



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2020 – 07:43 PM BST

PDB ID : 6Y60
Title : Structure of Human Polyomavirus 12 VP1 in complex with 3'-Sialyllactosamine
Authors : Stroh, L.J.; Rustmeier, N.H.; Stehle, T.
Deposited on : 2020-02-26
Resolution : 1.80 Å(reported)

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

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<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 : 2.13.1
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) : 2.13.1

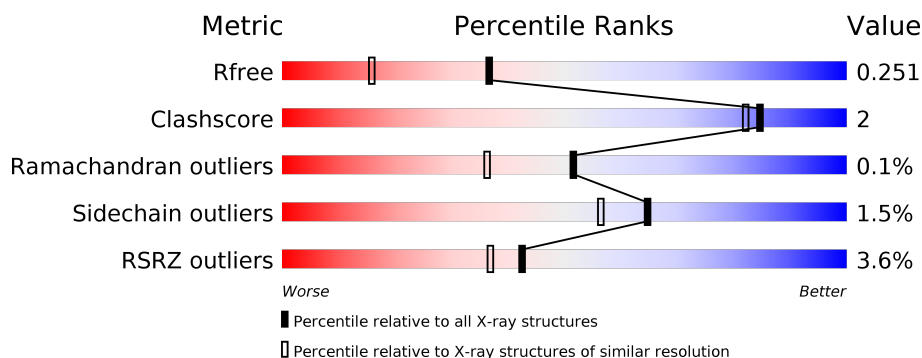
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria 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	AAA	277	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>6%</div> <div>6%</div> </div> </div>
1	BBB	277	<div> <div>2%</div> <div> <div></div> <div>86%</div> <div>7%</div> <div>6%</div> </div> </div>
1	CCC	277	<div> <div>6%</div> <div> <div></div> <div>85%</div> <div>8%</div> <div>7%</div> </div> </div>
1	DDD	277	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>5%</div> <div>6%</div> </div> </div>
1	EEE	277	<div> <div>%</div> <div> <div></div> <div>88%</div> <div>5%</div> <div>6%</div> </div> </div>
1	FFF	277	<div> <div>6%</div> <div> <div></div> <div>89%</div> <div>•</div> <div>6%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	GGG	277	
1	HHH	277	
1	III	277	
1	JJJ	277	
2	A	2	
2	B	2	
2	C	2	
2	D	2	
2	E	2	
2	F	2	
2	G	2	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 21950 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 Capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	260	Total	C	N	O	S	0	0	0
			1994	1264	336	380	14			
1	BBB	260	Total	C	N	O	S	0	0	0
			1986	1256	336	380	14			
1	CCC	258	Total	C	N	O	S	0	0	0
			1957	1239	331	374	13			
1	DDD	259	Total	C	N	O	S	0	0	0
			1974	1250	335	376	13			
1	EEE	260	Total	C	N	O	S	0	0	0
			2008	1269	340	385	14			
1	FFF	259	Total	C	N	O	S	0	1	0
			1991	1260	334	383	14			
1	GGG	258	Total	C	N	O	S	0	0	0
			1979	1253	335	377	14			
1	HHH	258	Total	C	N	O	S	0	0	0
			1949	1232	331	373	13			
1	III	258	Total	C	N	O	S	0	0	0
			1972	1248	335	376	13			
1	JJJ	260	Total	C	N	O	S	0	0	0
			1999	1266	336	383	14			

There are 40 discrepancies between the modelled and reference sequences:

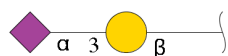
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	23	GLY	-	expression tag	UNP M4T5D3
AAA	24	SER	-	expression tag	UNP M4T5D3
AAA	25	HIS	-	expression tag	UNP M4T5D3
AAA	26	MET	-	expression tag	UNP M4T5D3
BBB	23	GLY	-	expression tag	UNP M4T5D3
BBB	24	SER	-	expression tag	UNP M4T5D3
BBB	25	HIS	-	expression tag	UNP M4T5D3
BBB	26	MET	-	expression tag	UNP M4T5D3
CCC	23	GLY	-	expression tag	UNP M4T5D3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	24	SER	-	expression tag	UNP M4T5D3
CCC	25	HIS	-	expression tag	UNP M4T5D3
CCC	26	MET	-	expression tag	UNP M4T5D3
DDD	23	GLY	-	expression tag	UNP M4T5D3
DDD	24	SER	-	expression tag	UNP M4T5D3
DDD	25	HIS	-	expression tag	UNP M4T5D3
DDD	26	MET	-	expression tag	UNP M4T5D3
EEE	23	GLY	-	expression tag	UNP M4T5D3
EEE	24	SER	-	expression tag	UNP M4T5D3
EEE	25	HIS	-	expression tag	UNP M4T5D3
EEE	26	MET	-	expression tag	UNP M4T5D3
FFF	23	GLY	-	expression tag	UNP M4T5D3
FFF	24	SER	-	expression tag	UNP M4T5D3
FFF	25	HIS	-	expression tag	UNP M4T5D3
FFF	26	MET	-	expression tag	UNP M4T5D3
GGG	23	GLY	-	expression tag	UNP M4T5D3
GGG	24	SER	-	expression tag	UNP M4T5D3
GGG	25	HIS	-	expression tag	UNP M4T5D3
GGG	26	MET	-	expression tag	UNP M4T5D3
HHH	23	GLY	-	expression tag	UNP M4T5D3
HHH	24	SER	-	expression tag	UNP M4T5D3
HHH	25	HIS	-	expression tag	UNP M4T5D3
HHH	26	MET	-	expression tag	UNP M4T5D3
III	23	GLY	-	expression tag	UNP M4T5D3
III	24	SER	-	expression tag	UNP M4T5D3
III	25	HIS	-	expression tag	UNP M4T5D3
III	26	MET	-	expression tag	UNP M4T5D3
JJJ	23	GLY	-	expression tag	UNP M4T5D3
JJJ	24	SER	-	expression tag	UNP M4T5D3
JJJ	25	HIS	-	expression tag	UNP M4T5D3
JJJ	26	MET	-	expression tag	UNP M4T5D3

- Molecule 2 is an oligosaccharide called N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose.



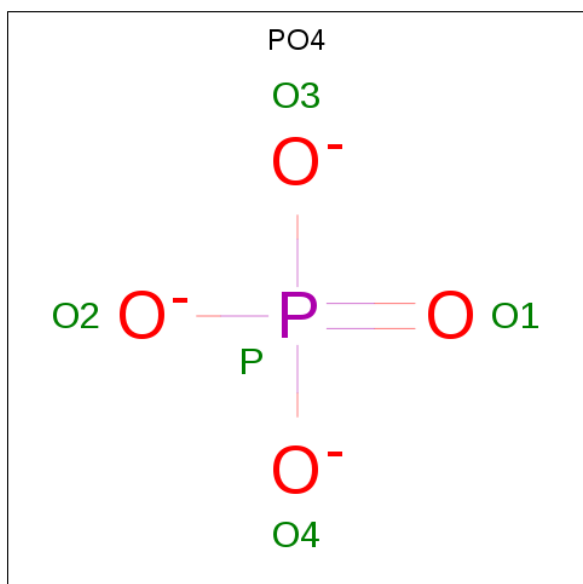
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	A	2	Total	C	N	O	0	0	0
			32	17	1	14			

Continued on next page...

Continued from previous page...

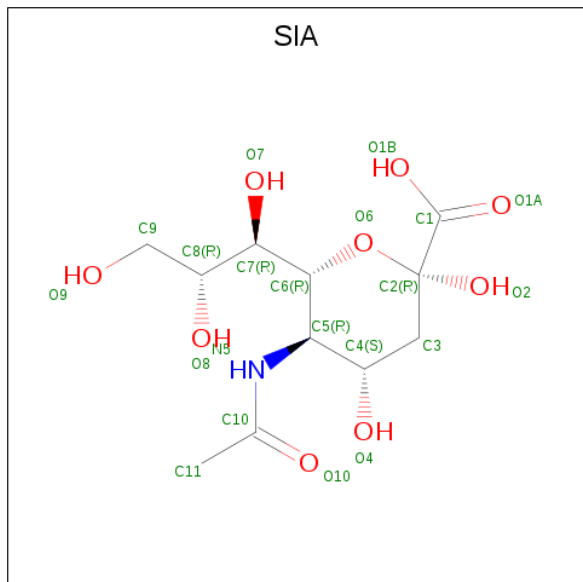
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	2	Total	C	N	O	0	0	0
			32	17	1	14			
2	C	2	Total	C	N	O	0	0	0
			32	17	1	14			
2	D	2	Total	C	N	O	0	0	0
			32	17	1	14			
2	E	2	Total	C	N	O	0	0	0
			32	17	1	14			
2	F	2	Total	C	N	O	0	0	0
			32	17	1	14			
2	G	2	Total	C	N	O	0	0	0
			32	17	1	14			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	AAA	1	Total	O	P	0	0
			5	4	1		
3	BBB	1	Total	O	P	0	0
			5	4	1		
3	CCC	1	Total	O	P	0	0
			5	4	1		
3	DDD	1	Total	O	P	0	0
			5	4	1		
3	EEE	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is N-acetyl-alpha-neuraminic acid (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$) (labeled as "Ligand of Interest" by author).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	BBB	1	Total	C	N	O	0	0
			21	11	1	9		
4	CCC	1	Total	C	N	O	0	0
			21	11	1	9		
4	GGG	1	Total	C	N	O	0	0
			21	11	1	9		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	AAA	206	Total	O	0	0
			206	206		
5	BBB	164	Total	O	0	1
			165	165		
5	CCC	145	Total	O	0	0
			145	145		
5	DDD	159	Total	O	0	0
			159	159		
5	EEE	206	Total	O	0	0
			206	206		
5	FFF	198	Total	O	0	0
			198	198		
5	GGG	164	Total	O	0	0
			164	164		

Continued on next page...

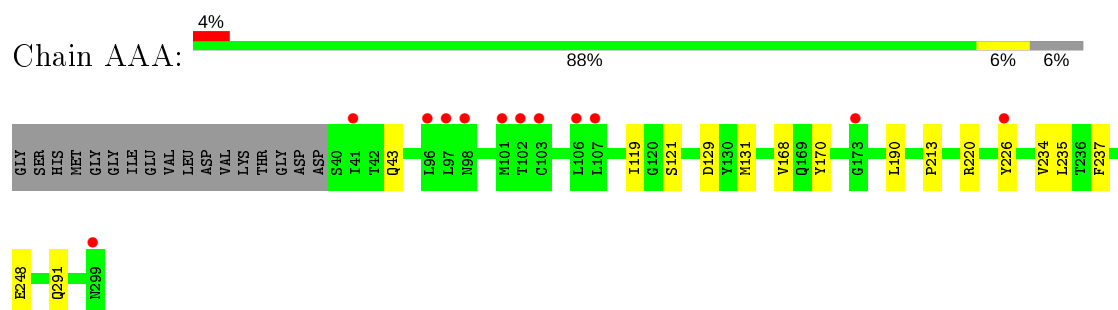
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	HHH	183	Total 183	O 183	0	0
5	III	189	Total 189	O 189	0	0
5	JJJ	211	Total 214	O 214	0	3

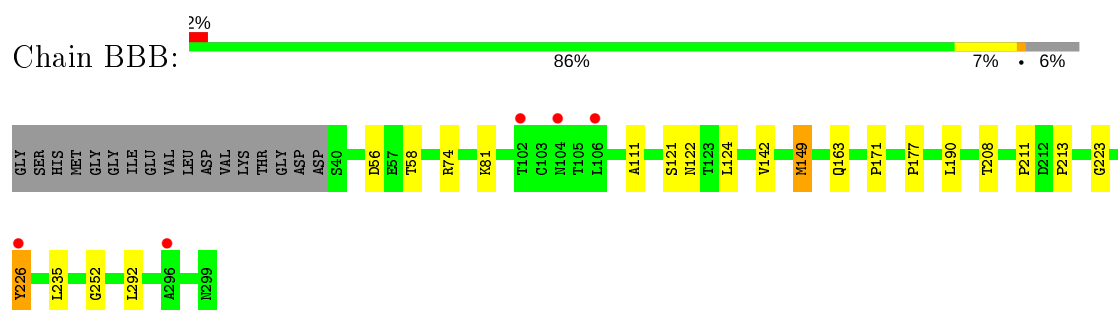
3 Residue-property plots [i](#)

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.

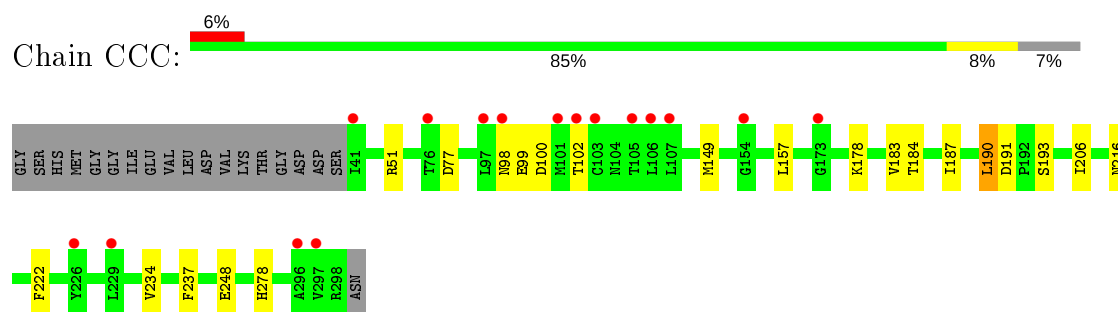
- Molecule 1: Capsid protein VP1



- Molecule 1: Capsid protein VP1

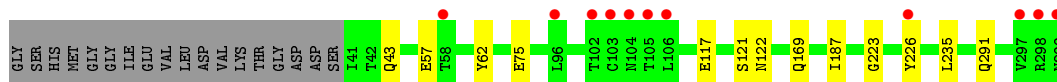


- Molecule 1: Capsid protein VP1

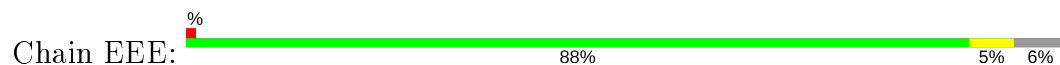


- Molecule 1: Capsid protein VP1

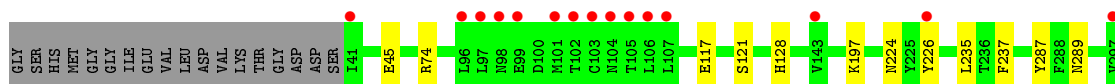




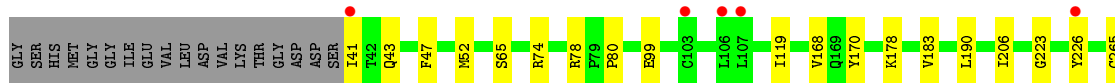
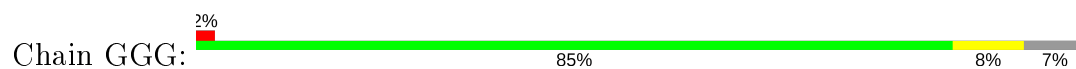
- Molecule 1: Capsid protein VP1



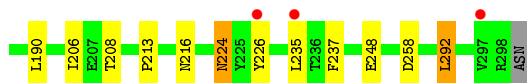
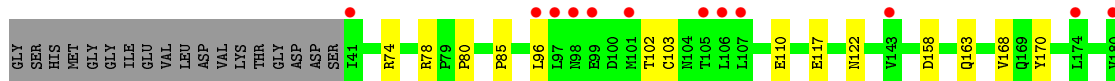
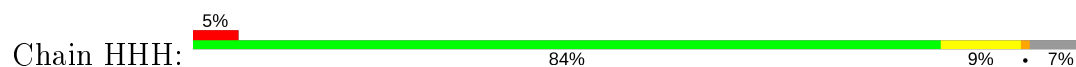
- Molecule 1: Capsid protein VP1



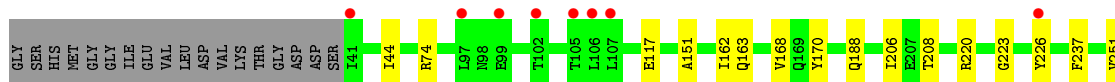
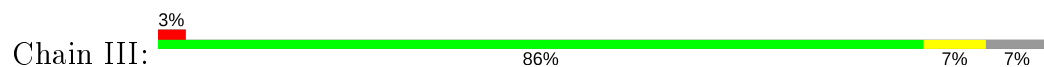
- Molecule 1: Capsid protein VP1



- Molecule 1: Capsid protein VP1

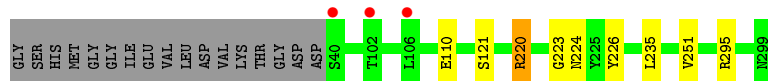


- Molecule 1: Capsid protein VP1





- Molecule 1: Capsid protein VP1



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose



- Molecule 2: N-acetyl-alpha-neuraminic acid-(2-3)-beta-D-galactopyranose

Chain G: 


GAL1
ST12

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	83.32Å 141.72Å 251.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.37 – 1.80 47.37 – 1.80	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.37-1.80) 99.4 (47.37-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.97 (at 1.79Å)	Xtriage
Refinement program	REFMAC 5.8.0253	Depositor
R, R_{free}	0.203 , 0.243 0.210 , 0.251	Depositor DCC
R_{free} test set	2741 reflections (1.00%)	wwPDB-VP
Wilson B-factor (Å ²)	25.2	Xtriage
Anisotropy	0.616	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 50.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	21950	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.65% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: PO4, GAL, SIA

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	AAA	0.70	0/2044	0.87	2/2791 (0.1%)
1	BBB	0.71	0/2036	0.88	0/2784
1	CCC	0.72	0/2007	0.88	2/2747 (0.1%)
1	DDD	0.68	0/2024	0.89	0/2768
1	EEE	0.70	0/2058	0.88	0/2811
1	FFF	0.69	0/2041	0.88	0/2789
1	GGG	0.73	0/2029	0.89	0/2772
1	HHH	0.71	0/1999	0.88	0/2738
1	III	0.69	0/2022	0.89	0/2765
1	JJJ	0.70	0/2049	0.88	2/2799 (0.1%)
All	All	0.70	0/20309	0.88	6/27764 (0.0%)

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	220	ARG	NE-CZ-NH1	6.25	123.43	120.30
1	JJJ	220	ARG	NE-CZ-NH2	-5.36	117.62	120.30
1	CCC	51	ARG	NE-CZ-NH1	5.31	122.95	120.30
1	JJJ	220	ARG	NE-CZ-NH1	5.21	122.90	120.30
1	CCC	77	ASP	CB-CG-OD2	-5.09	113.72	118.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	AAA	1994	0	1894	10	0
1	BBB	1986	0	1885	12	0
1	CCC	1957	0	1844	14	0
1	DDD	1974	0	1874	10	0
1	EEE	2008	0	1927	9	0
1	FFF	1991	0	1879	6	0
1	GGG	1979	0	1888	13	0
1	HHH	1949	0	1830	14	0
1	III	1972	0	1872	11	0
1	JJJ	1999	0	1904	6	0
2	A	32	0	28	0	0
2	B	32	0	28	1	0
2	C	32	0	28	0	0
2	D	32	0	28	0	0
2	E	32	0	28	0	0
2	F	32	0	28	0	0
2	G	32	0	28	1	0
3	AAA	5	0	0	0	0
3	BBB	5	0	0	0	0
3	CCC	5	0	0	0	0
3	DDD	5	0	0	0	0
3	EEE	5	0	0	0	0
4	BBB	21	0	18	0	0
4	CCC	21	0	18	0	0
4	GGG	21	0	18	0	0
5	AAA	206	0	0	0	0
5	BBB	165	0	0	0	0
5	CCC	145	0	0	2	0
5	DDD	159	0	0	1	0
5	EEE	206	0	0	0	0
5	FFF	198	0	0	2	0
5	GGG	164	0	0	2	0
5	HHH	183	0	0	2	0
5	III	189	0	0	0	0
5	JJJ	214	0	0	1	0
All	All	21950	0	19047	91	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 91 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:DDD:43:GLN:HE22	1:DDD:291:GLN:HG3	1.45	0.81
1:BBB:121:SER:HB3	1:BBB:235:LEU:HD23	1.66	0.75
1:HHH:224:ASN:ND2	5:HHH:401:HOH:O	2.26	0.68
1:AAA:43:GLN:HE22	1:AAA:291:GLN:HG3	1.58	0.68
1:CCC:278:HIS:ND1	5:CCC:401:HOH:O	2.29	0.65

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	AAA	258/277 (93%)	250 (97%)	8 (3%)	0	100	100
1	BBB	258/277 (93%)	253 (98%)	5 (2%)	0	100	100
1	CCC	256/277 (92%)	241 (94%)	14 (6%)	1 (0%)	34	21
1	DDD	257/277 (93%)	245 (95%)	12 (5%)	0	100	100
1	EEE	258/277 (93%)	247 (96%)	11 (4%)	0	100	100
1	FFF	258/277 (93%)	251 (97%)	7 (3%)	0	100	100
1	GGG	256/277 (92%)	248 (97%)	8 (3%)	0	100	100
1	HHH	256/277 (92%)	244 (95%)	11 (4%)	1 (0%)	34	21
1	III	256/277 (92%)	249 (97%)	7 (3%)	0	100	100
1	JJJ	258/277 (93%)	250 (97%)	8 (3%)	0	100	100
All	All	2571/2770 (93%)	2478 (96%)	91 (4%)	2 (0%)	51	36

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	CCC	99	GLU
1	HHH	103	CYS

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	AAA	211/240 (88%)	209 (99%)	2 (1%)	78	75
1	BBB	213/240 (89%)	206 (97%)	7 (3%)	38	23
1	CCC	207/240 (86%)	206 (100%)	1 (0%)	88	87
1	DDD	210/240 (88%)	209 (100%)	1 (0%)	88	87
1	EEE	219/240 (91%)	216 (99%)	3 (1%)	67	59
1	FFF	211/240 (88%)	207 (98%)	4 (2%)	57	46
1	GGG	212/240 (88%)	209 (99%)	3 (1%)	67	59
1	HHH	206/240 (86%)	199 (97%)	7 (3%)	37	22
1	III	210/240 (88%)	207 (99%)	3 (1%)	67	59
1	JJJ	214/240 (89%)	213 (100%)	1 (0%)	88	87
All	All	2113/2400 (88%)	2081 (98%)	32 (2%)	65	56

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	FFF	74	ARG
1	FFF	226	TYR
1	III	117	GLU
1	FFF	224	ASN
1	GGG	74	ARG

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

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 ⓘ

14 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	GAL	A	1	2	12,12,12	0.73	0	17,17,17	1.04	0
2	SIA	A	2	2	17,20,21	0.74	0	21,28,31	1.17	3 (14%)
2	GAL	B	1	2	12,12,12	0.81	0	17,17,17	1.50	6 (35%)
2	SIA	B	2	2	17,20,21	0.83	0	21,28,31	1.21	2 (9%)
2	GAL	C	1	2	12,12,12	0.84	0	17,17,17	1.02	0
2	SIA	C	2	2	17,20,21	0.80	0	21,28,31	1.22	3 (14%)
2	GAL	D	1	2	12,12,12	0.78	0	17,17,17	1.26	3 (17%)
2	SIA	D	2	2	17,20,21	1.22	1 (5%)	21,28,31	0.89	1 (4%)
2	GAL	E	1	2	12,12,12	0.81	0	17,17,17	1.20	1 (5%)
2	SIA	E	2	2	17,20,21	1.06	0	21,28,31	0.75	0
2	GAL	F	1	2	12,12,12	0.85	0	17,17,17	1.32	1 (5%)
2	SIA	F	2	2	17,20,21	1.27	2 (11%)	21,28,31	1.03	1 (4%)
2	GAL	G	1	2	12,12,12	0.92	0	17,17,17	1.72	4 (23%)
2	SIA	G	2	2	17,20,21	0.90	0	21,28,31	1.23	1 (4%)

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	GAL	A	1	2	-	2/2/22/22	0/1/1/1
2	SIA	A	2	2	-	0/14/34/38	0/1/1/1
2	GAL	B	1	2	-	0/2/22/22	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	SIA	B	2	2	-	0/14/34/38	0/1/1/1
2	GAL	C	1	2	-	2/2/22/22	0/1/1/1
2	SIA	C	2	2	-	0/14/34/38	0/1/1/1
2	GAL	D	1	2	-	2/2/22/22	0/1/1/1
2	SIA	D	2	2	-	2/14/34/38	0/1/1/1
2	GAL	E	1	2	-	1/2/22/22	0/1/1/1
2	SIA	E	2	2	-	1/14/34/38	0/1/1/1
2	GAL	F	1	2	-	0/2/22/22	0/1/1/1
2	SIA	F	2	2	-	0/14/34/38	0/1/1/1
2	GAL	G	1	2	-	0/2/22/22	0/1/1/1
2	SIA	G	2	2	-	0/14/34/38	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	2	SIA	C4-C5	2.95	1.55	1.53
2	F	2	SIA	C7-C6	2.84	1.56	1.53
2	F	2	SIA	C4-C5	2.61	1.55	1.53

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	1	GAL	O5-C5-C6	3.59	115.36	106.44
2	G	1	GAL	O5-C5-C6	3.43	114.96	106.44
2	G	1	GAL	O2-C2-C3	-3.34	102.63	110.35
2	G	2	SIA	C6-O6-C2	2.94	117.62	111.34
2	G	1	GAL	O4-C4-C3	-2.87	103.71	110.35

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	1	GAL	O5-C5-C6-O6
2	A	1	GAL	O5-C5-C6-O6
2	C	1	GAL	C4-C5-C6-O6
2	D	1	GAL	C4-C5-C6-O6
2	D	1	GAL	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	2	SIA	1	0
2	G	1	GAL	1	0

5.6 Ligand geometry

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
3	PO4	CCC	302	-	4,4,4	0.80	0	6,6,6	0.83	0
3	PO4	BBB	302	-	4,4,4	0.49	0	6,6,6	0.56	0
4	SIA	GGG	301	-	18,21,21	1.65	2 (11%)	21,31,31	0.95	1 (4%)
4	SIA	BBB	301	-	18,21,21	1.50	4 (22%)	21,31,31	1.26	1 (4%)
3	PO4	AAA	303	-	4,4,4	1.09	1 (25%)	6,6,6	0.59	0
4	SIA	CCC	301	-	18,21,21	1.31	1 (5%)	21,31,31	1.03	0
3	PO4	DDD	303	-	4,4,4	0.68	0	6,6,6	0.86	0
3	PO4	EEE	303	-	4,4,4	0.69	0	6,6,6	0.64	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
4	SIA	GGG	301	-	-	0/14/38/38	0/1/1/1
4	SIA	BBB	301	-	-	0/14/38/38	0/1/1/1
4	SIA	CCC	301	-	-	1/14/38/38	0/1/1/1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	GGG	301	SIA	O2-C2	5.54	1.47	1.39
4	CCC	301	SIA	O2-C2	4.69	1.46	1.39
4	BBB	301	SIA	O2-C2	4.59	1.46	1.39

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	GGG	301	SIA	C3-C2	2.49	1.54	1.51
4	BBB	301	SIA	C4-C5	2.24	1.55	1.53

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	BBB	301	SIA	O6-C6-C5	-3.87	106.00	109.78
4	GGG	301	SIA	C6-C5-N5	2.40	114.89	110.91

There are no chirality outliers.

All (1) torsion outliers are listed below:

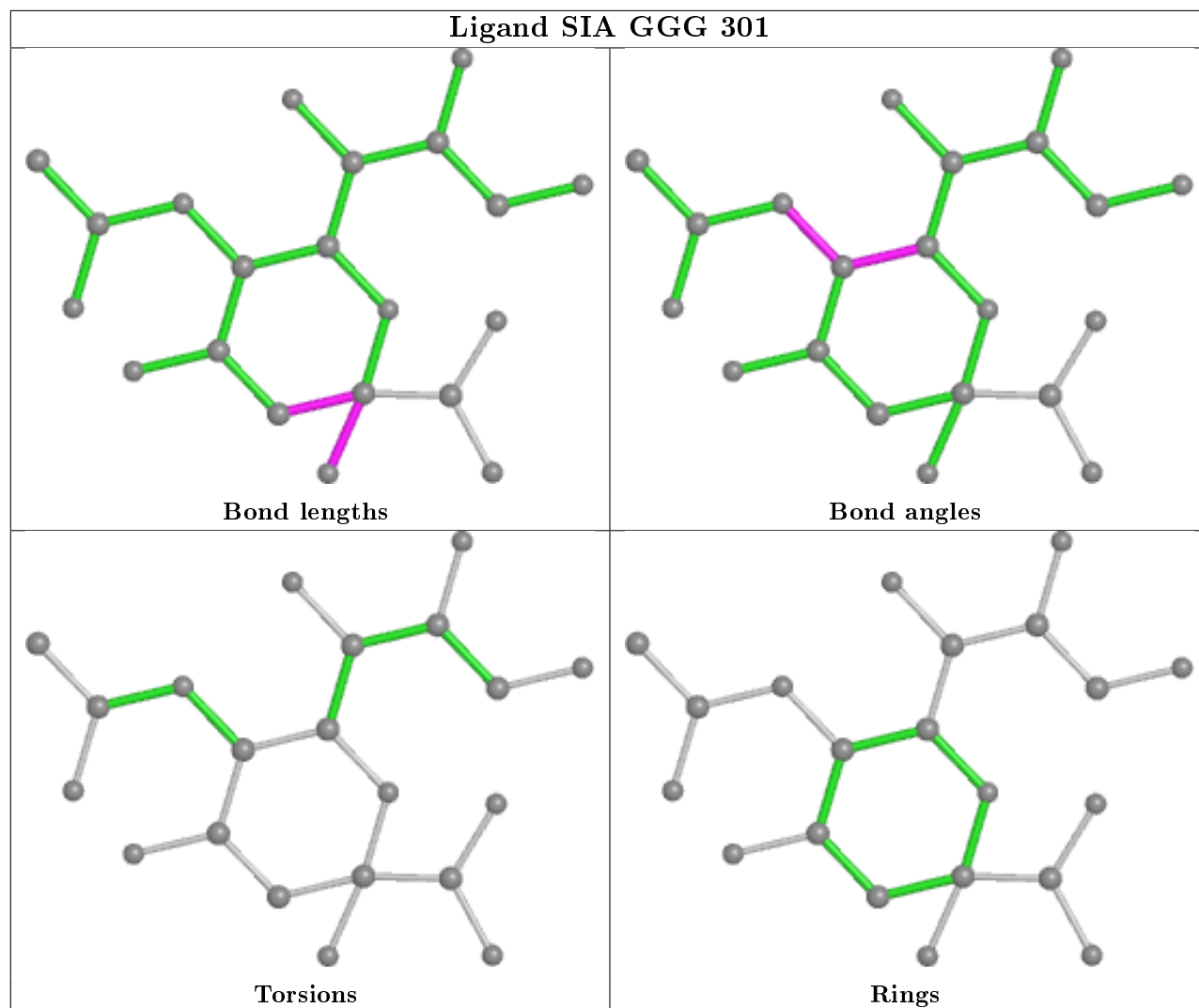
Mol	Chain	Res	Type	Atoms
4	CCC	301	SIA	O8-C8-C9-O9

There are no ring outliers.

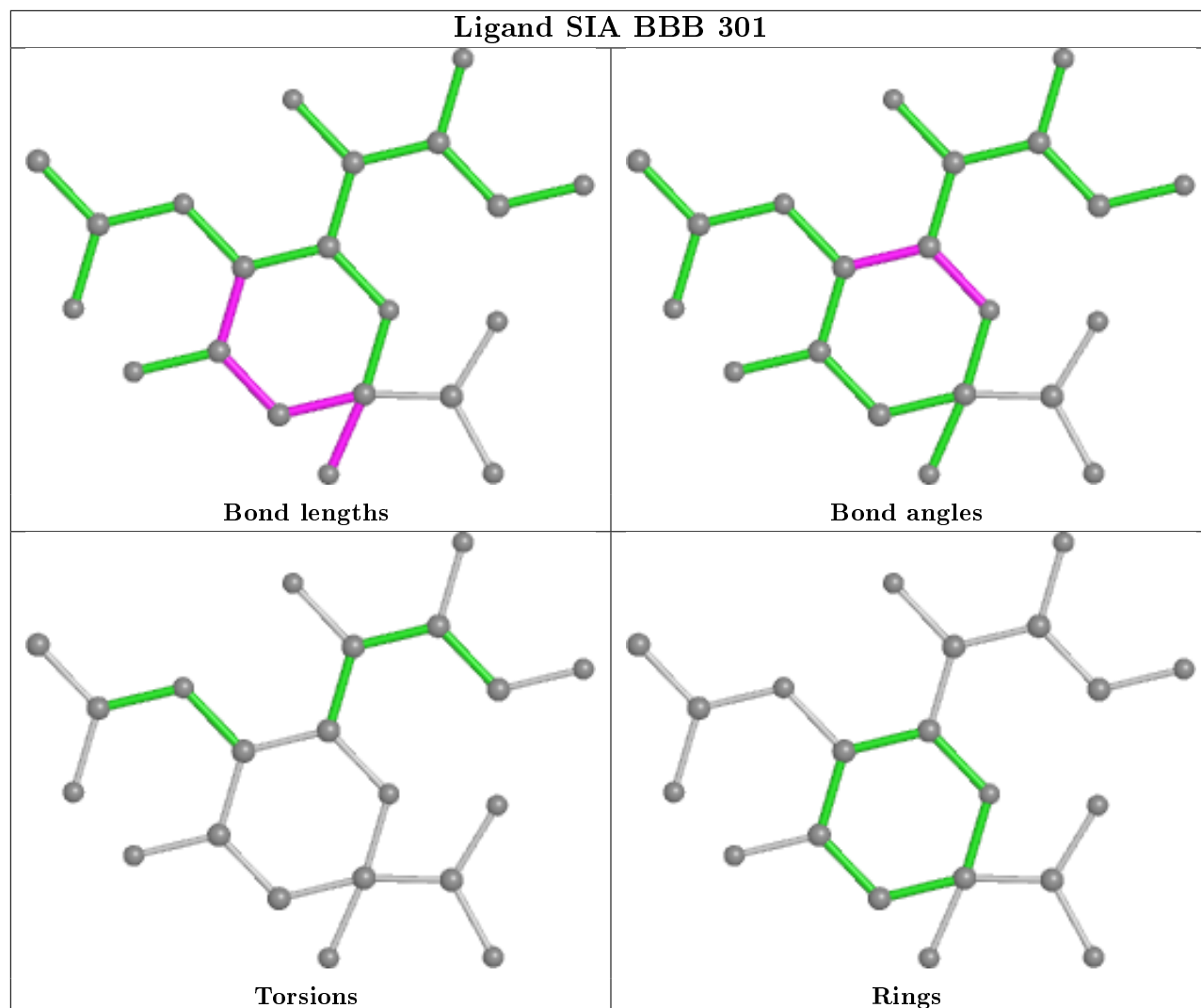
No monomer is involved in short contacts.

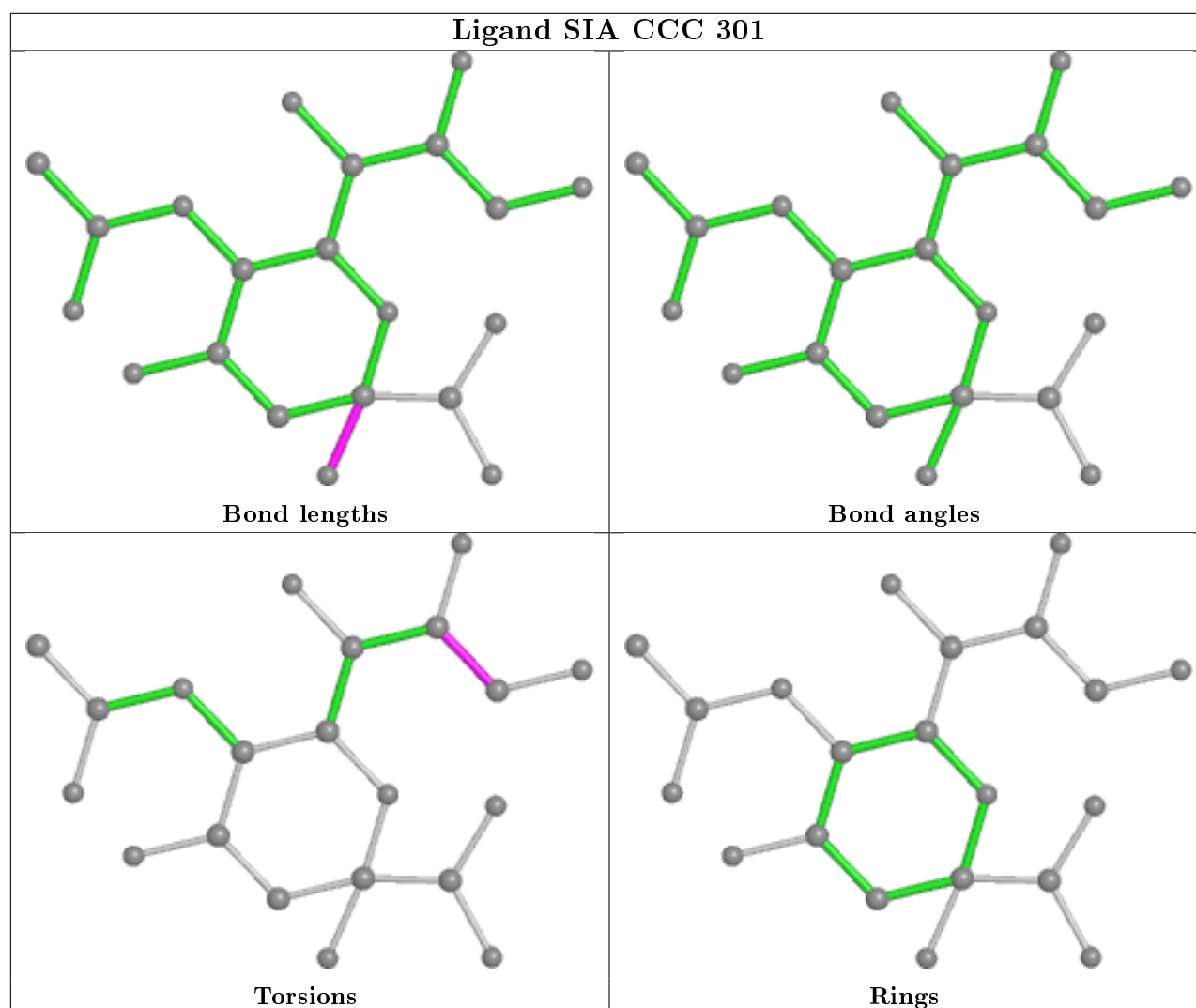
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. 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.

Ligand SIA GGG 301



Ligand SIA BBB 301





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, 95th 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 ²)	Q<0.9
1	AAA	260/277 (93%)	0.11	12 (4%) 32 26	21, 29, 47, 64	0
1	BBB	260/277 (93%)	0.23	5 (1%) 66 63	23, 34, 50, 62	0
1	CCC	258/277 (93%)	0.51	16 (6%) 20 16	24, 38, 58, 75	0
1	DDD	259/277 (93%)	0.16	11 (4%) 36 30	22, 34, 53, 73	0
1	EEE	260/277 (93%)	-0.11	3 (1%) 79 76	20, 27, 47, 60	0
1	FFF	259/277 (93%)	0.13	16 (6%) 20 16	21, 29, 46, 74	0
1	GGG	258/277 (93%)	0.30	5 (1%) 66 63	21, 35, 51, 66	0
1	HHH	258/277 (93%)	0.48	15 (5%) 23 18	23, 35, 54, 71	0
1	III	258/277 (93%)	0.14	8 (3%) 49 43	20, 31, 50, 69	0
1	JJJ	260/277 (93%)	-0.09	3 (1%) 79 76	18, 26, 39, 62	0
All	All	2590/2770 (93%)	0.19	94 (3%) 42 37	18, 32, 51, 75	0

The worst 5 of 94 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	102	THR	5.9
1	FFF	299	ASN	5.4
1	BBB	102	THR	5.1
1	HHH	96	LEU	5.0
1	FFF	106	LEU	4.9

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, 95th 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 ²)	Q<0.9
2	GAL	F	1	12/12	0.79	0.19	43,55,60,65	0
2	GAL	E	1	12/12	0.82	0.24	46,59,62,68	0
2	GAL	A	1	12/12	0.85	0.19	45,60,63,64	0
2	GAL	D	1	12/12	0.86	0.19	42,54,60,63	0
2	GAL	B	1	12/12	0.87	0.15	44,58,62,72	0
2	GAL	C	1	12/12	0.88	0.13	39,54,65,71	0
2	SIA	E	2	20/21	0.89	0.13	29,36,49,51	0
2	GAL	G	1	12/12	0.91	0.13	36,50,54,62	0
2	SIA	F	2	20/21	0.92	0.09	27,31,41,42	0
2	SIA	D	2	20/21	0.92	0.11	26,31,41,42	0
2	SIA	B	2	20/21	0.94	0.08	30,34,42,47	0
2	SIA	G	2	20/21	0.95	0.08	25,30,37,42	0
2	SIA	A	2	20/21	0.95	0.07	24,31,38,41	0
2	SIA	C	2	20/21	0.96	0.07	24,31,37,39	0

6.4 Ligands

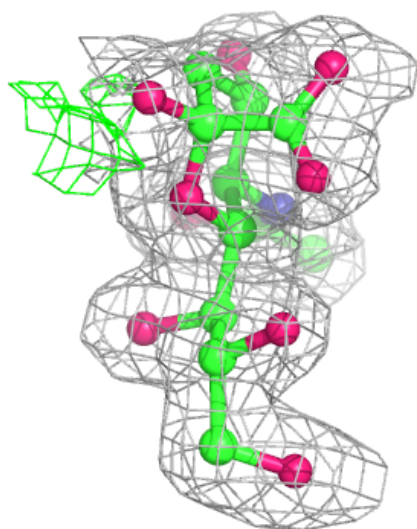
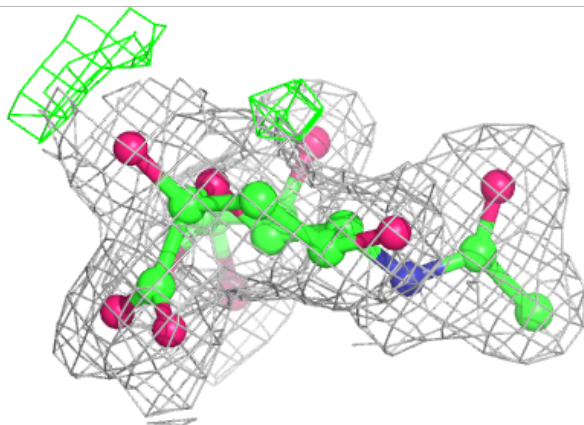
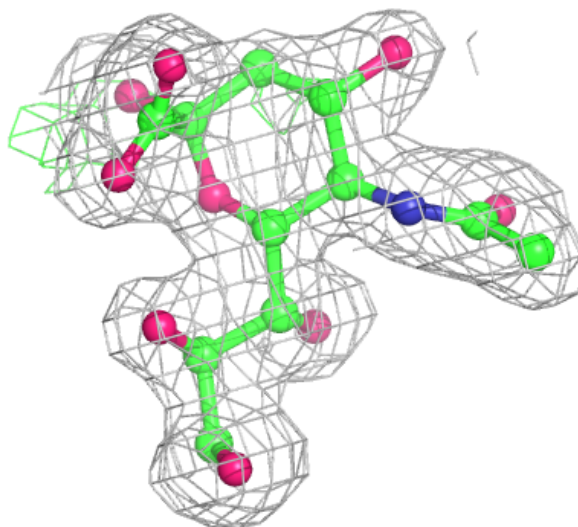
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
4	SIA	GGG	301	21/21	0.90	0.12	28,33,42,46	0
4	SIA	CCC	301	21/21	0.90	0.10	34,42,47,50	0
4	SIA	BBB	301	21/21	0.91	0.11	29,35,39,40	0
3	PO4	EEE	303	5/5	0.98	0.07	32,32,38,39	0
3	PO4	CCC	302	5/5	0.98	0.06	33,35,39,42	0
3	PO4	AAA	303	5/5	0.98	0.07	34,35,38,41	0
3	PO4	BBB	302	5/5	0.99	0.07	39,40,42,42	0
3	PO4	DDD	303	5/5	0.99	0.06	28,32,36,40	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.

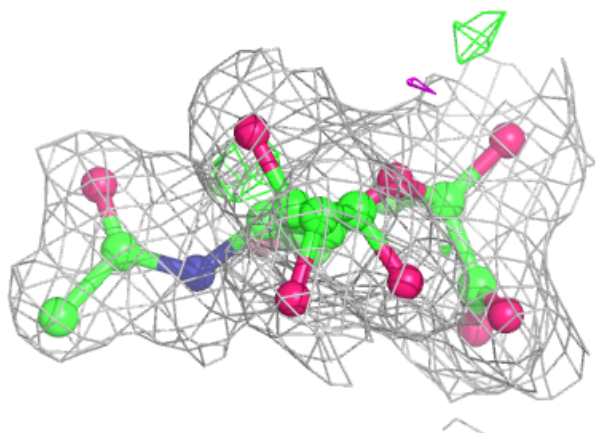
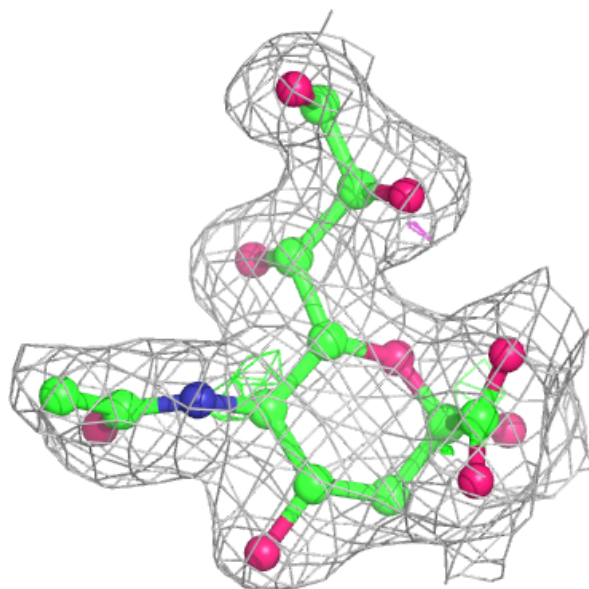
Electron density around SIA GGG 301:

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



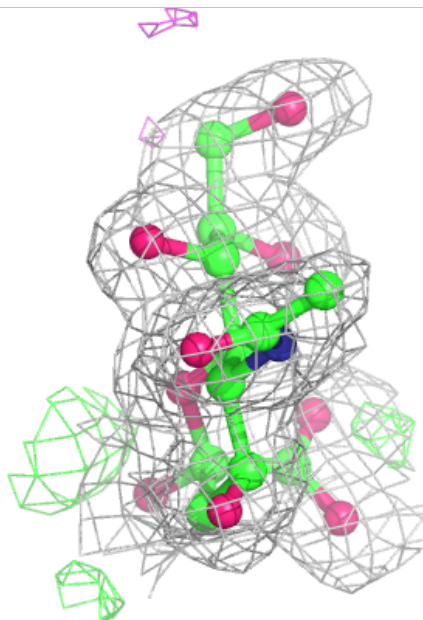
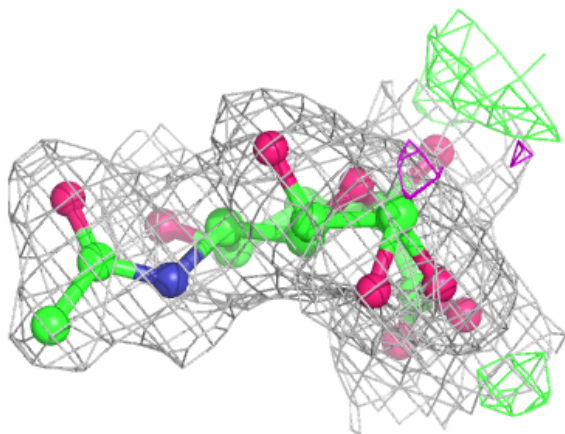
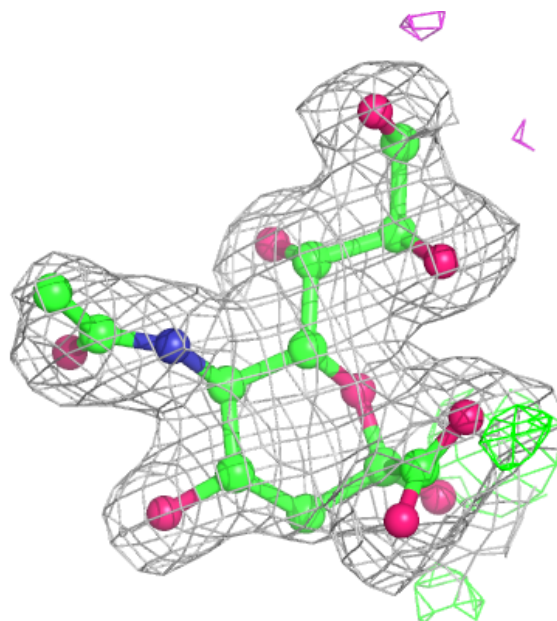
Electron density around SIA CCC 301:

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



Electron density around SIA BBB 301:

$2mF_o-DF_c$ (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.