



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 23, 2024 – 06:04 AM EDT

PDB ID : 1SID  
Title : MURINE POLYOMAVIRUS COMPLEXED WITH 3'SIALYL LACTOSE  
Authors : Stehle, T.; Harrison, S.C.  
Deposited on : 1995-12-12  
Resolution : 3.65 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)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

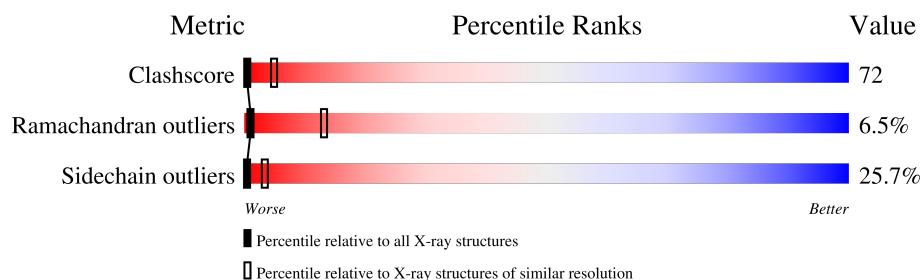
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.65 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	180529	1328 (3.80-3.52)
Ramachandran outliers	177936	1306 (3.80-3.52)
Sidechain outliers	177891	1303 (3.80-3.52)

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  $\leq 5\%$ .

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	383	23% 50% 20% . .
1	B	383	25% 46% 22% . .
1	C	383	22% 48% 21% . 7%
1	D	383	20% 44% 22% . 11%
1	E	383	20% 52% 21% . .
1	F	383	20% 46% 23% . 8%
2	G	3	67% 33%
2	H	3	33% 67%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain	
2	I	3	<div><div></div></div>	67%
2	J	3	<div><div></div></div>	67%
2	K	3	<div><div></div></div>	67%
2	L	3	<div><div></div></div>	67%

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 17003 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.

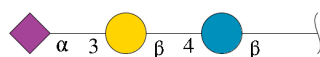
- Molecule 1 is a protein called POLYOMAVIRUS COAT PROTEIN VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	366	Total	C	N	O	S	0	0	0
			2849	1804	479	550	16			
1	B	367	Total	C	N	O	S	0	0	0
			2857	1808	481	552	16			
1	C	357	Total	C	N	O	S	0	0	0
			2784	1761	468	539	16			
1	D	340	Total	C	N	O	S	0	0	0
			2645	1674	445	511	15			
1	E	367	Total	C	N	O	S	0	0	0
			2857	1808	481	552	16			
1	F	354	Total	C	N	O	S	0	0	0
			2753	1740	461	536	16			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	6	ALA	SER	conflict	UNP P49302
B	6	ALA	SER	conflict	UNP P49302
C	6	ALA	SER	conflict	UNP P49302
D	6	ALA	SER	conflict	UNP P49302
E	6	ALA	SER	conflict	UNP P49302
F	6	ALA	SER	conflict	UNP P49302

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



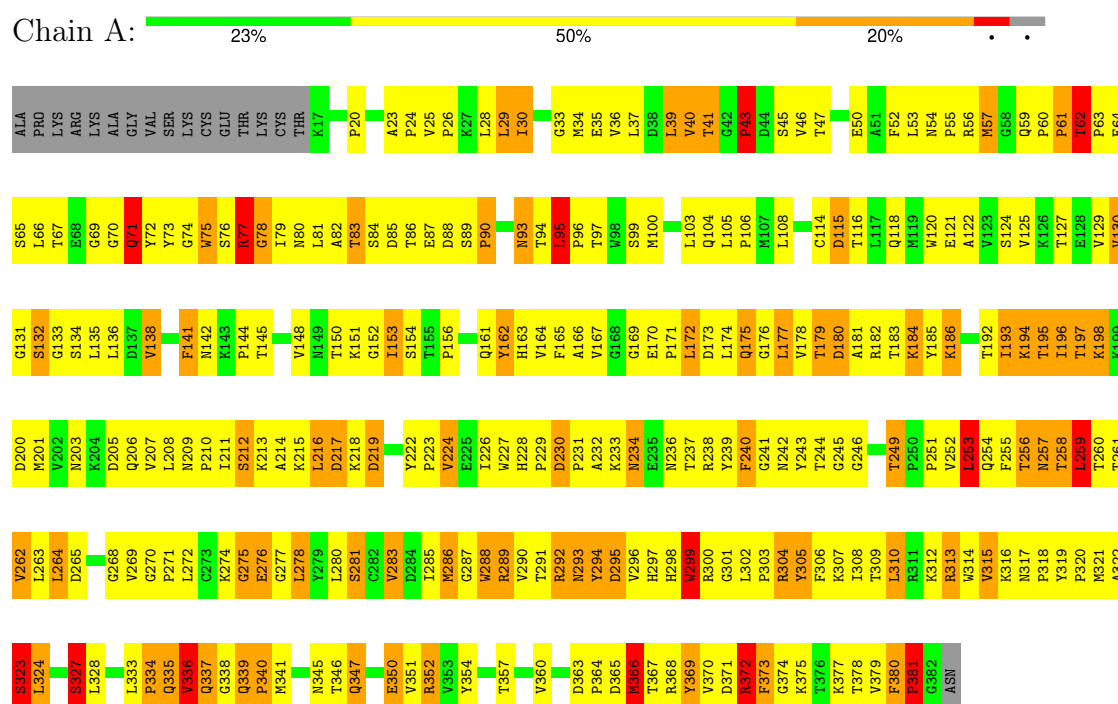
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	3	Total 43	C 23	N 1	O 19	0	0	0
2	H	3	Total 43	C 23	N 1	O 19	0	0	0
2	I	3	Total 43	C 23	N 1	O 19	0	0	0
2	J	3	Total 43	C 23	N 1	O 19	0	0	0
2	K	3	Total 43	C 23	N 1	O 19	0	0	0
2	L	3	Total 43	C 23	N 1	O 19	0	0	0

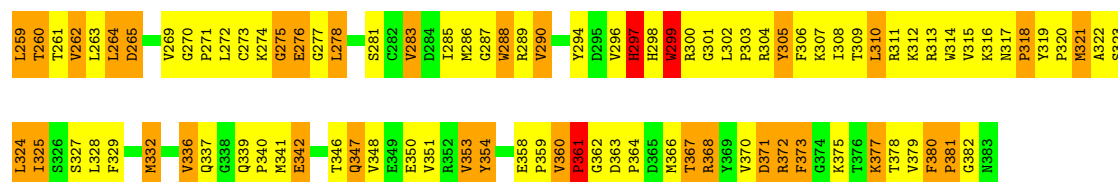
### 3 Residue-property plots

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. 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. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS was not executed.

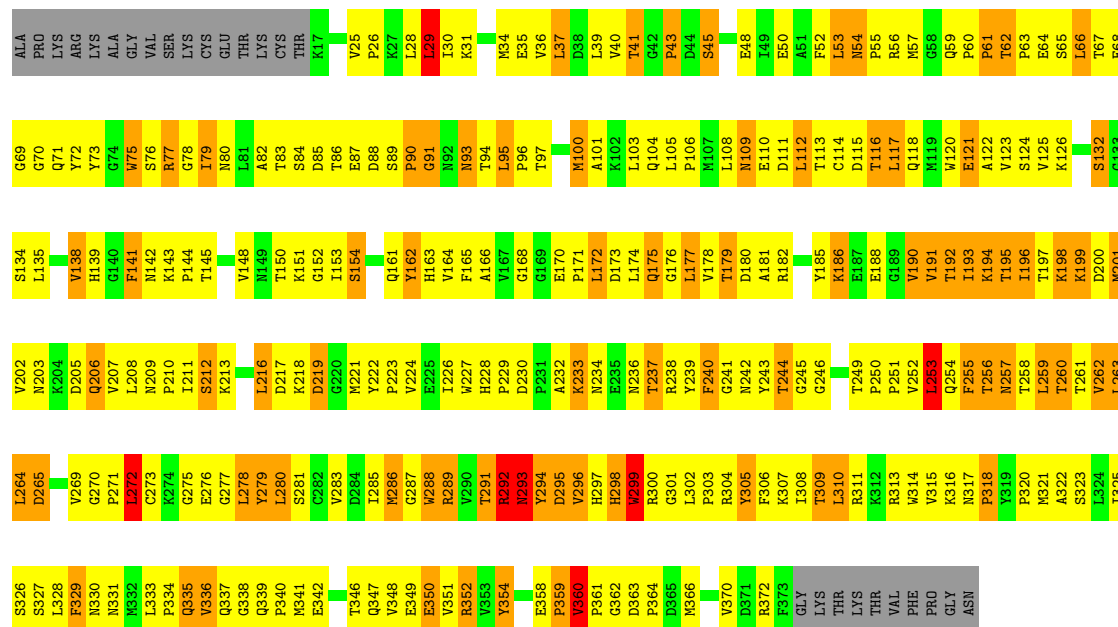
#### • Molecule 1: POLYOMAVIRUS COAT PROTEIN VP1





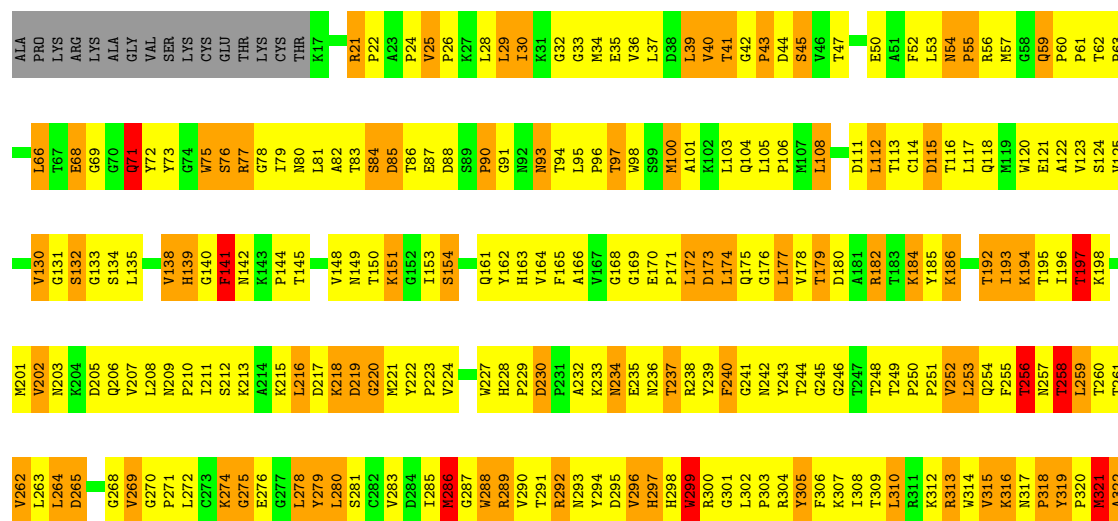
• Molecule 1: POLYOMAVIRUS COAT PROTEIN VP1

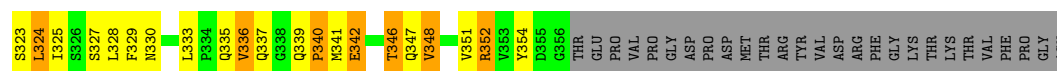
Chain C: 22% 48% 21% 7%



• Molecule 1: POLYOMAVIRUS COAT PROTEIN VP1

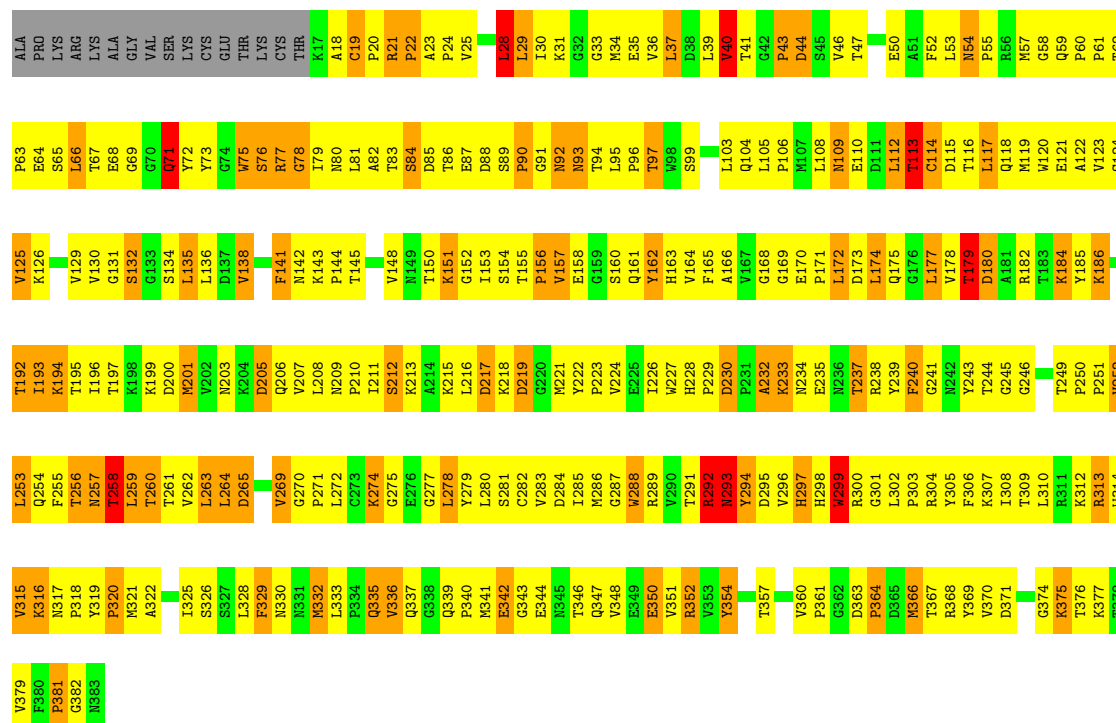
Chain D: 20% 44% 22% 11%





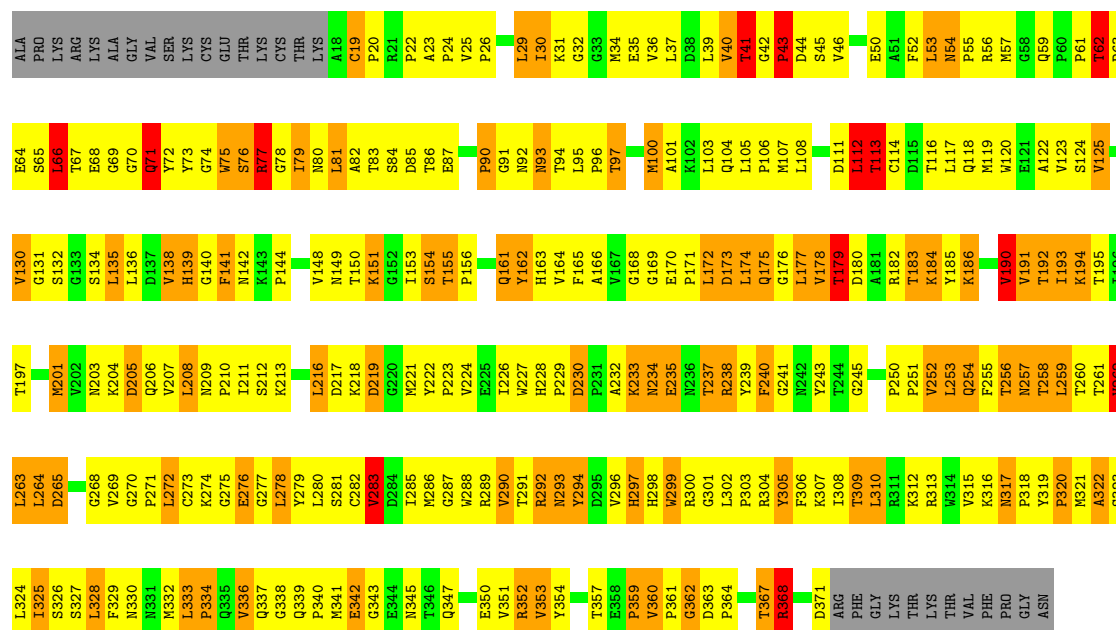
• Molecule 1: POLYOMAVIRUS COAT PROTEIN VP1

Chain E: 20% 52% 21%



• Molecule 1: POLYOMAVIRUS COAT PROTEIN VP1

Chain F: 20% 46% 23% 8%





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

Chain G:  67% 33%

BGC1  
GAL2  
STIA3

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

Chain H:  33% 67%

BGC1  
GAL2  
STIA3

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

Chain I:  33% 67%

BGC1  
GAL2  
STIA3

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

Chain J:  33% 67%

BGC1  
GAL2  
STIA3

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

Chain K:  33% 67%

BGC1  
GAL2  
STIA3

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

Chain L:  33% 67%

BGC1  
GAL2  
STIA3

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	570.00Å 570.00Å 570.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 3.65	Depositor
% Data completeness (in resolution range)	73.0 (12.00-3.65)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	X-PLOR	Depositor
R, $R_{free}$	0.236 , 0.253	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	17003	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	77.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GAL, BGC, 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	A	0.96	2/2920 (0.1%)	1.26	22/3981 (0.6%)
1	B	0.95	0/2927	1.25	19/3989 (0.5%)
1	C	0.97	0/2852	1.26	21/3888 (0.5%)
1	D	1.01	2/2708 (0.1%)	1.27	23/3690 (0.6%)
1	E	1.00	1/2928 (0.0%)	1.34	23/3992 (0.6%)
1	F	0.96	2/2820 (0.1%)	1.25	21/3847 (0.5%)
All	All	0.98	7/17155 (0.0%)	1.27	129/23387 (0.6%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	B	0	3
1	C	0	4
1	D	0	5
1	E	0	2
1	F	0	3
All	All	0	22

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	336	VAL	C-N	13.07	1.64	1.34
1	D	342	GLU	CD-OE2	6.37	1.32	1.25
1	A	336	VAL	C-N	-5.83	1.20	1.34
1	D	299	TRP	CB-CG	5.46	1.60	1.50
1	F	234	ASN	CB-CG	5.37	1.63	1.51

The worst 5 of 129 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	336	VAL	O-C-N	21.97	157.85	122.70
1	A	336	VAL	O-C-N	-17.79	94.24	122.70
1	E	336	VAL	CA-C-N	-16.45	81.00	117.20
1	C	29	LEU	CA-CB-CG	-11.50	88.85	115.30
1	F	264	LEU	CA-CB-CG	-11.09	89.80	115.30

There are no chirality outliers.

5 of 22 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	162	TYR	Sidechain
1	A	305	TYR	Sidechain
1	A	336	VAL	Peptide,Mainchain
1	A	71	GLN	Mainchain
1	B	71	GLN	Mainchain

## 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2849	0	2814	457	2
1	B	2857	0	2820	454	0
1	C	2784	0	2743	441	0
1	D	2645	0	2616	404	0
1	E	2857	0	2821	472	0
1	F	2753	0	2710	420	0
2	G	43	0	37	1	0
2	H	43	0	37	5	0
2	I	43	0	37	3	0
2	J	43	0	37	1	0
2	K	43	0	37	3	0
2	L	43	0	37	5	0
All	All	17003	0	16746	2434	2

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 72.

The worst 5 of 2434 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:A:328:LEU:CD2	1:A:333:LEU:HD21	1.37	1.52
1:A:328:LEU:HD21	1:A:333:LEU:CD2	1.40	1.50
1:A:175:GLN:HB2	1:A:230:ASP:HB2	1.26	1.11
1:F:175:GLN:HB2	1:F:230:ASP:HB2	1.33	1.10
1:B:336:VAL:HA	1:B:337:GLN:N	1.68	1.09

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:274:LYS:CD	1:A:328:LEU:O[2_555]	1.75	0.45
1:A:217:ASP:O	1:A:336:VAL:O[2_555]	2.05	0.15

## 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	364/383 (95%)	270 (74%)	67 (18%)	27 (7%)	1	10
1	B	363/383 (95%)	280 (77%)	62 (17%)	21 (6%)	1	14
1	C	353/383 (92%)	275 (78%)	60 (17%)	18 (5%)	1	16
1	D	336/383 (88%)	264 (79%)	58 (17%)	14 (4%)	2	19
1	E	365/383 (95%)	278 (76%)	62 (17%)	25 (7%)	1	12
1	F	350/383 (91%)	271 (77%)	45 (13%)	34 (10%)	0	6
All	All	2131/2298 (93%)	1638 (77%)	354 (17%)	139 (6%)	1	13

5 of 139 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	43	PRO
1	A	62	THR
1	A	148	VAL
1	A	183	THR
1	A	276	GLU

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	321/335 (96%)	244 (76%)	77 (24%)	0	4
1	B	322/335 (96%)	237 (74%)	85 (26%)	0	3
1	C	314/335 (94%)	238 (76%)	76 (24%)	0	4
1	D	298/335 (89%)	214 (72%)	84 (28%)	0	2
1	E	322/335 (96%)	241 (75%)	81 (25%)	0	3
1	F	311/335 (93%)	228 (73%)	83 (27%)	0	3
All	All	1888/2010 (94%)	1402 (74%)	486 (26%)	0	3

5 of 486 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	335	GLN
1	F	192	THR
1	D	234	ASN
1	F	179	THR
1	F	328	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 79 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	59	GLN
1	F	59	GLN
1	E	104	GLN
1	E	297	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	317	ASN

### 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 ⓘ

18 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	BGC	G	1	2	12,12,12	0.66	0	17,17,17	0.87	1 (5%)
2	GAL	G	2	2	11,11,12	0.33	0	15,15,17	1.08	1 (6%)
2	SIA	G	3	2	20,20,21	0.81	0	21,28,31	1.07	2 (9%)
2	BGC	H	1	2	12,12,12	0.69	0	17,17,17	0.75	1 (5%)
2	GAL	H	2	2	11,11,12	0.50	0	15,15,17	1.45	1 (6%)
2	SIA	H	3	2	20,20,21	0.81	0	21,28,31	1.06	1 (4%)
2	BGC	I	1	2	12,12,12	0.70	0	17,17,17	0.81	1 (5%)
2	GAL	I	2	2	11,11,12	0.40	0	15,15,17	1.59	1 (6%)
2	SIA	I	3	2	20,20,21	0.79	0	21,28,31	0.94	1 (4%)
2	BGC	J	1	2	12,12,12	0.68	0	17,17,17	0.84	1 (5%)
2	GAL	J	2	2	11,11,12	0.24	0	15,15,17	1.33	1 (6%)
2	SIA	J	3	2	20,20,21	1.09	1 (5%)	21,28,31	1.36	2 (9%)
2	BGC	K	1	2	12,12,12	0.79	0	17,17,17	1.06	1 (5%)
2	GAL	K	2	2	11,11,12	0.25	0	15,15,17	1.36	1 (6%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	SIA	K	3	2	20,20,21	0.84	0	21,28,31	1.10	2 (9%)
2	BGC	L	1	2	12,12,12	0.82	0	17,17,17	0.87	1 (5%)
2	GAL	L	2	2	11,11,12	0.35	0	15,15,17	1.06	1 (6%)
2	SIA	L	3	2	20,20,21	1.09	1 (5%)	21,28,31	1.23	2 (9%)

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	BGC	G	1	2	-	0/2/22/22	0/1/1/1
2	GAL	G	2	2	-	0/2/19/22	0/1/1/1
2	SIA	G	3	2	-	2/18/34/38	0/1/1/1
2	BGC	H	1	2	-	0/2/22/22	0/1/1/1
2	GAL	H	2	2	-	0/2/19/22	0/1/1/1
2	SIA	H	3	2	-	1/18/34/38	0/1/1/1
2	BGC	I	1	2	-	0/2/22/22	0/1/1/1
2	GAL	I	2	2	-	0/2/19/22	0/1/1/1
2	SIA	I	3	2	-	7/18/34/38	0/1/1/1
2	BGC	J	1	2	-	1/2/22/22	0/1/1/1
2	GAL	J	2	2	-	0/2/19/22	0/1/1/1
2	SIA	J	3	2	-	3/18/34/38	0/1/1/1
2	BGC	K	1	2	-	0/2/22/22	0/1/1/1
2	GAL	K	2	2	-	1/2/19/22	0/1/1/1
2	SIA	K	3	2	-	4/18/34/38	0/1/1/1
2	BGC	L	1	2	-	1/2/22/22	0/1/1/1
2	GAL	L	2	2	-	0/2/19/22	0/1/1/1
2	SIA	L	3	2	-	7/18/34/38	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	3	SIA	C2-C1	3.07	1.56	1.52
2	J	3	SIA	C2-C1	2.83	1.55	1.52

The worst 5 of 22 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	2	GAL	C1-C2-C3	-5.32	101.90	109.64

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	2	GAL	C1-C2-C3	-4.85	102.59	109.64
2	J	2	GAL	C1-C2-C3	-4.83	102.62	109.64
2	K	2	GAL	C1-C2-C3	-4.67	102.84	109.64
2	J	3	SIA	C8-C7-C6	-4.48	104.64	113.05

There are no chirality outliers.

5 of 27 torsion outliers are listed below:

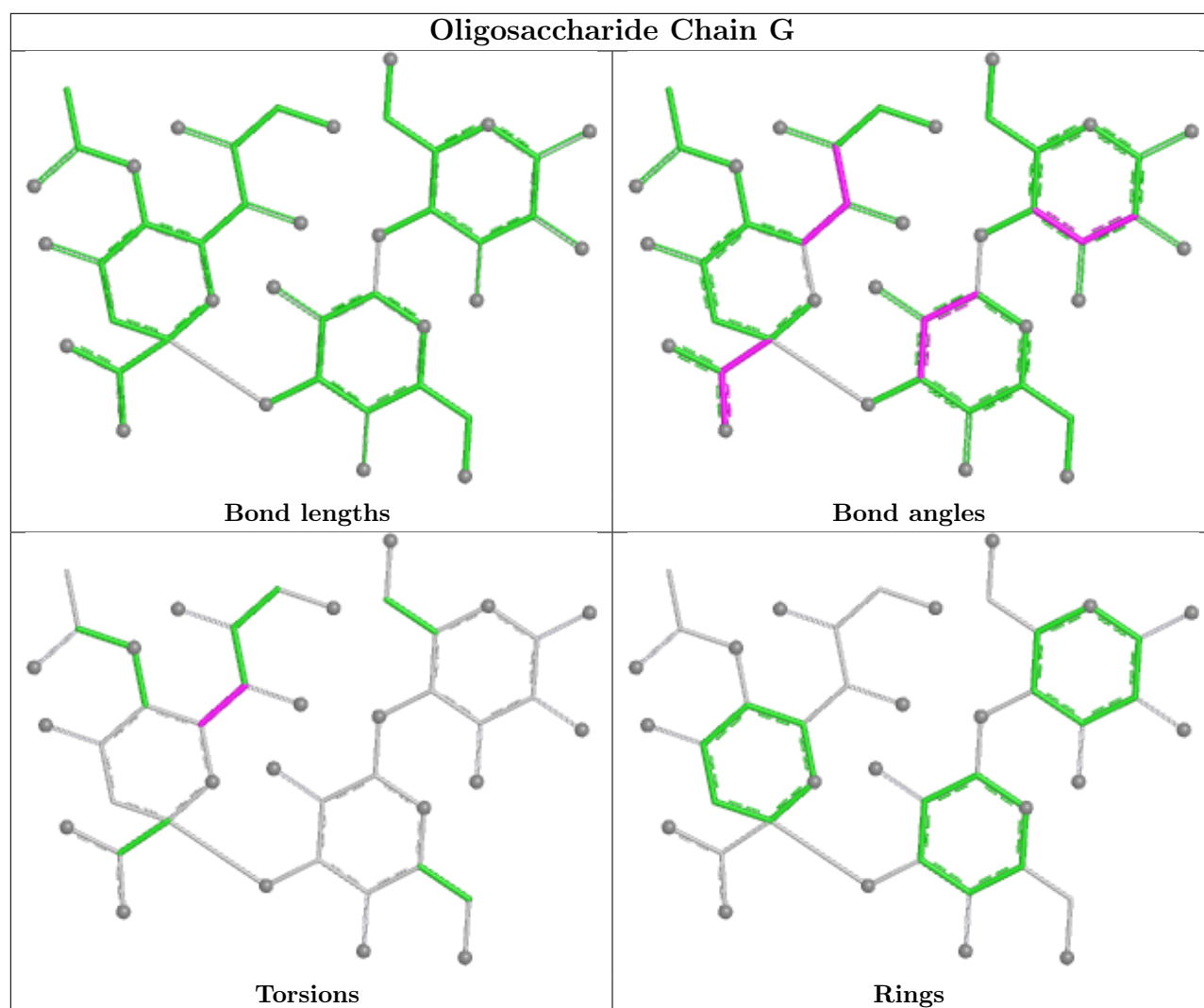
Mol	Chain	Res	Type	Atoms
2	I	3	SIA	C6-C7-C8-C9
2	I	3	SIA	C6-C7-C8-O8
2	I	3	SIA	O7-C7-C8-C9
2	I	3	SIA	O7-C7-C8-O8
2	J	3	SIA	C5-C6-C7-O7

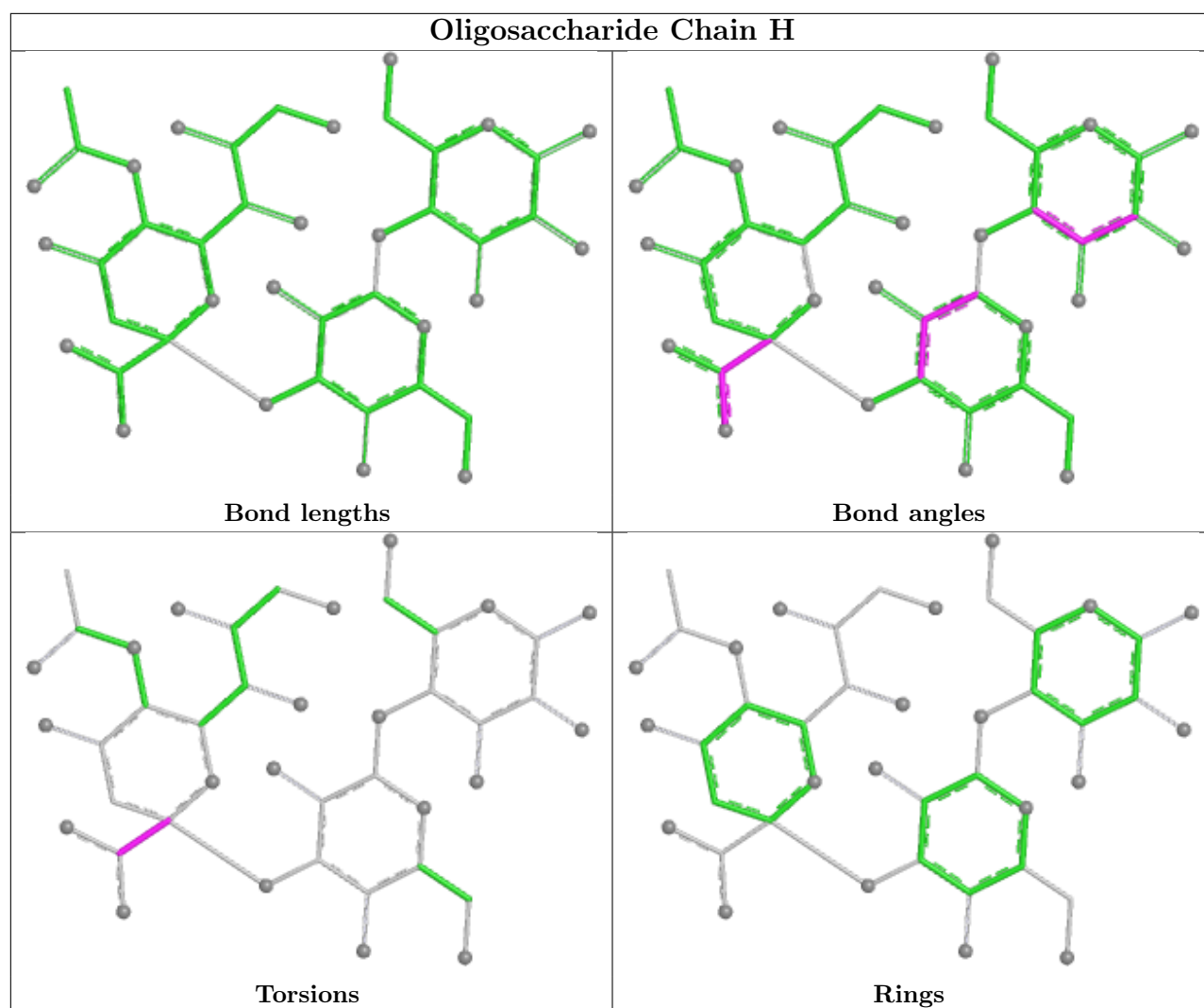
There are no ring outliers.

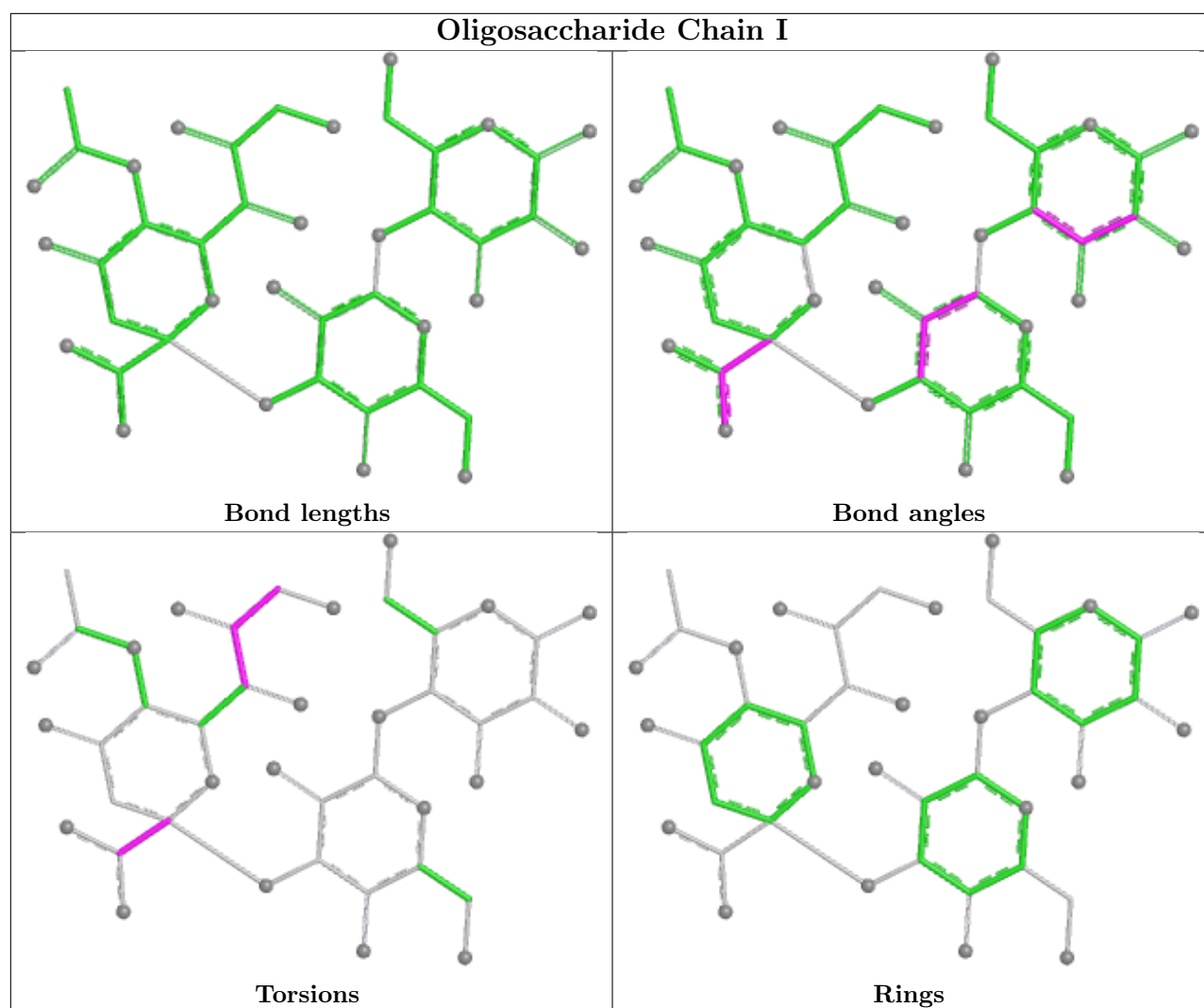
11 monomers are involved in 18 short contacts:

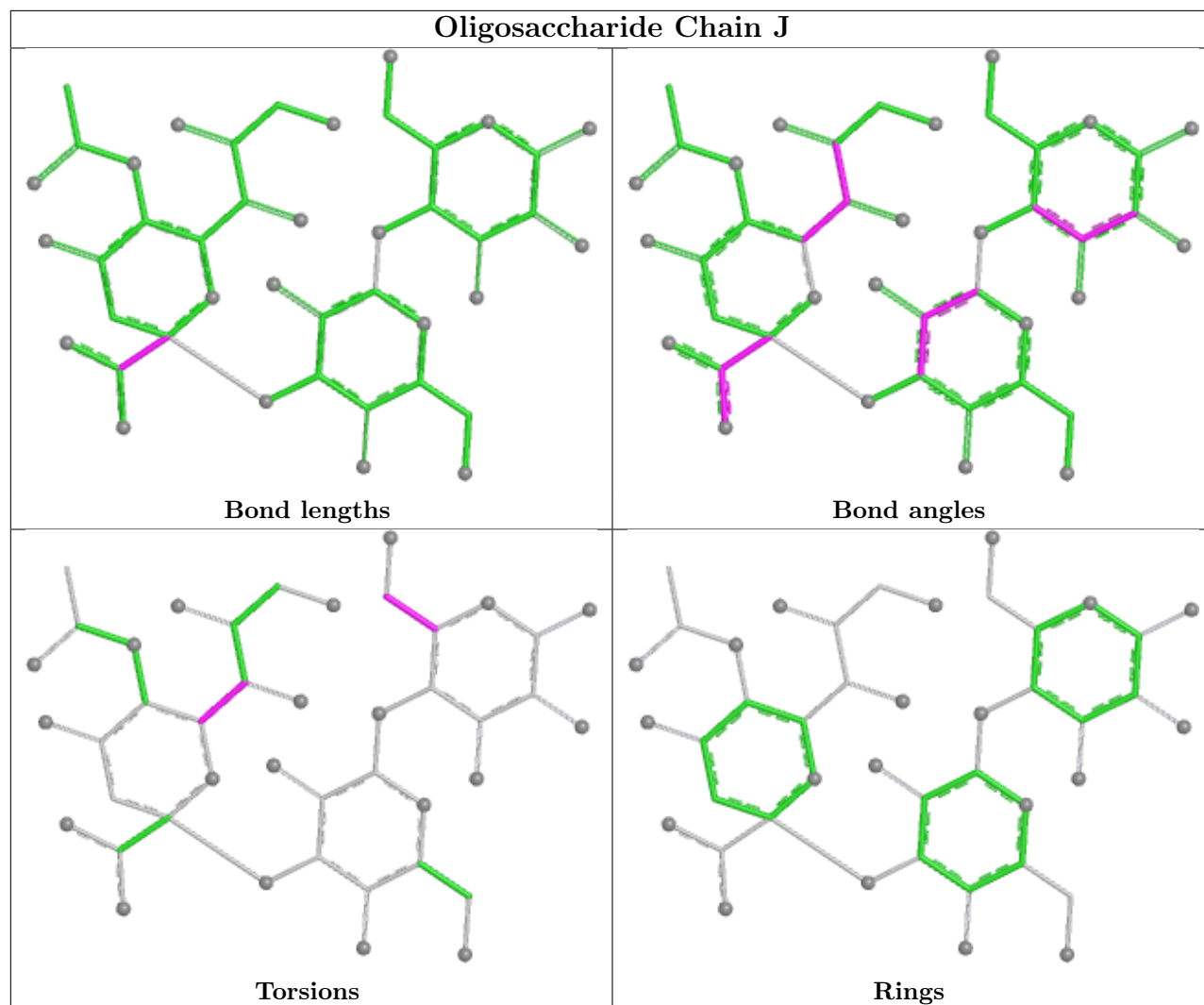
Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	K	3	SIA	2	0
2	J	2	GAL	1	0
2	I	2	GAL	3	0
2	L	2	GAL	4	0
2	L	3	SIA	2	0
2	G	3	SIA	1	0
2	H	2	GAL	4	0
2	H	3	SIA	2	0
2	K	2	GAL	2	0
2	I	3	SIA	1	0
2	J	3	SIA	1	0

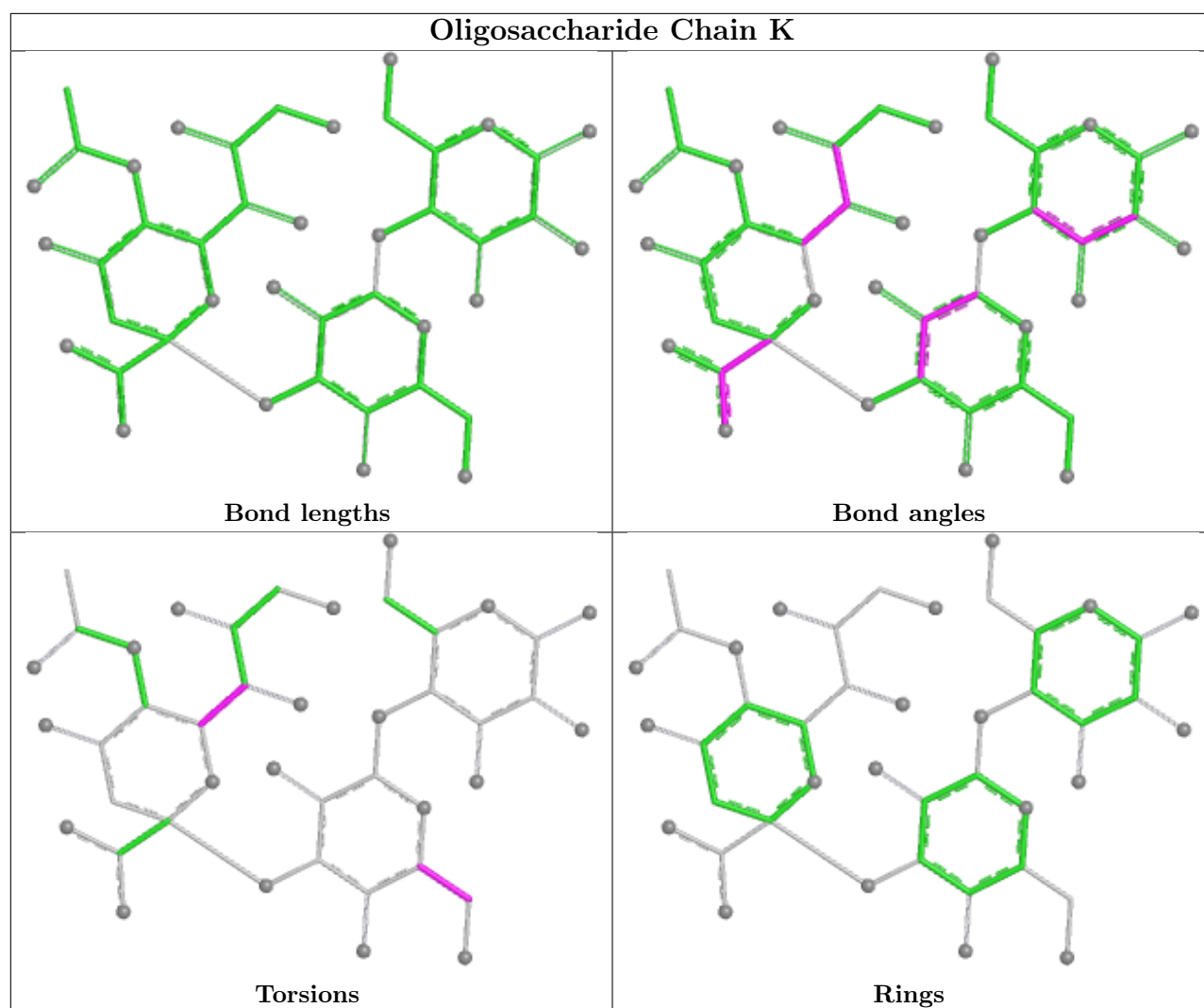
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

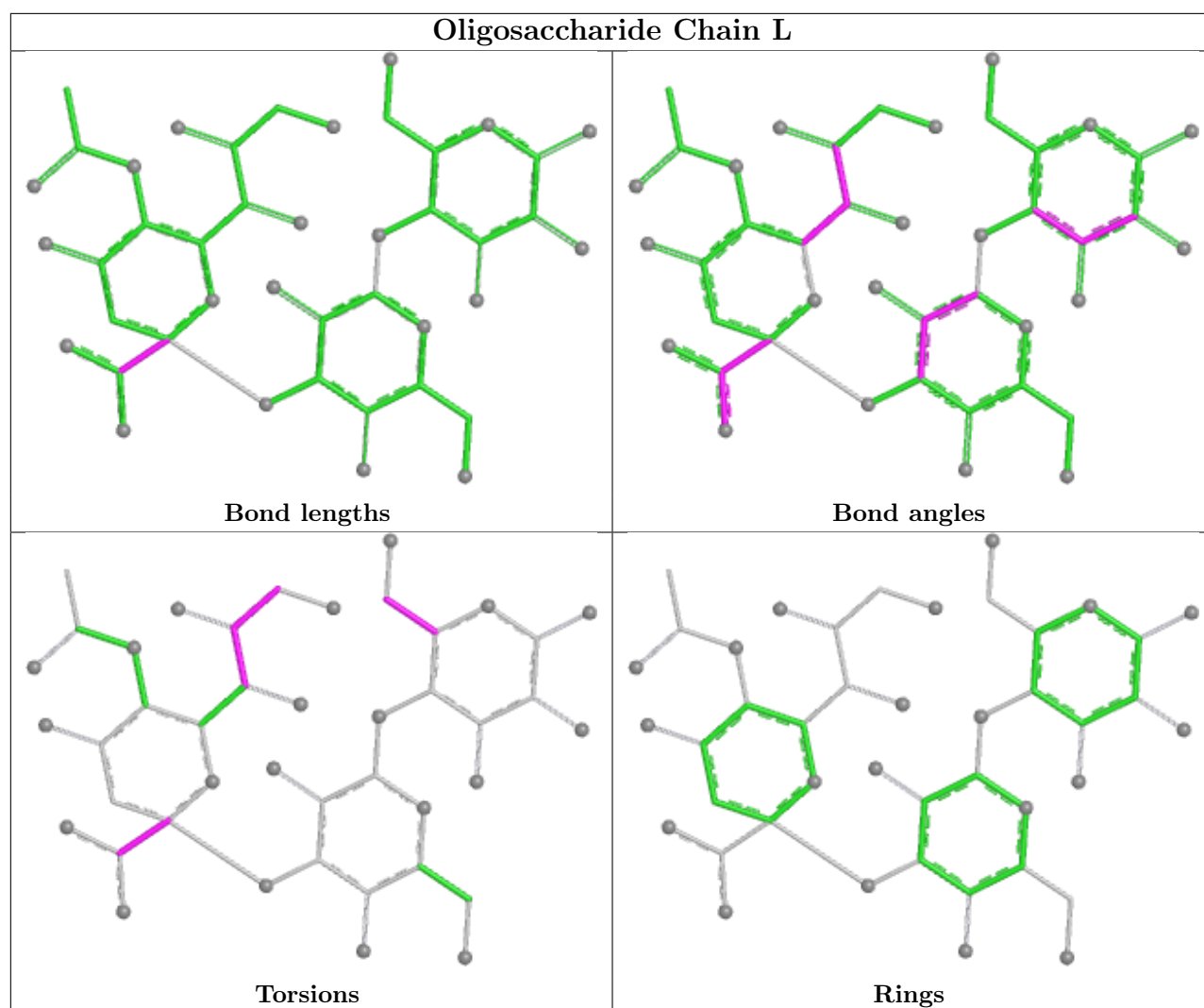












## 5.6 Ligand geometry [i](#)

There are no ligands in this entry.

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	D	1
1	B	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Number of breaks
1	F	1
1	C	1
1	E	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	D	336:VAL	C	337:GLN	N	4.77
1	B	336:VAL	C	337:GLN	N	3.16
1	F	336:VAL	C	337:GLN	N	2.96
1	C	336:VAL	C	337:GLN	N	2.91
1	E	336:VAL	C	337:GLN	N	1.64



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

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

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.