0.70	0	15,15,17	0.86	1 (6%)
6	MAN	E	6	6	11,11,12	0.69	0	15,15,17	1.05	2 (13%)
7	NAG	F	1	1,7	14,14,15	0.30	0	17,19,21	0.44	0
7	NAG	F	2	7	14,14,15	0.33	0	17,19,21	0.45	0
7	BMA	F	3	7	11,11,12	0.45	0	15,15,17	0.74	0
7	MAN	F	4	7	11,11,12	0.56	0	15,15,17	0.92	2 (13%)
7	MAN	F	5	7	11,11,12	0.66	0	15,15,17	0.96	1 (6%)
5	NAG	G	1	5	14,14,15	0.57	0	17,19,21	1.03	1 (5%)
5	NAG	G	2	5	14,14,15	1.60	2 (14%)	17,19,21	1.22	3 (17%)
5	BMA	G	3	5	11,11,12	0.67	0	15,15,17	1.79	1 (6%)
5	NAG	I	1	1,5	14,14,15	0.41	0	17,19,21	0.55	0
5	NAG	I	2	5	14,14,15	0.63	1 (7%)	17,19,21	1.30	1 (5%)
5	BMA	I	3	5	11,11,12	0.68	0	15,15,17	0.72	0
8	NAG	L	1	8	14,14,15	0.55	0	17,19,21	0.63	0
8	NAG	L	2	8	14,14,15	0.67	0	17,19,21	0.42	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	E	1	1,6	-	0/6/23/26	0/1/1/1
6	NAG	E	2	6	-	4/6/23/26	0/1/1/1
6	BMA	E	3	6	-	2/2/19/22	0/1/1/1
6	MAN	E	4	6	-	2/2/19/22	0/1/1/1
6	MAN	E	5	6	-	0/2/19/22	0/1/1/1
6	MAN	E	6	6	-	2/2/19/22	0/1/1/1
7	NAG	F	1	1,7	-	0/6/23/26	0/1/1/1
7	NAG	F	2	7	-	1/6/23/26	0/1/1/1
7	BMA	F	3	7	-	2/2/19/22	0/1/1/1
7	MAN	F	4	7	-	1/2/19/22	0/1/1/1
7	MAN	F	5	7	-	2/2/19/22	0/1/1/1
5	NAG	G	1	5	-	4/6/23/26	0/1/1/1
5	NAG	G	2	5	-	4/6/23/26	0/1/1/1
5	BMA	G	3	5	-	2/2/19/22	0/1/1/1
5	NAG	I	1	1,5	-	0/6/23/26	0/1/1/1
5	NAG	I	2	5	-	4/6/23/26	0/1/1/1
5	BMA	I	3	5	-	2/2/19/22	0/1/1/1
8	NAG	L	1	8	-	2/6/23/26	0/1/1/1
8	NAG	L	2	8	-	3/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	2	NAG	C1-C2	4.09	1.57	1.52
5	G	2	NAG	O5-C1	-3.79	1.37	1.43
5	I	2	NAG	O5-C1	2.25	1.47	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	3	BMA	C1-O5-C5	5.97	120.19	112.19
5	I	2	NAG	C1-O5-C5	5.00	118.88	112.19
5	G	1	NAG	C1-O5-C5	3.66	117.08	112.19
5	G	2	NAG	C4-C3-C2	3.00	115.42	111.02
6	E	4	MAN	C1-O5-C5	2.73	115.84	112.19

There are no chirality outliers.

5 of 37 torsion outliers are listed below:

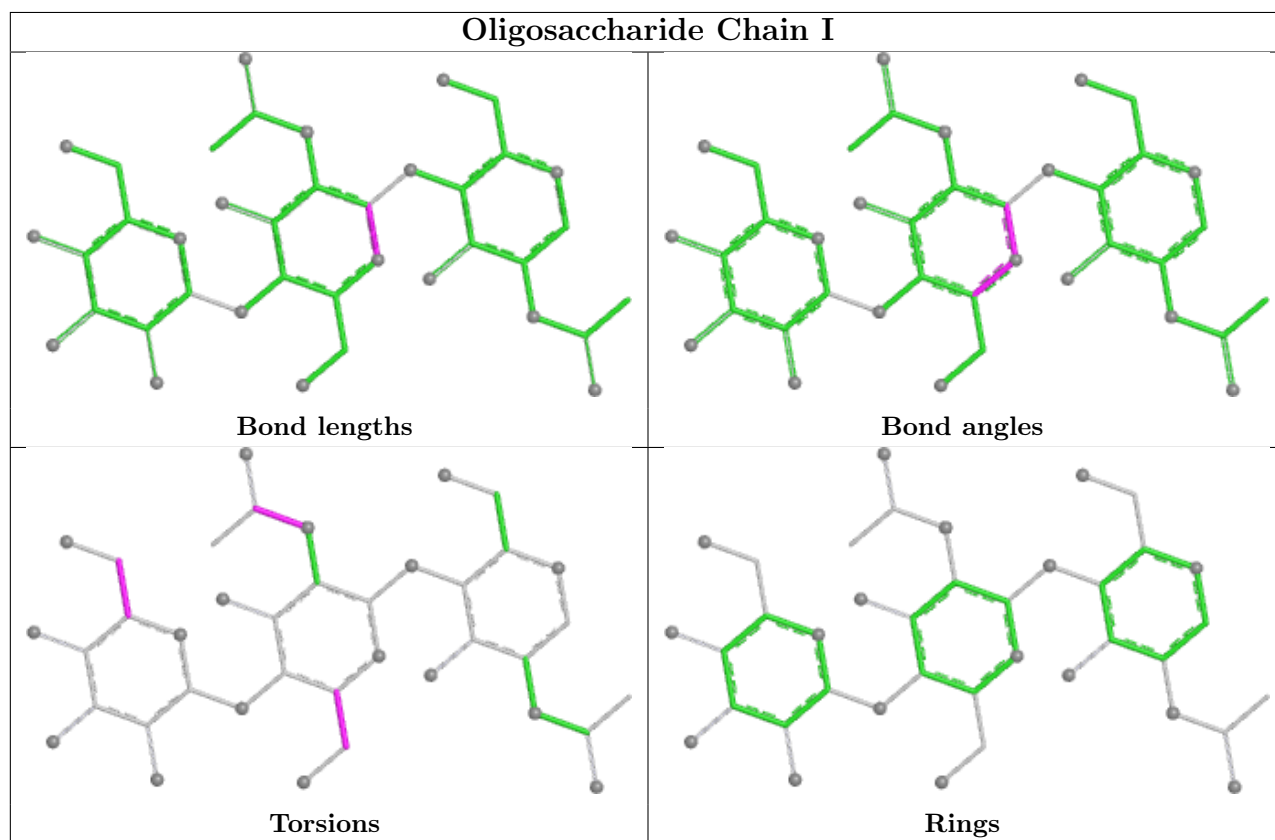
Mol	Chain	Res	Type	Atoms
6	E	6	MAN	O5-C5-C6-O6
5	G	2	NAG	O5-C5-C6-O6
8	L	1	NAG	O5-C5-C6-O6
6	E	3	BMA	O5-C5-C6-O6
8	L	1	NAG	C4-C5-C6-O6

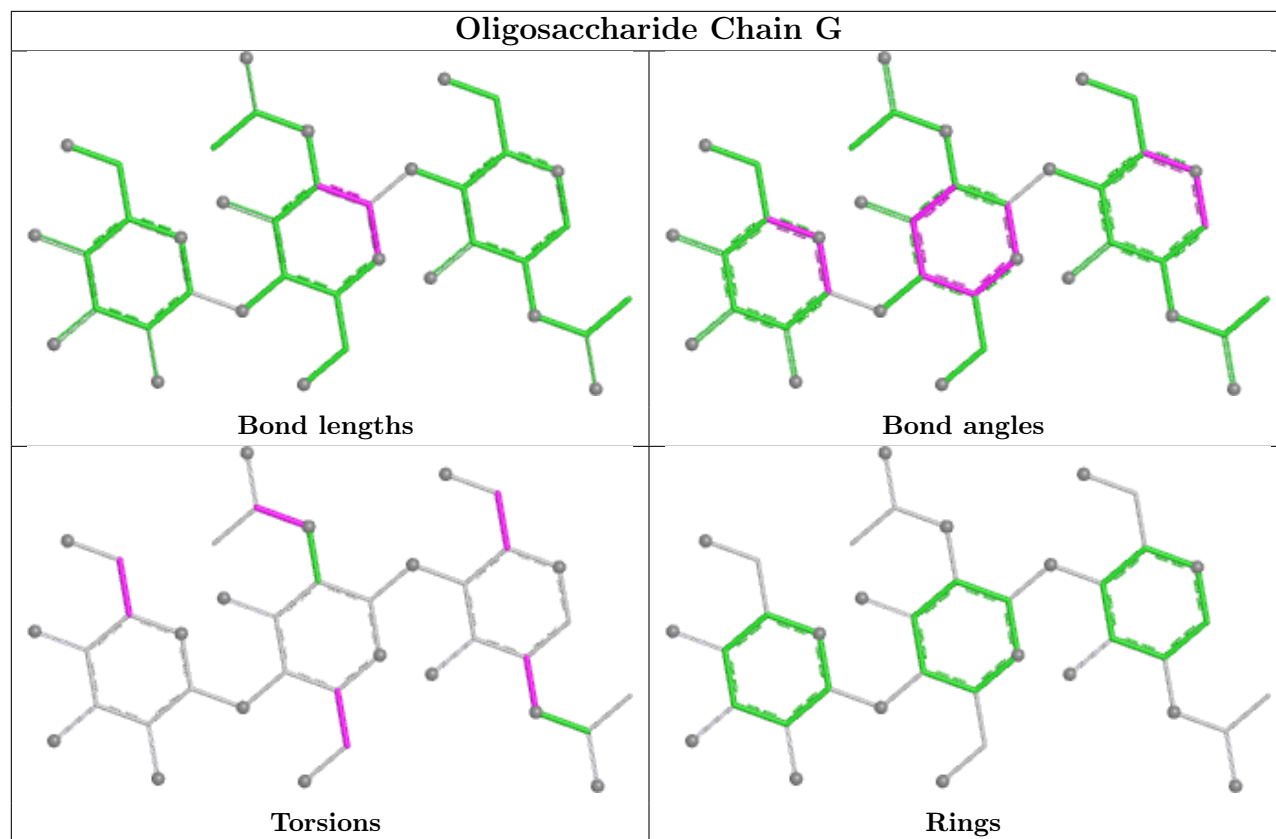
There are no ring outliers.

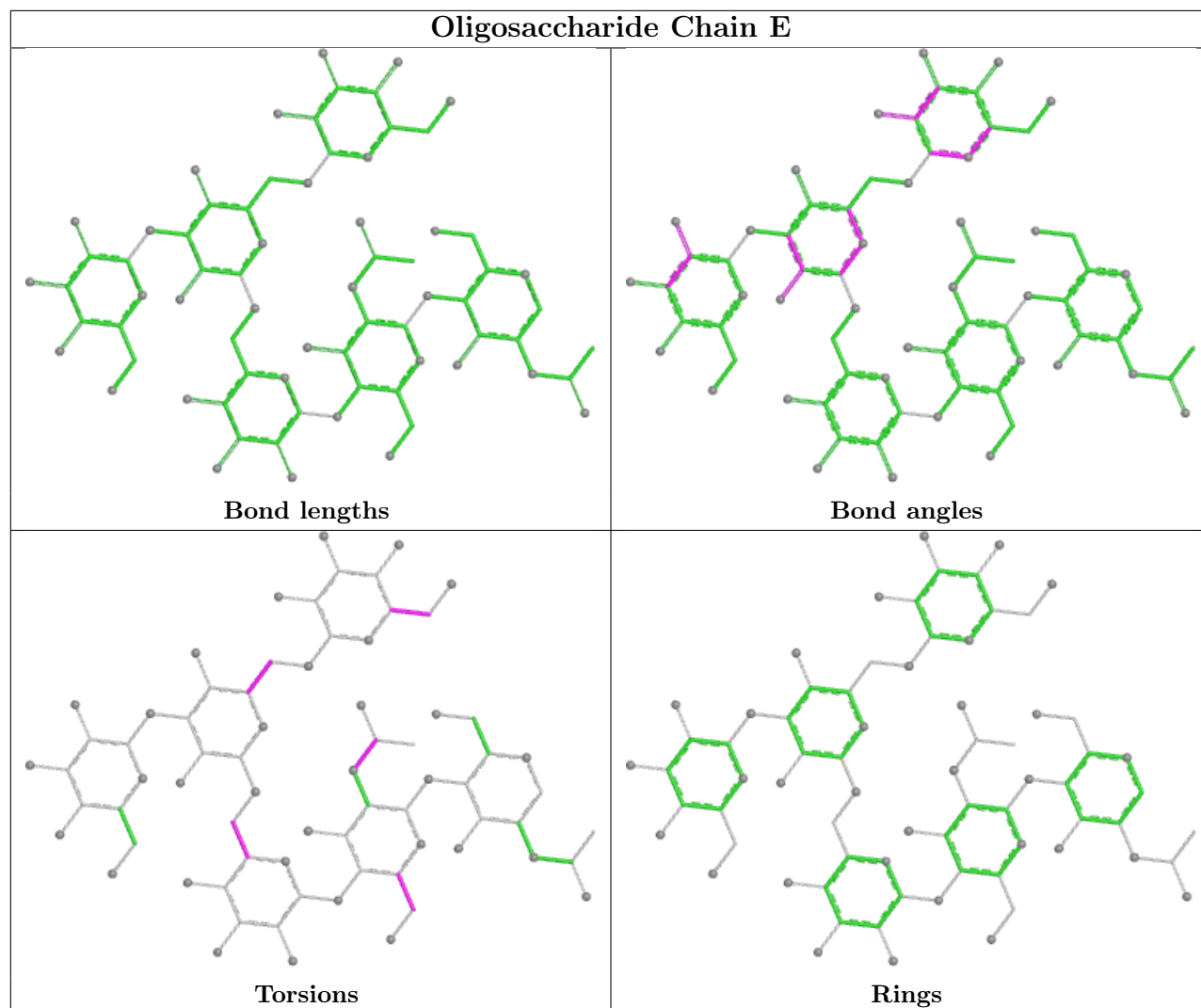
1 monomer is involved in 1 short contact:

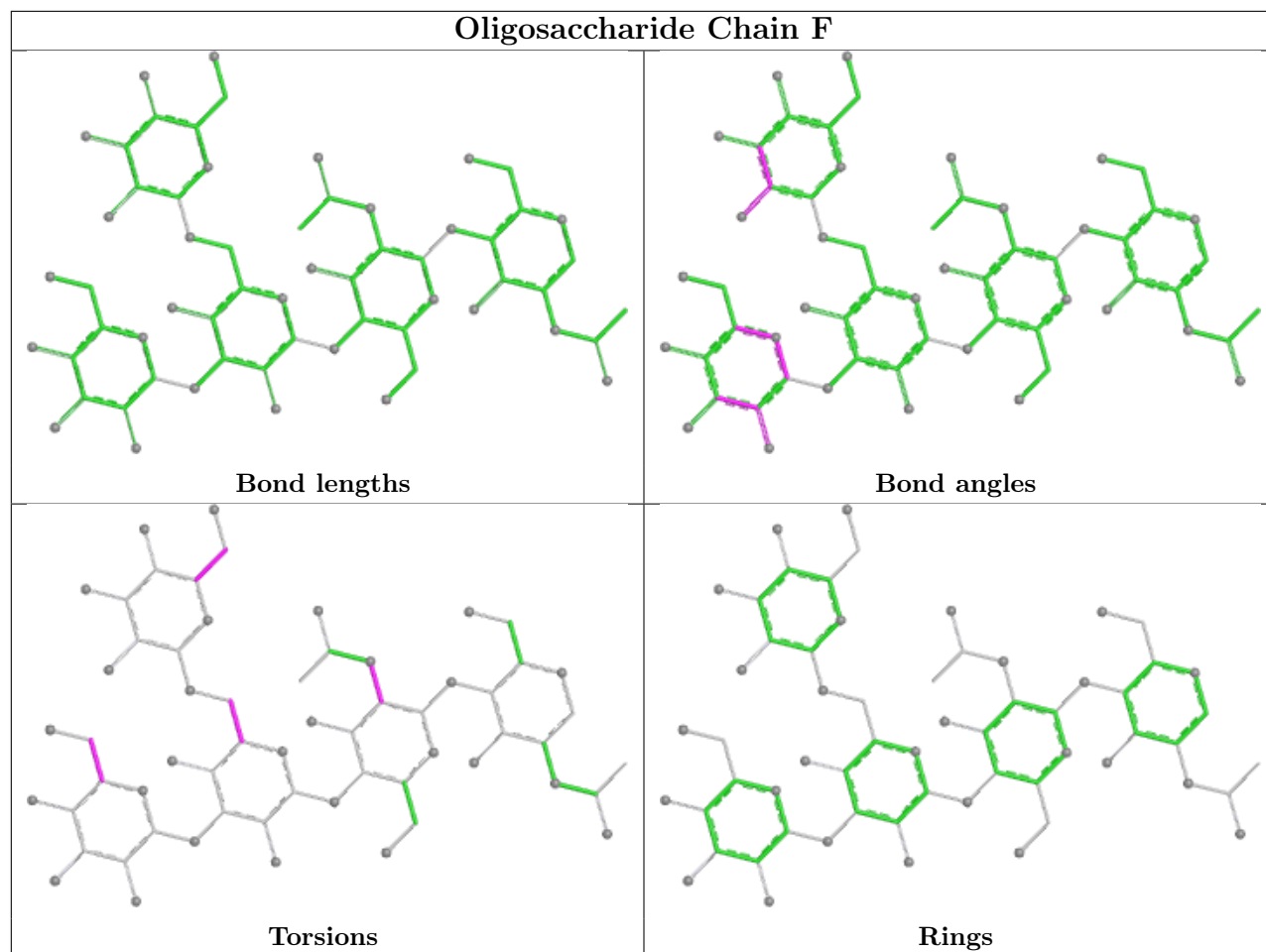
Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	1	NAG	1	0

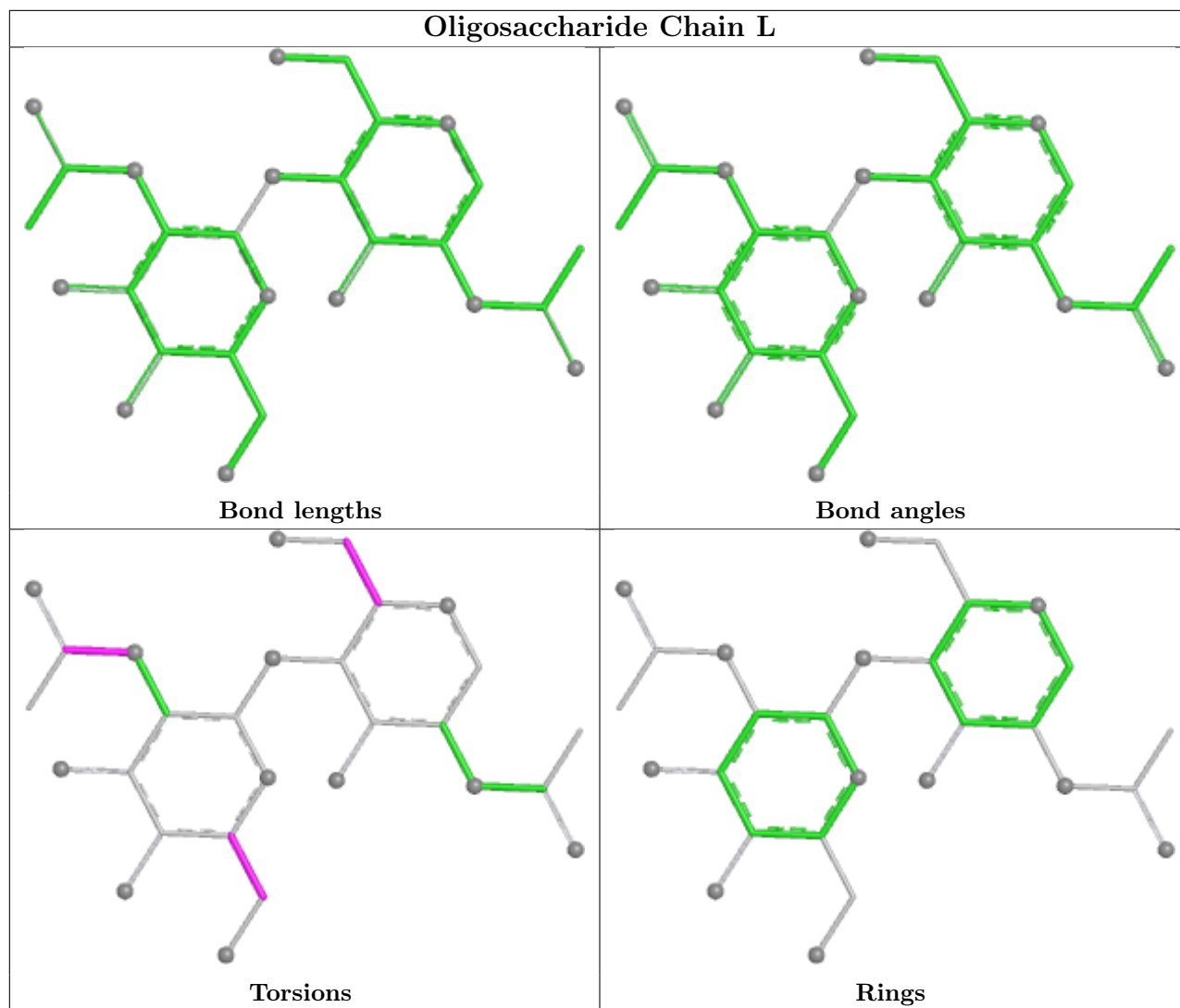
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.











## 5.6 Ligand geometry [i](#)

Of 20 ligands modelled in this entry, 7 are monoatomic - leaving 13 for Mogul analysis.

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	$\# Z  > 2$
15	ACT	B	2009	-	3,3,3	1.40	0	3,3,3	1.35	0
14	NAG	B	2006	2	14,14,15	0.29	0	17,19,21	0.41	0
10	MAN	A	2005	-	11,11,12	0.63	0	15,15,17	1.44	2 (13%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
11	GOL	A	2008	-	5,5,5	0.93	0	5,5,5	1.08	0
11	GOL	D	301	-	5,5,5	0.89	0	5,5,5	1.08	0
11	GOL	C	301	-	5,5,5	0.91	0	5,5,5	1.12	0
11	GOL	A	2007	-	5,5,5	0.89	0	5,5,5	1.09	0
13	A1A6B	B	2004	12	36,37,37	1.14	1 (2%)	37,51,51	1.91	4 (10%)
11	GOL	B	2007	-	5,5,5	0.93	0	5,5,5	1.08	0
14	NAG	B	2005	-	14,14,15	0.52	0	17,19,21	0.74	1 (5%)
11	GOL	A	2006	-	5,5,5	0.90	0	5,5,5	1.13	0
15	ACT	D	302	-	3,3,3	1.61	1 (33%)	3,3,3	1.28	0
11	GOL	B	2008	-	5,5,5	0.96	0	5,5,5	1.04	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
14	NAG	B	2006	2	-	3/6/23/26	0/1/1/1
10	MAN	A	2005	-	-	0/2/19/22	0/1/1/1
11	GOL	A	2008	-	-	4/4/4/4	-
11	GOL	D	301	-	-	4/4/4/4	-
11	GOL	C	301	-	-	0/4/4/4	-
11	GOL	A	2007	-	-	4/4/4/4	-
13	A1A6B	B	2004	12	-	3/24/42/42	0/5/5/5
11	GOL	B	2007	-	-	2/4/4/4	-
14	NAG	B	2005	-	-	0/6/23/26	0/1/1/1
11	GOL	A	2006	-	-	2/4/4/4	-
11	GOL	B	2008	-	-	2/4/4/4	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	B	2004	A1A6B	C31-N32	2.93	1.50	1.45
15	D	302	ACT	CH3-C	2.40	1.58	1.49

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	B	2004	A1A6B	C13-C11-C10	9.33	139.23	121.75

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
10	A	2005	MAN	C1-O5-C5	4.04	117.59	112.19
13	B	2004	A1A6B	C12-C11-C10	-3.12	115.90	121.75
13	B	2004	A1A6B	C28-C27-C26	-2.80	117.73	121.39
14	B	2005	NAG	C1-O5-C5	2.72	115.83	112.19

There are no chirality outliers.

5 of 24 torsion outliers are listed below:

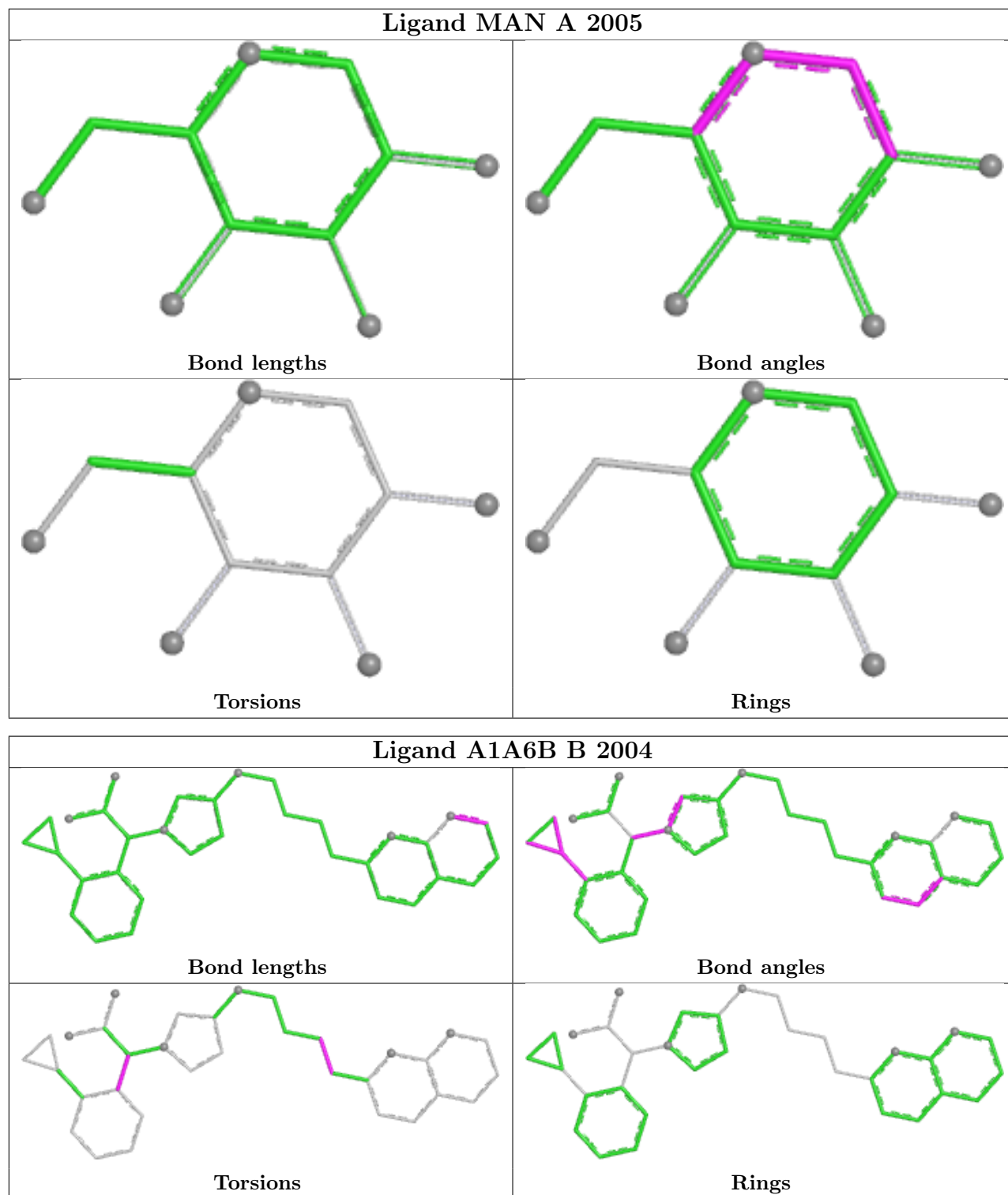
Mol	Chain	Res	Type	Atoms
11	A	2006	GOL	C1-C2-C3-O3
11	A	2006	GOL	O2-C2-C3-O3
11	A	2007	GOL	O1-C1-C2-C3
11	A	2007	GOL	C1-C2-C3-O3
11	A	2008	GOL	O1-C1-C2-O2

There are no ring outliers.

4 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
15	B	2009	ACT	1	0
14	B	2006	NAG	1	0
11	A	2008	GOL	1	0
15	D	302	ACT	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 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

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	595/605 (98%)	0.53	70 (11%) 10 10	25, 64, 116, 147	1 (0%)
2	B	466/481 (96%)	0.68	72 (15%) 6 6	44, 68, 126, 147	0
3	C	214/214 (100%)	0.64	15 (7%) 24 22	55, 75, 111, 138	0
4	D	213/218 (97%)	0.56	16 (7%) 22 20	51, 73, 102, 118	0
All	All	1488/1518 (98%)	0.60	173 (11%) 11 10	25, 69, 117, 147	1 (0%)

The worst 5 of 173 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	217	PRO	6.5
2	B	55	LEU	6.3
2	B	39	GLY	6.2
1	A	400	CYS	5.8
2	B	8	LEU	5.7

### 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](#)

SUGAR-RSR INFOmissingINFO

### 6.4 Ligands [i](#)

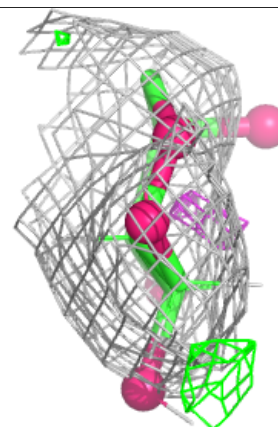
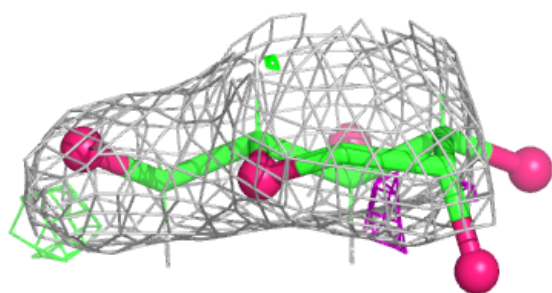
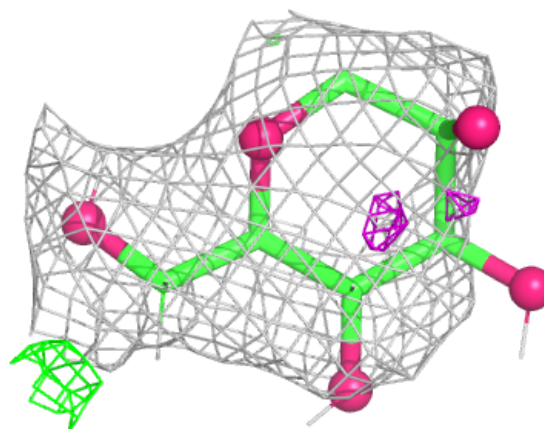
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
14	NAG	B	2005	14/15	0.26	0.24	98,112,130,135	0
14	NAG	B	2006	14/15	0.57	0.20	88,104,129,131	0
10	MAN	A	2005	11/12	0.65	0.17	73,92,111,114	0
11	GOL	B	2007	6/6	0.68	0.21	67,76,99,103	0
15	ACT	B	2009	4/4	0.80	0.22	71,71,71,71	0
15	ACT	D	302	4/4	0.81	0.24	53,54,56,58	0
11	GOL	C	301	6/6	0.82	0.21	66,69,83,83	0
11	GOL	A	2007	6/6	0.84	0.18	61,63,80,83	0
11	GOL	D	301	6/6	0.85	0.17	66,68,79,79	0
11	GOL	A	2008	6/6	0.88	0.16	59,62,73,73	0
11	GOL	B	2008	6/6	0.91	0.12	64,66,82,84	0
13	A1A6B	B	2004	33/33	0.96	0.13	47,55,67,67	0
11	GOL	A	2006	6/6	0.97	0.08	48,51,62,62	0
12	MG	B	2001	1/1	0.98	0.05	42,42,42,42	0
9	CA	A	2003	1/1	0.99	0.04	59,59,59,59	0
9	CA	A	2004	1/1	0.99	0.03	55,55,55,55	0
9	CA	B	2002	1/1	0.99	0.05	42,42,42,42	0
9	CA	B	2003	1/1	0.99	0.02	49,49,49,49	0
9	CA	A	2001	1/1	0.99	0.04	50,50,50,50	0
9	CA	A	2002	1/1	0.99	0.03	59,59,59,59	0

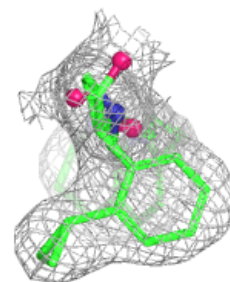
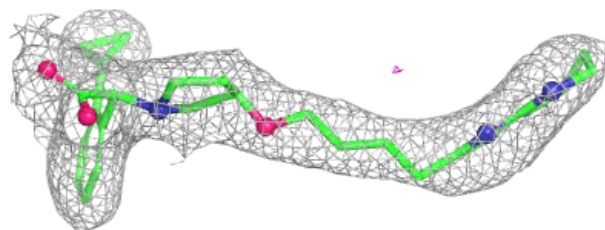
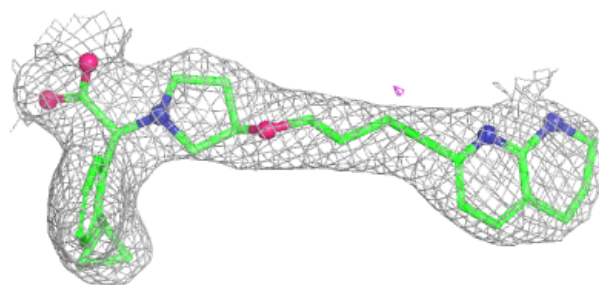
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around MAN A 2005:**

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

**Electron density around A1A6B B 2004:**

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



## 6.5 Other polymers [i](#)

There are no such residues in this entry.