



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 21, 2024 – 12:10 am GMT

PDB ID : 7PRS  
Title : Crystal Structure of the B subunit of heat labile enterotoxin LT-IIc from Escherichia coli in complex with Sialyl-lacto-N-neotetraose d  
Authors : Varrot, A.  
Deposited on : 2021-09-22  
Resolution : 2.00 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 : 1.8.4, CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

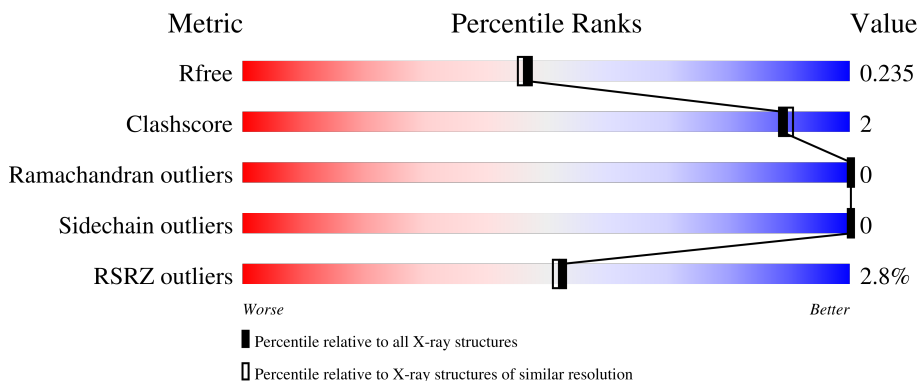
# 1 Overall quality at a glance

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	DDD	104	92% (5% poor fit)
1	EEE	104	91% (5% poor fit)
1	FFF	104	92% (7% poor fit)
1	GGG	104	90% (2% poor fit)
1	HHH	104	91% (0% poor fit)

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	III	104	<p>5% 91% 5% 5%</p>
1	JJJ	104	<p>% 88% 7% 5%</p>
1	KKK	104	<p>3% 91% 5% 5%</p>
1	LLL	104	<p>% 89% 6% 5%</p>
1	MMM	104	<p>3% 93% 5% 5%</p>
2	DaD	5	<p>100%</p>
3	GaG	4	<p>100%</p>
3	JaJ	4	<p>100%</p>
4	HaH	3	<p>100%</p>

## 2 Entry composition i

There are 7 unique types of molecules in this entry. The entry contains 8462 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 Heat-labile enterotoxin IIA, B chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	DDD	99	Total 750	C 474	N 127	O 145	S 4	0	0	0
1	EEE	99	Total 748	C 473	N 126	O 145	S 4	0	0	0
1	FFF	99	Total 745	C 470	N 125	O 146	S 4	0	1	0
1	GGG	99	Total 749	C 475	N 126	O 144	S 4	0	1	0
1	HHH	99	Total 748	C 473	N 126	O 145	S 4	0	0	0
1	III	99	Total 752	C 477	N 127	O 144	S 4	0	1	0
1	JJJ	99	Total 758	C 480	N 129	O 145	S 4	0	0	0
1	KKK	99	Total 748	C 473	N 126	O 145	S 4	0	0	0
1	LLL	99	Total 748	C 473	N 126	O 145	S 4	0	0	0
1	MMM	99	Total 750	C 474	N 127	O 145	S 4	0	0	0

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	99	HIS	-	expression tag	UNP H6W8F2
DDD	100	HIS	-	expression tag	UNP H6W8F2
DDD	101	HIS	-	expression tag	UNP H6W8F2
DDD	102	HIS	-	expression tag	UNP H6W8F2
DDD	103	HIS	-	expression tag	UNP H6W8F2
DDD	104	HIS	-	expression tag	UNP H6W8F2
EEE	99	HIS	-	expression tag	UNP H6W8F2
EEE	100	HIS	-	expression tag	UNP H6W8F2
EEE	101	HIS	-	expression tag	UNP H6W8F2

*Continued on next page...*

*Continued from previous page...*

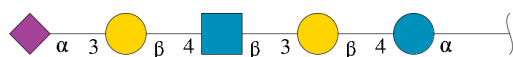
Chain	Residue	Modelled	Actual	Comment	Reference
EEE	102	HIS	-	expression tag	UNP H6W8F2
EEE	103	HIS	-	expression tag	UNP H6W8F2
EEE	104	HIS	-	expression tag	UNP H6W8F2
FFF	99	HIS	-	expression tag	UNP H6W8F2
FFF	100	HIS	-	expression tag	UNP H6W8F2
FFF	101	HIS	-	expression tag	UNP H6W8F2
FFF	102	HIS	-	expression tag	UNP H6W8F2
FFF	103	HIS	-	expression tag	UNP H6W8F2
FFF	104	HIS	-	expression tag	UNP H6W8F2
GGG	99	HIS	-	expression tag	UNP H6W8F2
GGG	100	HIS	-	expression tag	UNP H6W8F2
GGG	101	HIS	-	expression tag	UNP H6W8F2
GGG	102	HIS	-	expression tag	UNP H6W8F2
GGG	103	HIS	-	expression tag	UNP H6W8F2
GGG	104	HIS	-	expression tag	UNP H6W8F2
HHH	99	HIS	-	expression tag	UNP H6W8F2
HHH	100	HIS	-	expression tag	UNP H6W8F2
HHH	101	HIS	-	expression tag	UNP H6W8F2
HHH	102	HIS	-	expression tag	UNP H6W8F2
HHH	103	HIS	-	expression tag	UNP H6W8F2
HHH	104	HIS	-	expression tag	UNP H6W8F2
III	99	HIS	-	expression tag	UNP H6W8F2
III	100	HIS	-	expression tag	UNP H6W8F2
III	101	HIS	-	expression tag	UNP H6W8F2
III	102	HIS	-	expression tag	UNP H6W8F2
III	103	HIS	-	expression tag	UNP H6W8F2
III	104	HIS	-	expression tag	UNP H6W8F2
JJJ	99	HIS	-	expression tag	UNP H6W8F2
JJJ	100	HIS	-	expression tag	UNP H6W8F2
JJJ	101	HIS	-	expression tag	UNP H6W8F2
JJJ	102	HIS	-	expression tag	UNP H6W8F2
JJJ	103	HIS	-	expression tag	UNP H6W8F2
JJJ	104	HIS	-	expression tag	UNP H6W8F2
KKK	99	HIS	-	expression tag	UNP H6W8F2
KKK	100	HIS	-	expression tag	UNP H6W8F2
KKK	101	HIS	-	expression tag	UNP H6W8F2
KKK	102	HIS	-	expression tag	UNP H6W8F2
KKK	103	HIS	-	expression tag	UNP H6W8F2
KKK	104	HIS	-	expression tag	UNP H6W8F2
LLL	99	HIS	-	expression tag	UNP H6W8F2
LLL	100	HIS	-	expression tag	UNP H6W8F2
LLL	101	HIS	-	expression tag	UNP H6W8F2

*Continued on next page...*

Continued from previous page...

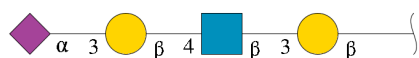
Chain	Residue	Modelled	Actual	Comment	Reference
LLL	102	HIS	-	expression tag	UNP H6W8F2
LLL	103	HIS	-	expression tag	UNP H6W8F2
LLL	104	HIS	-	expression tag	UNP H6W8F2
MMM	99	HIS	-	expression tag	UNP H6W8F2
MMM	100	HIS	-	expression tag	UNP H6W8F2
MMM	101	HIS	-	expression tag	UNP H6W8F2
MMM	102	HIS	-	expression tag	UNP H6W8F2
MMM	103	HIS	-	expression tag	UNP H6W8F2
MMM	104	HIS	-	expression tag	UNP H6W8F2

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose.



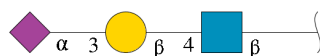
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	DaD	5	68	37	2	29	0	0	0

- Molecule 3 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose.



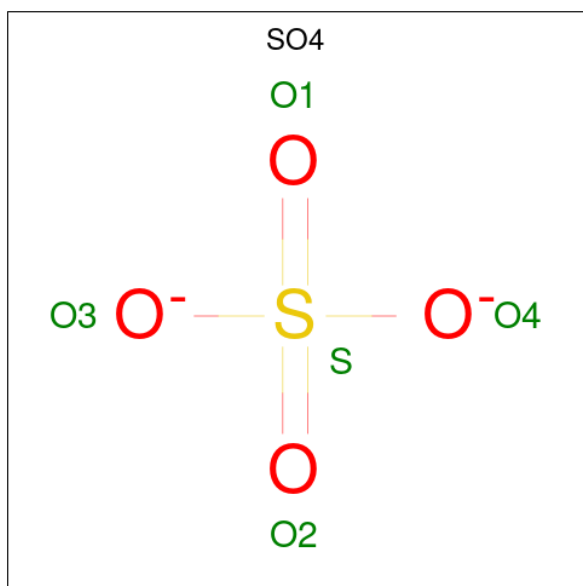
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	GaG	4	57	31	2	24	0	0	0
3	JaJ	4	57	31	2	24	0	0	0

- Molecule 4 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



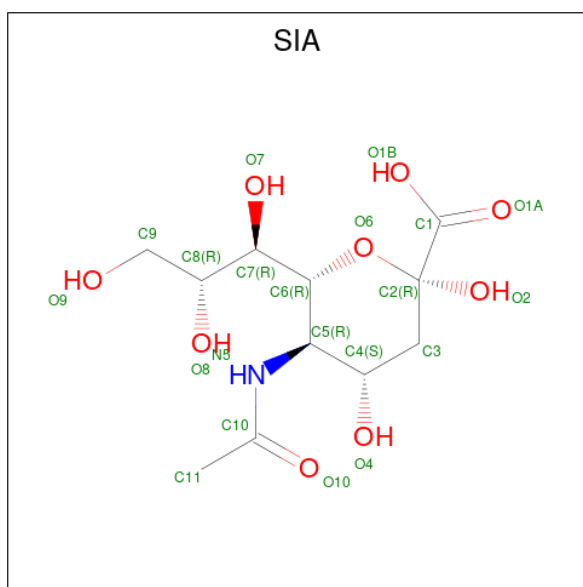
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	HaH	3	45	25	2	18	0	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	S		
5	DDD	1	5	4	1	0	0
5	EEE	1	5	4	1	0	0
5	FFF	1	5	4	1	0	0
5	HHH	1	5	4	1	0	0

- Molecule 6 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: C<sub>11</sub>H<sub>19</sub>NO<sub>9</sub>) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
6	III	1	21	11	1	9	0	0
6	JJJ	1	21	11	1	9	0	0
6	KKK	1	21	11	1	9	0	0
6	LLL	1	20	11	1	8	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	DDD	71	73	73	0	2
7	EEE	61	64	64	0	3
7	FFF	45	46	46	0	1
7	GGG	56	58	58	0	2
7	HHH	73	75	75	0	2
7	III	57	59	59	0	2
7	JJJ	68	69	69	0	1
7	KKK	69	70	70	0	1

*Continued on next page...*



*Continued from previous page...*

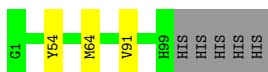
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
7	LLL	54	Total	O	0	2
			56	56		
7	MMM	62	Total	O	0	4
			66	66		

### 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.

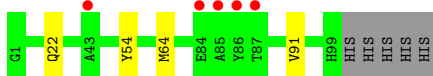
- Molecule 1: Heat-labile enterotoxin IIA, B chain

Chain DDD:  92% 5%



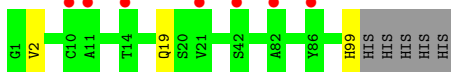
- Molecule 1: Heat-labile enterotoxin IIA, B chain

Chain EEE:  91% 5%



- Molecule 1: Heat-labile enterotoxin IIA, B chain

Chain FFF:  92% 5%



- Molecule 1: Heat-labile enterotoxin IIA, B chain

Chain GGG:  90% 5% 5%

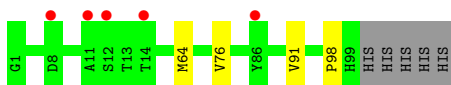
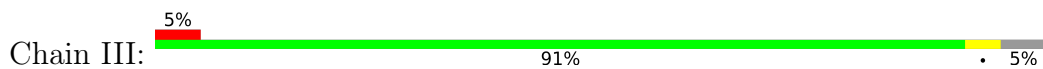


- Molecule 1: Heat-labile enterotoxin IIA, B chain

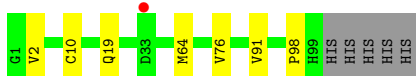
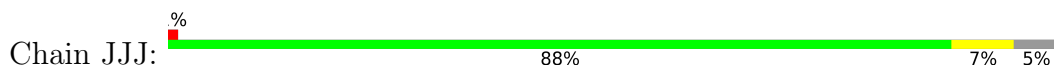
Chain HHH:  91% 5%



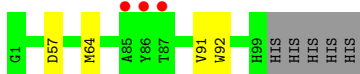
- Molecule 1: Heat-labile enterotoxin IIA, B chain



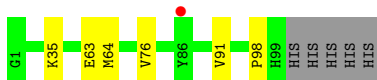
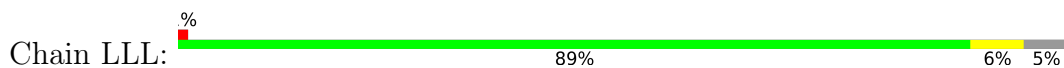
- Molecule 1: Heat-labile enterotoxin IIA, B chain



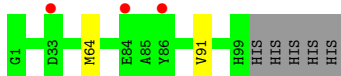
- Molecule 1: Heat-labile enterotoxin IIA, B chain



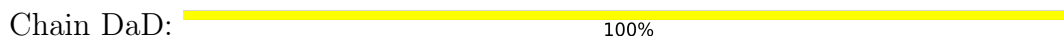
- Molecule 1: Heat-labile enterotoxin IIA, B chain



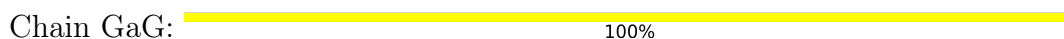
- Molecule 1: Heat-labile enterotoxin IIA, B chain



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose-(1-4)-alpha-D-glucopyranose



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose



- Molecule 3: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-3)-beta-D-galactopyranose

Chain JaJ:  100%

GAL1  
NAG2  
GAL3  
SIA4

- Molecule 4: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain HaH:  100%

NAG1  
GAL2  
SIA3

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.09Å 121.27Å 126.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.54 – 2.00 38.51 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (38.54-2.00) 100.0 (38.51-2.00)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.33 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.188 , 0.230 0.196 , 0.235	Depositor DCC
$R_{free}$ test set	4010 reflections (5.23%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.0	Xtrriage
Anisotropy	0.388	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 49.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for -h,l,k	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8462	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	39.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 46.63 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.1140e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: SIA, GLC, SO4, GAL, NAG

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	DDD	0.90	0/766	0.93	1/1043 (0.1%)
1	EEE	0.87	0/764	0.90	1/1041 (0.1%)
1	FFF	0.84	0/761	0.93	1/1041 (0.1%)
1	GGG	0.90	1/765 (0.1%)	0.91	0/1044
1	HHH	0.92	1/764 (0.1%)	0.93	1/1041 (0.1%)
1	III	0.88	0/768	0.90	0/1047
1	JJJ	0.87	0/774	0.92	0/1051
1	KKK	0.86	0/764	0.91	0/1041
1	LLL	0.90	1/764 (0.1%)	0.91	0/1041
1	MMM	0.92	0/766	0.96	0/1043
All	All	0.89	3/7656 (0.0%)	0.92	4/10433 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	GGG	63	GLU	CD-OE1	-8.51	1.16	1.25
1	HHH	84	GLU	CD-OE1	-6.88	1.18	1.25
1	LLL	63	GLU	CD-OE1	-5.11	1.20	1.25

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	EEE	54	TYR	CB-CG-CD1	-5.84	117.49	121.00
1	DDD	54	TYR	CB-CG-CD2	-5.80	117.52	121.00
1	HHH	54	TYR	CB-CG-CD1	-5.70	117.58	121.00
1	FFF	99	HIS	CA-C-O	5.25	131.13	120.10

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	DDD	750	0	734	2	0
1	EEE	748	0	727	3	0
1	FFF	745	0	710	1	0
1	GGG	749	0	724	3	0
1	HHH	748	0	727	2	0
1	III	752	0	733	3	0
1	JJJ	758	0	756	5	0
1	KKK	748	0	727	4	0
1	LLL	748	0	727	4	0
1	MMM	750	0	734	2	0
2	DaD	68	0	58	0	0
3	GaG	57	0	49	0	0
3	JaJ	57	0	49	0	0
4	HaH	45	0	38	0	0
5	DDD	5	0	0	0	0
5	EEE	5	0	0	0	0
5	FFF	5	0	0	0	0
5	HHH	5	0	0	0	0
6	III	21	0	18	0	0
6	JJJ	21	0	18	0	0
6	KKK	21	0	18	1	0
6	LLL	20	0	17	0	0
7	DDD	73	0	0	0	0
7	EEE	64	0	0	1	0
7	FFF	46	0	0	0	0
7	GGG	58	0	0	0	0
7	HHH	75	0	0	0	0
7	III	59	0	0	0	0
7	JJJ	69	0	0	1	0
7	KKK	70	0	0	0	0
7	LLL	56	0	0	0	0
7	MMM	66	0	0	0	0
All	All	8462	0	7564	28	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 28 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:KKK:57:ASP:OD1	1:LLL:35:LYS:HE3	1.84	0.76
1:JJJ:76:VAL:HG13	1:JJJ:98:PRO:HA	1.76	0.67
1:GGG:64:MET:HE1	1:GGG:91:VAL:HG12	1.76	0.66
1:EEE:64:MET:CE	1:EEE:91:VAL:HG12	2.27	0.64
1:KKK:64:MET:CE	1:KKK:91:VAL:HG12	2.28	0.62

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	DDD	97/104 (93%)	97 (100%)	0	0	100	100
1	EEE	97/104 (93%)	97 (100%)	0	0	100	100
1	FFF	98/104 (94%)	98 (100%)	0	0	100	100
1	GGG	98/104 (94%)	98 (100%)	0	0	100	100
1	HHH	97/104 (93%)	97 (100%)	0	0	100	100
1	III	98/104 (94%)	98 (100%)	0	0	100	100
1	JJJ	97/104 (93%)	97 (100%)	0	0	100	100
1	KKK	97/104 (93%)	97 (100%)	0	0	100	100
1	LLL	97/104 (93%)	97 (100%)	0	0	100	100
1	MMM	97/104 (93%)	97 (100%)	0	0	100	100
All	All	973/1040 (94%)	973 (100%)	0	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	DDD	82/89 (92%)	82 (100%)	0	100	100
1	EEE	81/89 (91%)	81 (100%)	0	100	100
1	FFF	79/89 (89%)	79 (100%)	0	100	100
1	GGG	80/89 (90%)	80 (100%)	0	100	100
1	HHH	81/89 (91%)	81 (100%)	0	100	100
1	III	81/89 (91%)	81 (100%)	0	100	100
1	JJJ	84/89 (94%)	84 (100%)	0	100	100
1	KKK	81/89 (91%)	81 (100%)	0	100	100
1	LLL	81/89 (91%)	81 (100%)	0	100	100
1	MMM	82/89 (92%)	82 (100%)	0	100	100
All	All	812/890 (91%)	812 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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](#)

16 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
2	GLC	DaD	1	2	12,12,12	0.44	0	17,17,17	1.32	2 (11%)
2	GAL	DaD	2	2	11,11,12	0.47	0	15,15,17	1.18	1 (6%)
2	NAG	DaD	3	2	14,14,15	1.14	1 (7%)	17,19,21	1.35	3 (17%)
2	GAL	DaD	4	2	11,11,12	1.01	0	15,15,17	1.85	5 (33%)
2	SIA	DaD	5	2	20,20,21	0.77	1 (5%)	24,28,31	1.26	3 (12%)
3	GAL	GaG	1	3	12,12,12	0.82	0	17,17,17	1.37	4 (23%)
3	NAG	GaG	2	3	14,14,15	0.95	0	17,19,21	1.40	2 (11%)
3	GAL	GaG	3	3	11,11,12	0.93	0	15,15,17	1.75	4 (26%)
3	SIA	GaG	4	3	20,20,21	0.62	1 (5%)	24,28,31	1.16	2 (8%)
4	NAG	HaH	1	4	14,14,15	1.42	1 (7%)	17,19,21	1.90	4 (23%)
4	GAL	HaH	2	4	11,11,12	0.85	0	15,15,17	1.95	3 (20%)
4	SIA	HaH	3	4	20,20,21	0.72	1 (5%)	24,28,31	1.30	4 (16%)
3	GAL	JaJ	1	3	12,12,12	0.61	0	17,17,17	1.29	3 (17%)
3	NAG	JaJ	2	3	14,14,15	1.22	1 (7%)	17,19,21	1.70	4 (23%)
3	GAL	JaJ	3	3	11,11,12	0.96	1 (9%)	15,15,17	1.29	2 (13%)
3	SIA	JaJ	4	3	20,20,21	1.15	2 (10%)	24,28,31	0.93	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
2	GLC	DaD	1	2	-	0/2/22/22	0/1/1/1
2	GAL	DaD	2	2	-	1/2/19/22	0/1/1/1
2	NAG	DaD	3	2	-	0/6/23/26	0/1/1/1
2	GAL	DaD	4	2	-	2/2/19/22	0/1/1/1
2	SIA	DaD	5	2	-	5/18/34/38	0/1/1/1
3	GAL	GaG	1	3	-	1/2/22/22	0/1/1/1
3	NAG	GaG	2	3	-	0/6/23/26	0/1/1/1
3	GAL	GaG	3	3	-	2/2/19/22	0/1/1/1
3	SIA	GaG	4	3	-	0/18/34/38	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	HaH	1	4	-	0/6/23/26	0/1/1/1
4	GAL	HaH	2	4	-	2/2/19/22	0/1/1/1
4	SIA	HaH	3	4	-	4/18/34/38	0/1/1/1
3	GAL	JaJ	1	3	-	2/2/22/22	0/1/1/1
3	NAG	JaJ	2	3	-	0/6/23/26	0/1/1/1
3	GAL	JaJ	3	3	-	0/2/19/22	0/1/1/1
3	SIA	JaJ	4	3	-	3/18/34/38	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	JaJ	4	SIA	O1B-C1	-3.73	1.18	1.30
2	DaD	3	NAG	C1-C2	3.15	1.57	1.52
3	JaJ	4	SIA	C2-C1	2.92	1.55	1.52
4	HaH	1	NAG	C1-C2	2.59	1.56	1.52
3	GaG	4	SIA	O1B-C1	-2.20	1.23	1.30

The worst 5 of 46 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	HaH	2	GAL	C1-O5-C5	5.87	120.14	112.19
4	HaH	1	NAG	O5-C1-C2	-4.54	104.11	111.29
2	DaD	1	GLC	C4-C3-C2	-3.97	103.90	110.82
4	HaH	1	NAG	O5-C5-C6	3.60	112.84	107.20
3	JaJ	2	NAG	O5-C5-C6	3.53	112.73	107.20

There are no chirality outliers.

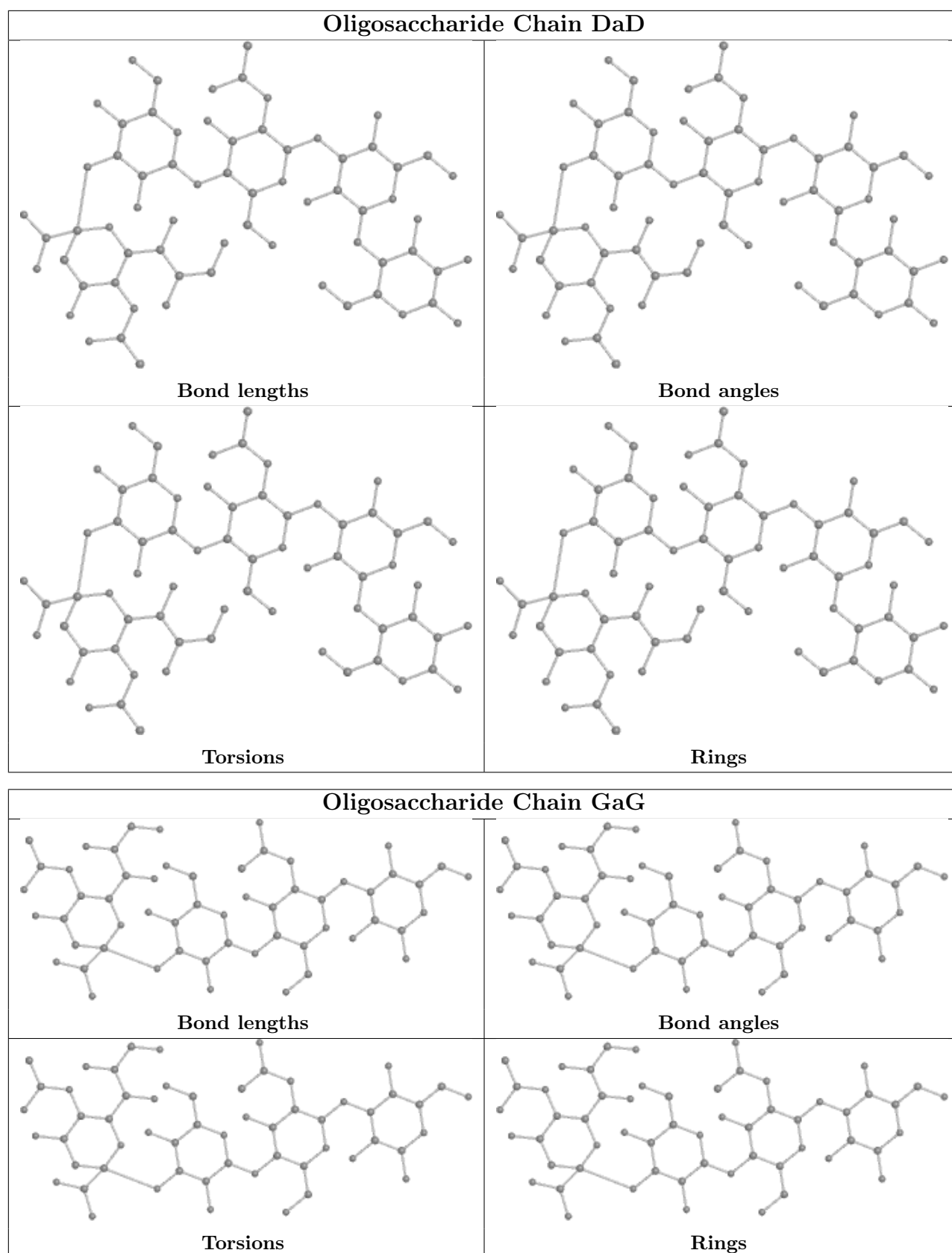
5 of 22 torsion outliers are listed below:

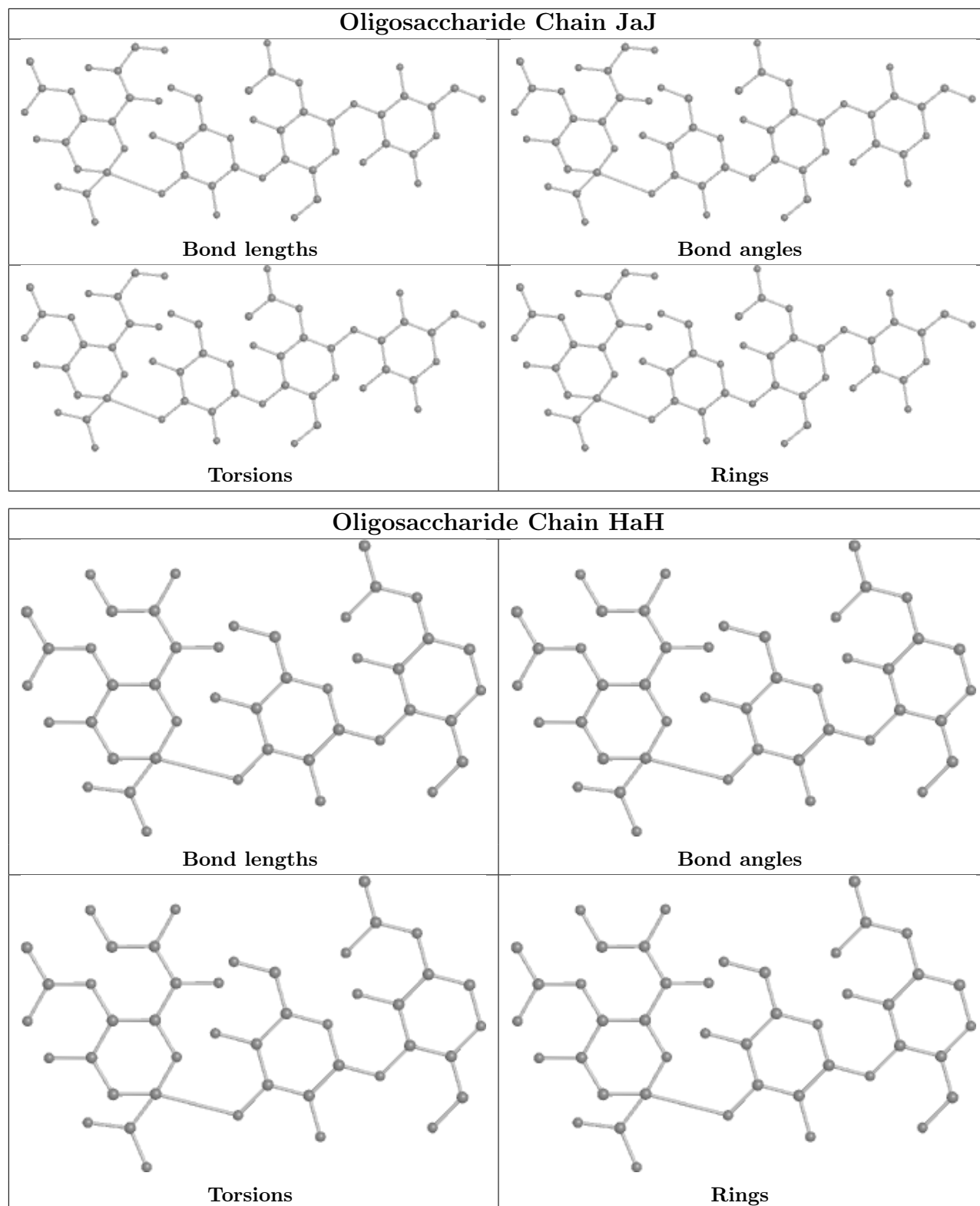
Mol	Chain	Res	Type	Atoms
4	HaH	2	GAL	O5-C5-C6-O6
4	HaH	2	GAL	C4-C5-C6-O6
3	JaJ	1	GAL	O5-C5-C6-O6
2	DaD	4	GAL	O5-C5-C6-O6
3	GaG	3	GAL	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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](#)

8 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	SIA	III	201	-	21,21,21	0.59	0	25,31,31	1.01	1 (4%)
5	SO4	FFF	201	-	4,4,4	0.32	0	6,6,6	0.49	0
5	SO4	DDD	201	-	4,4,4	0.46	0	6,6,6	0.22	0
6	SIA	LLL	201	-	20,20,21	0.81	1 (5%)	24,28,31	0.85	1 (4%)
5	SO4	HHH	201	-	4,4,4	0.52	0	6,6,6	0.27	0
6	SIA	JJJ	201	-	21,21,21	0.71	1 (4%)	25,31,31	0.96	2 (8%)
6	SIA	KKK	201	-	21,21,21	0.73	1 (4%)	25,31,31	0.97	1 (4%)
5	SO4	EEE	201	-	4,4,4	0.45	0	6,6,6	0.12	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	SIA	III	201	-	-	7/20/38/38	0/1/1/1
6	SIA	KKK	201	-	-	6/20/38/38	0/1/1/1
6	SIA	LLL	201	-	-	8/18/34/38	0/1/1/1
6	SIA	JJJ	201	-	-	8/20/38/38	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	LLL	201	SIA	O1B-C1	-2.65	1.21	1.30
6	JJJ	201	SIA	O1B-C1	-2.24	1.21	1.30
6	KKK	201	SIA	O1B-C1	-2.18	1.22	1.30

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	III	201	SIA	O1A-C1-C2	-4.05	117.46	123.59
6	KKK	201	SIA	O1A-C1-C2	-3.12	118.87	123.59
6	JJJ	201	SIA	O1A-C1-C2	-2.58	119.69	123.59

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	JJJ	201	SIA	O2-C2-O6	-2.36	103.42	109.42
6	LLL	201	SIA	O1B-C1-C2	2.27	119.50	113.03

There are no chirality outliers.

5 of 29 torsion outliers are listed below:

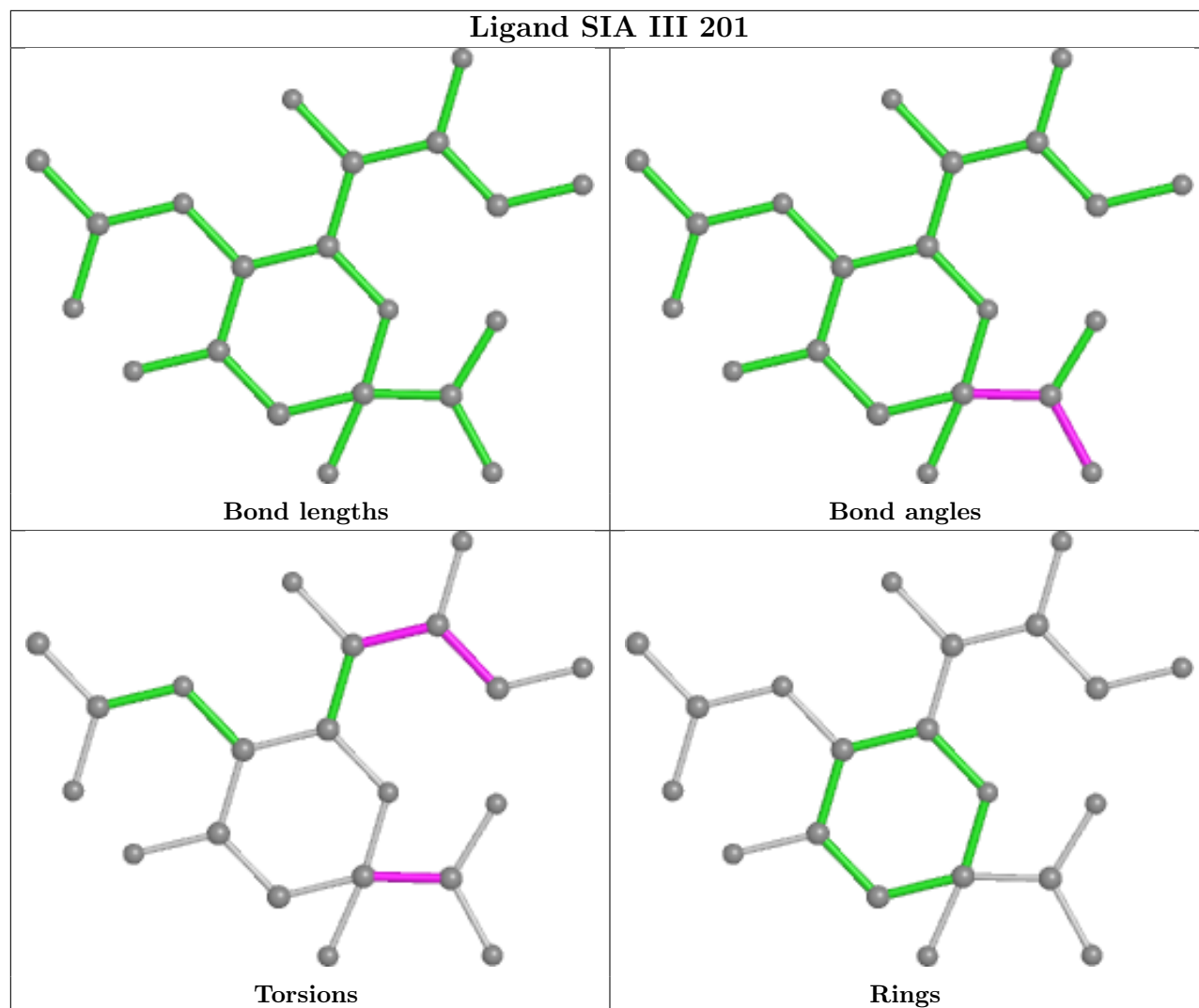
Mol	Chain	Res	Type	Atoms
6	III	201	SIA	C6-C7-C8-O8
6	III	201	SIA	O7-C7-C8-C9
6	III	201	SIA	O7-C7-C8-O8
6	III	201	SIA	C7-C8-C9-O9
6	III	201	SIA	O8-C8-C9-O9

There are no ring outliers.

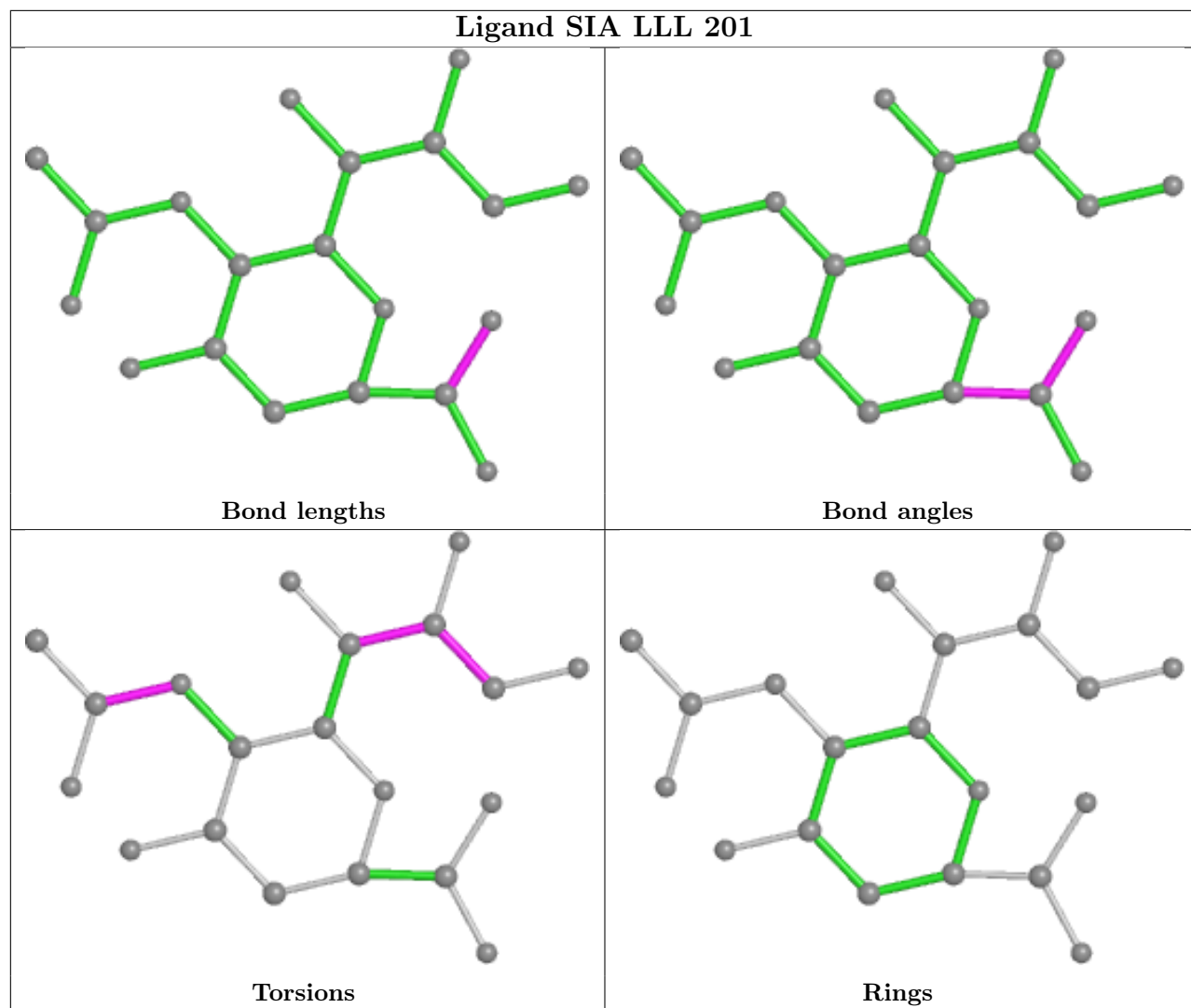
1 monomer is involved in 1 short contact:

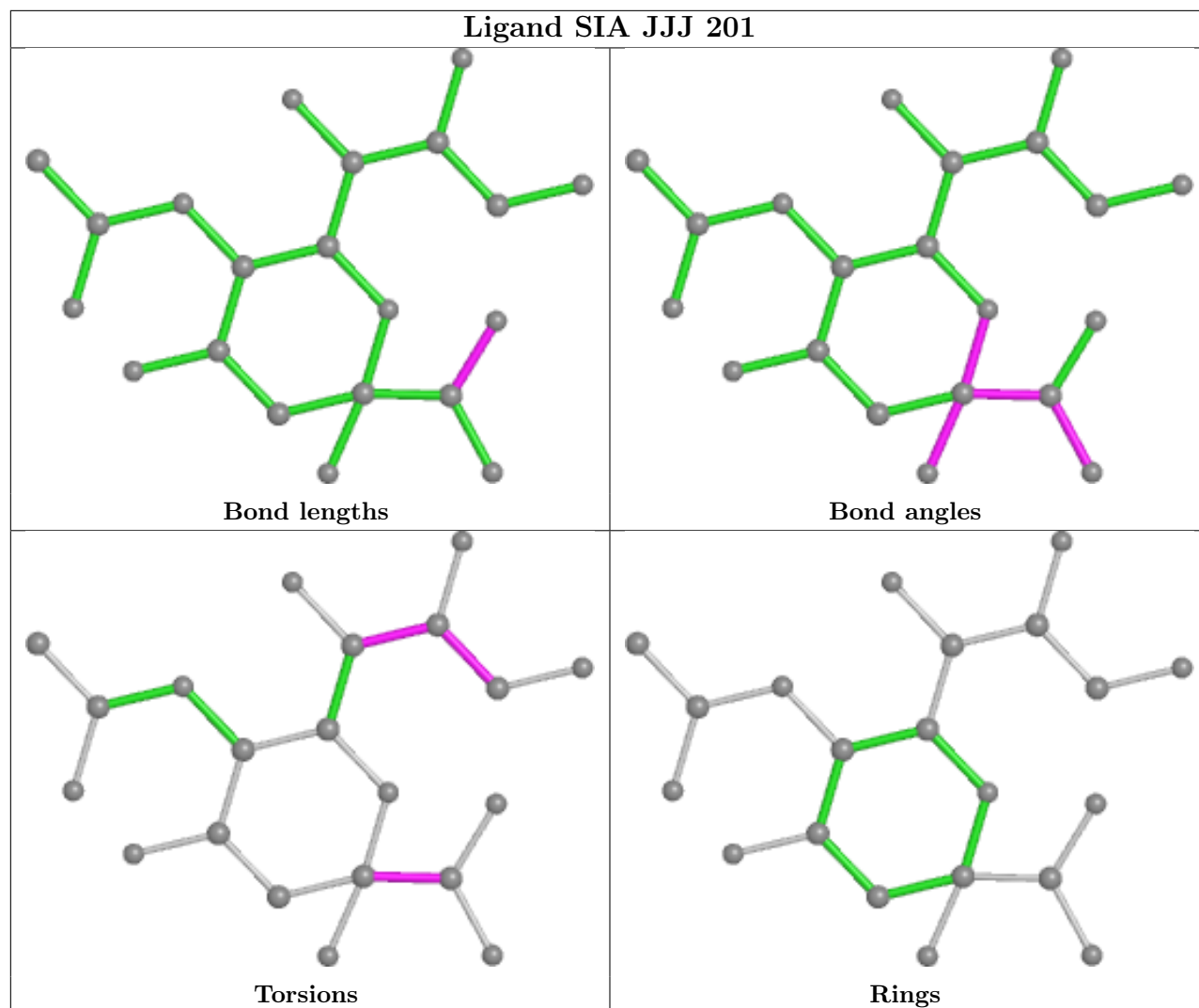
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	KKK	201	SIA	1	0

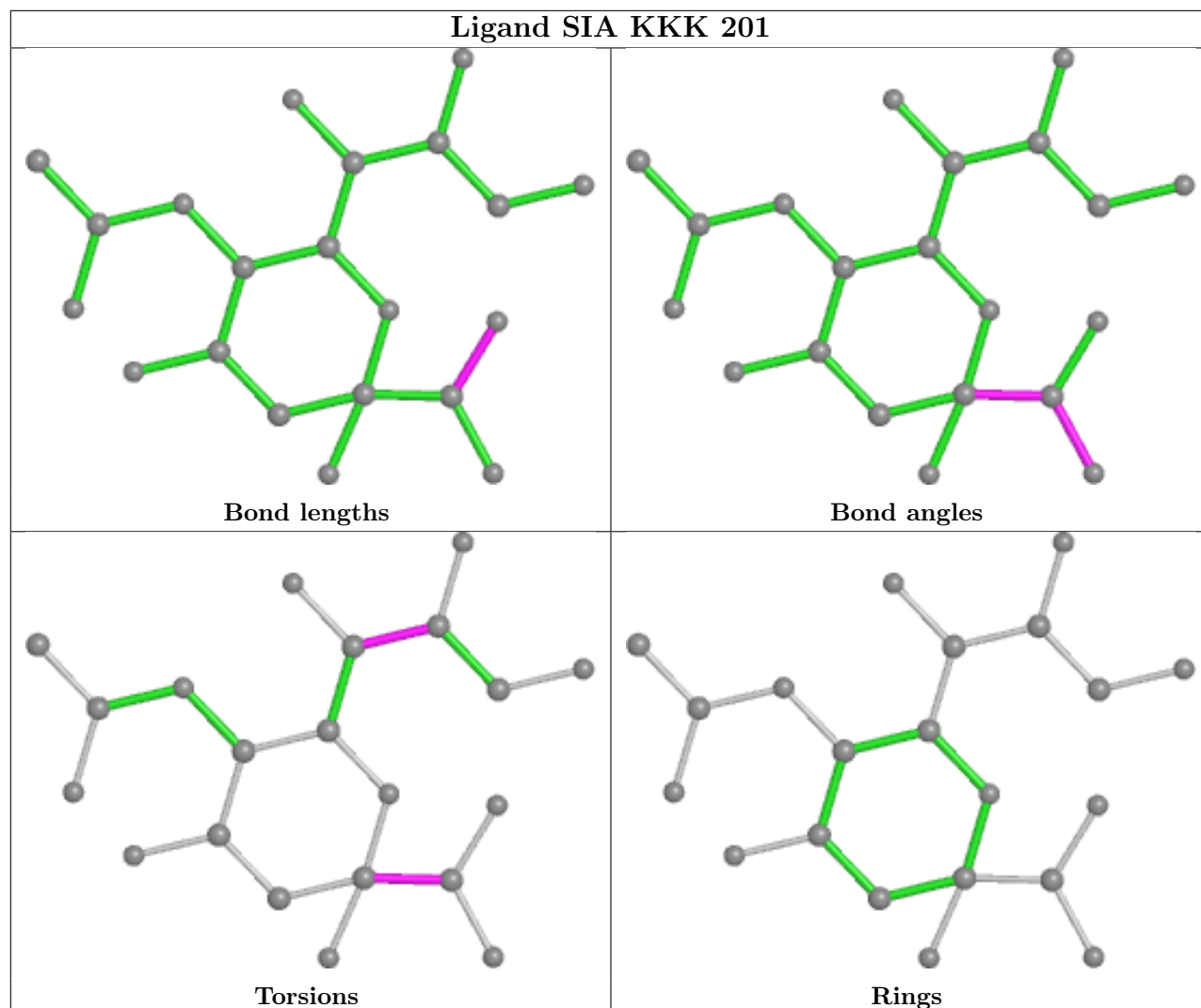
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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 [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	DDD	99/104 (95%)	-0.07	0 <b>100</b> <b>100</b>	22, 33, 47, 53	0
1	EEE	99/104 (95%)	0.24	5 (5%) 28 27	24, 37, 56, 60	2 (2%)
1	FFF	99/104 (95%)	0.49	7 (7%) 16 15	28, 42, 59, 71	0
1	GGG	99/104 (95%)	0.15	2 (2%) 65 63	26, 40, 57, 63	0
1	HHH	99/104 (95%)	-0.01	1 (1%) 82 81	23, 34, 48, 56	0
1	III	99/104 (95%)	0.36	5 (5%) 28 27	25, 38, 58, 64	0
1	JJJ	99/104 (95%)	0.09	1 (1%) 82 81	25, 36, 52, 67	0
1	KKK	99/104 (95%)	0.13	3 (3%) 50 49	25, 36, 53, 59	0
1	LLL	99/104 (95%)	0.04	1 (1%) 82 81	23, 35, 54, 61	0
1	MMM	99/104 (95%)	0.06	3 (3%) 50 49	23, 34, 49, 60	0
All	All	990/1040 (95%)	0.15	28 (2%) 53 51	22, 37, 55, 71	2 (0%)

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	FFF	11	ALA	4.0
1	MMM	86	TYR	3.3
1	EEE	86	TYR	3.2
1	EEE	84	GLU	3.1
1	FFF	86	TYR	3.0

### 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

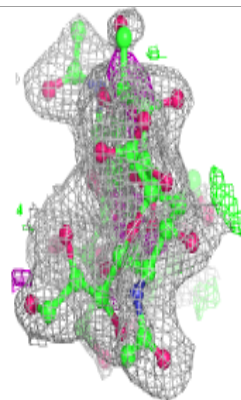
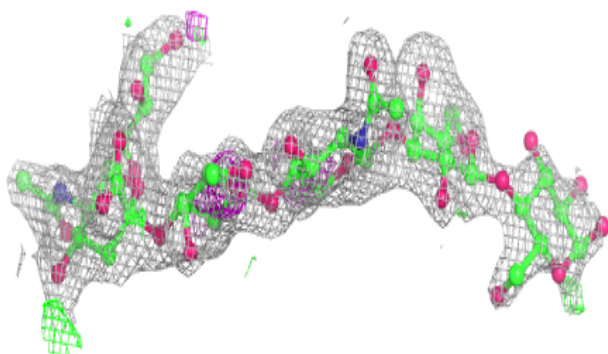
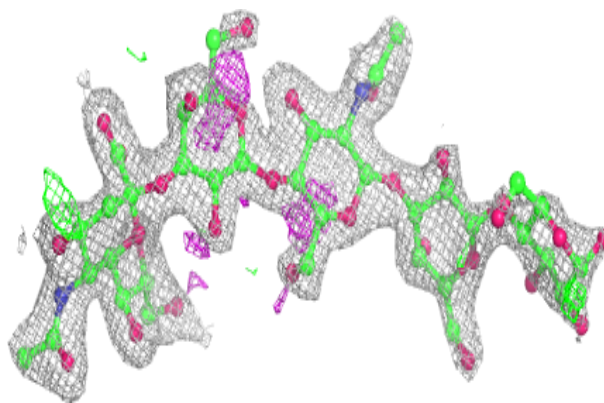
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
4	NAG	HaH	1	14/15	0.71	0.27	46,52,58,64	0
4	GAL	HaH	2	11/12	0.77	0.25	42,50,62,63	0
3	GAL	GaG	3	11/12	0.79	0.34	49,55,69,71	0
2	GLC	DaD	1	12/12	0.80	0.23	49,52,55,57	12
3	GAL	JaJ	1	12/12	0.81	0.18	40,54,64,67	12
2	GAL	DaD	2	11/12	0.85	0.13	42,54,61,61	11
3	GAL	GaG	1	12/12	0.86	0.20	44,69,80,81	0
2	GAL	DaD	4	11/12	0.87	0.32	39,50,59,61	0
3	GAL	JaJ	3	11/12	0.88	0.19	39,43,47,47	0
3	SIA	JaJ	4	20/21	0.88	0.13	33,42,45,48	0
2	NAG	DaD	3	14/15	0.89	0.21	35,42,47,51	0
4	SIA	HaH	3	20/21	0.89	0.15	39,46,52,52	0
3	SIA	GaG	4	20/21	0.91	0.17	39,44,50,55	0
3	NAG	JaJ	2	14/15	0.92	0.17	33,38,44,47	0
3	NAG	GaG	2	14/15	0.92	0.27	34,44,49,50	0
2	SIA	DaD	5	20/21	0.92	0.17	35,42,49,50	0

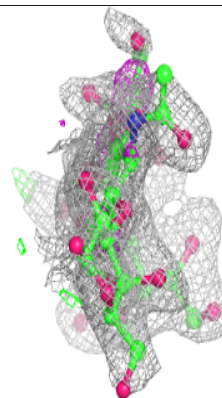
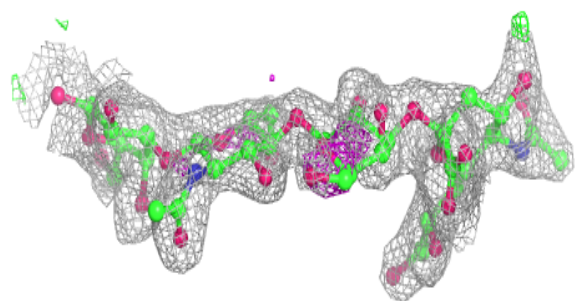
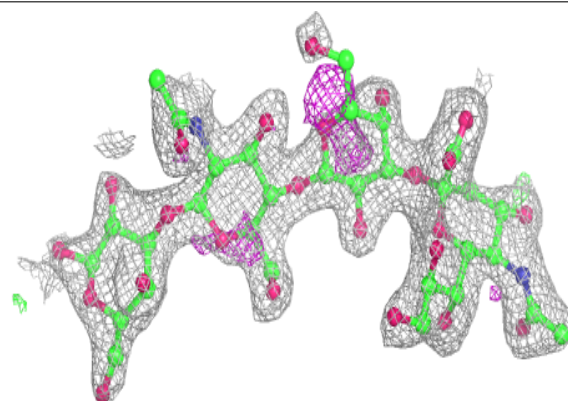
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain DaD:**

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

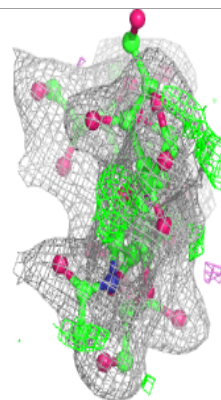
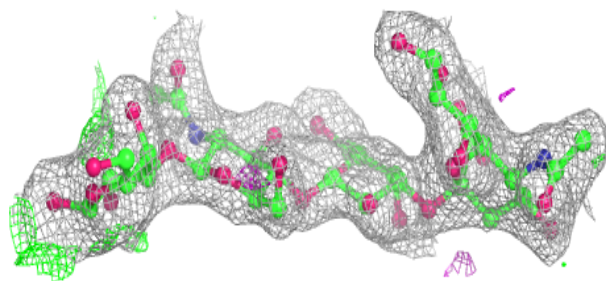
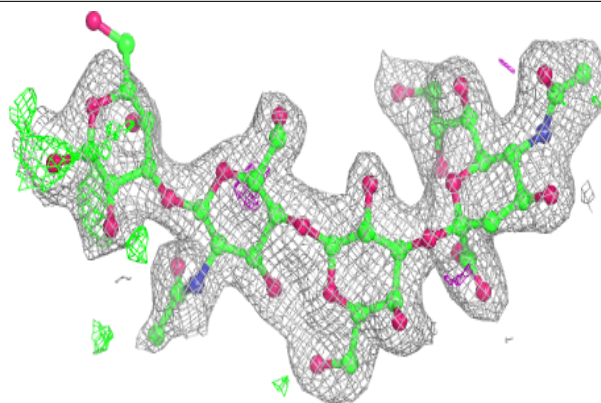
**Electron density around Chain GaG:**

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

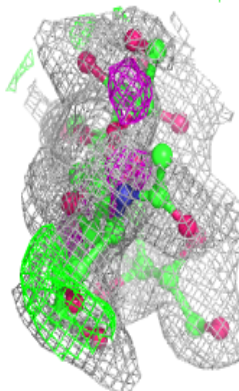
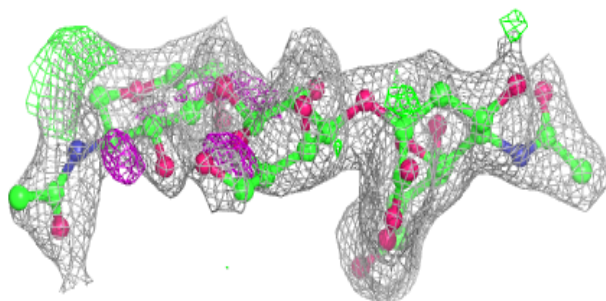
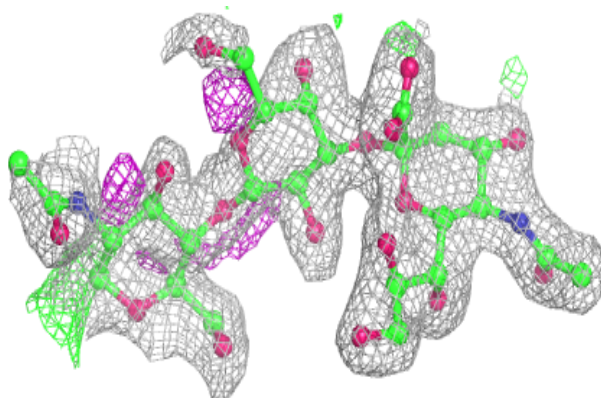


**Electron density around Chain JaJ:**

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

**Electron density around Chain HaH:**

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





## 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.

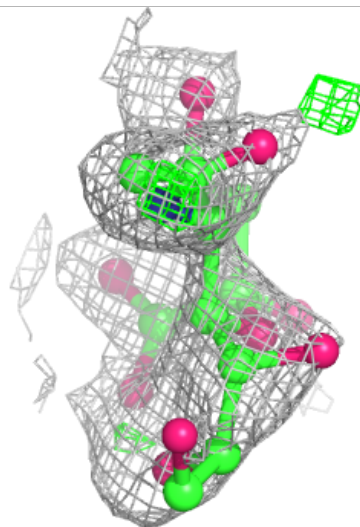
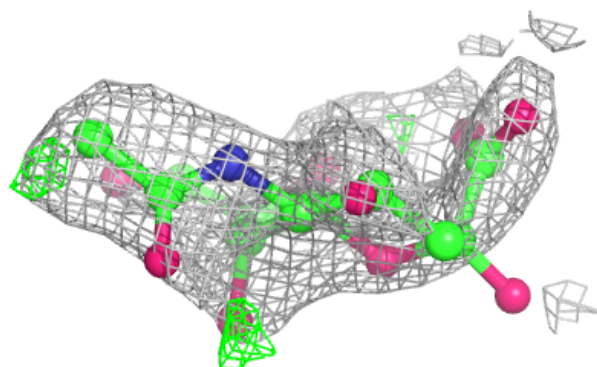
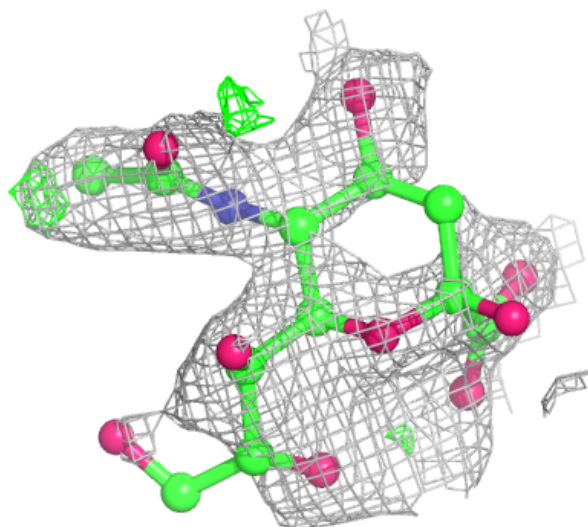
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	SIA	III	201	21/21	0.58	0.34	45,57,62,69	21
6	SIA	JJJ	201	21/21	0.65	0.28	51,63,67,69	20
5	SO4	FFF	201	5/5	0.69	0.32	31,35,46,48	5
6	SIA	LLL	201	20/21	0.78	0.31	47,58,62,68	20
6	SIA	KKK	201	21/21	0.79	0.29	56,63,70,95	0
5	SO4	EEE	201	5/5	0.92	0.36	43,44,48,58	5
5	SO4	DDD	201	5/5	0.95	0.33	27,31,39,44	5
5	SO4	HHH	201	5/5	0.97	0.13	27,29,42,55	5

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.



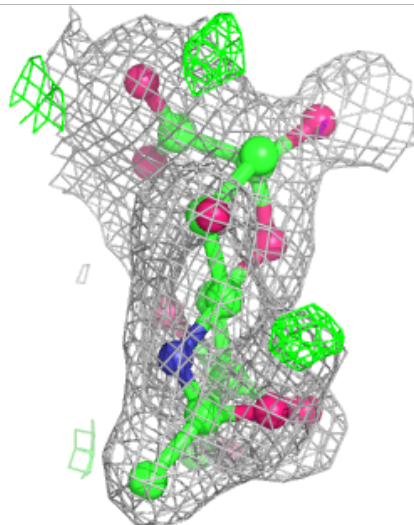
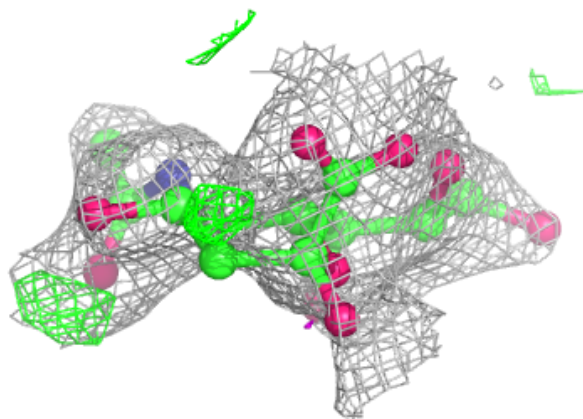
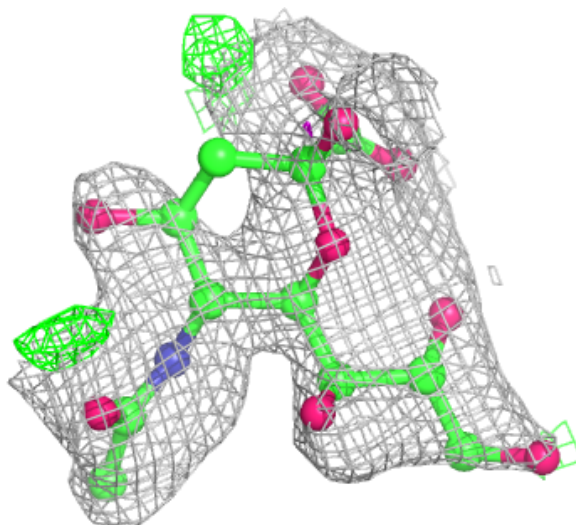
**Electron density around SIA III 201:**

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



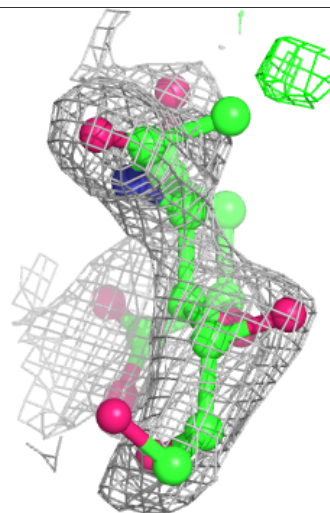
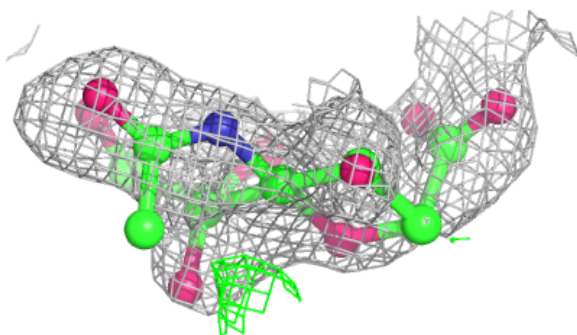
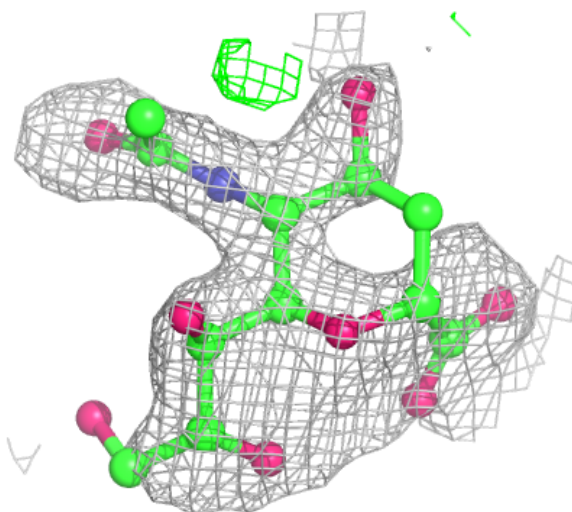
**Electron density around SIA JJJ 201:**

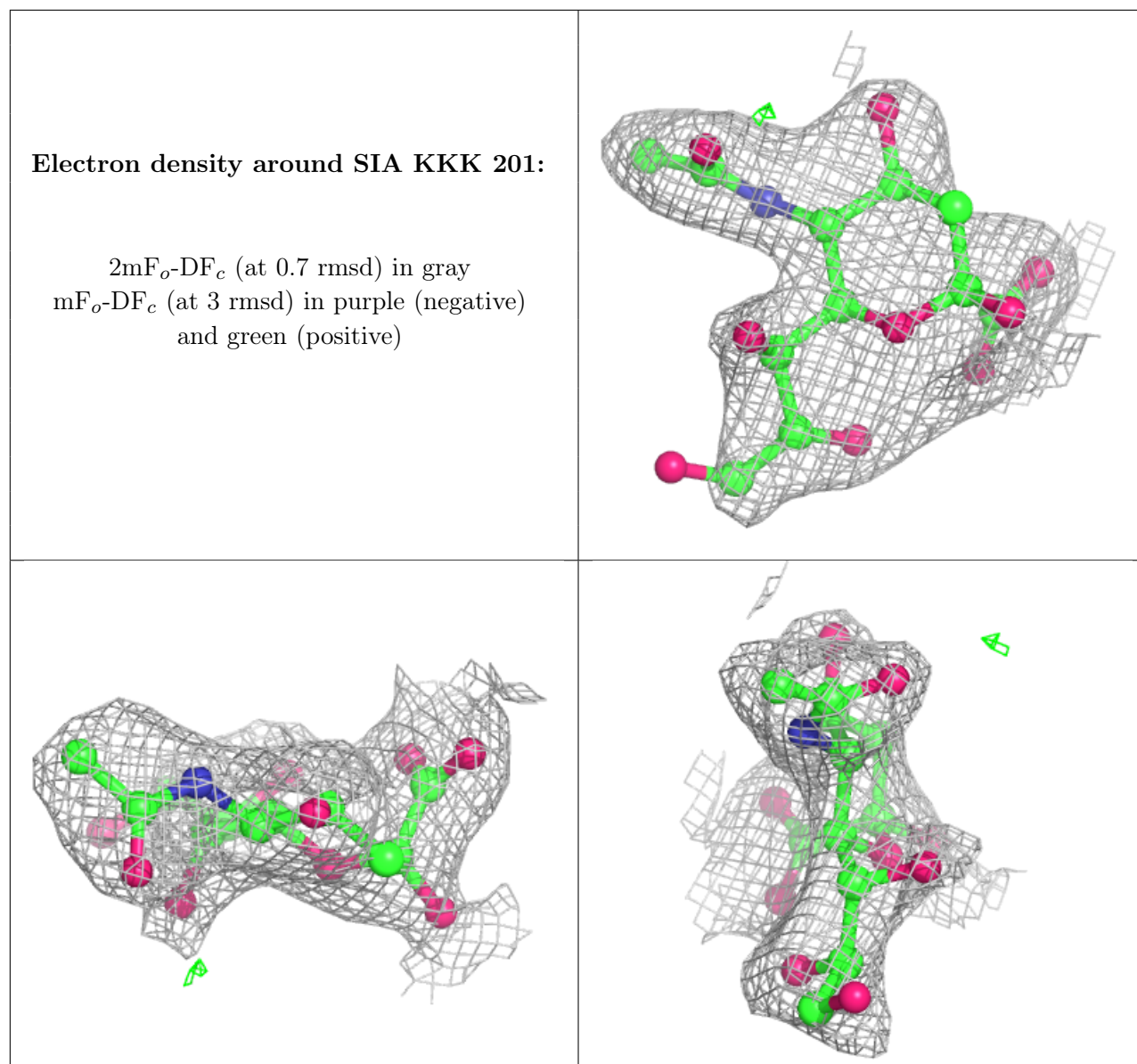
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around SIA LLL 201:**

$2mF_o-DF_c$  (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.