



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 6, 2024 – 07:56 pm GMT

PDB ID : 5NZF  
Title : Complex of H275Y/I223V mutant variant of neuraminidase from H1N1 influenza virus with oseltamivir  
Authors : Pachl, P.; Pokorna, J.  
Deposited on : 2017-05-14  
Resolution : 1.75 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

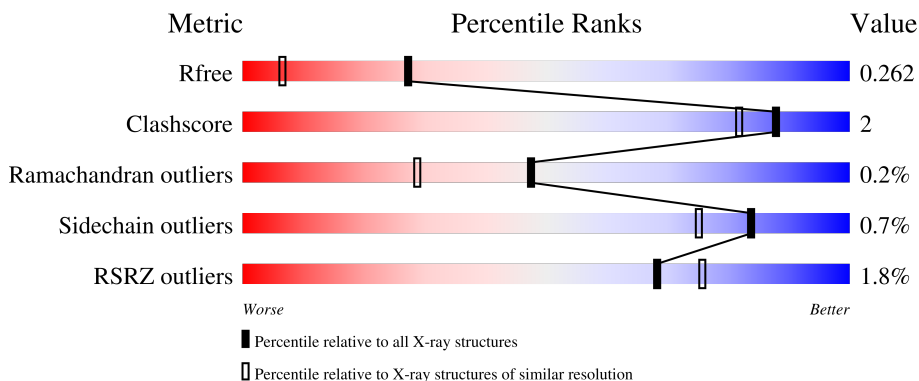
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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	2340 (1.76-1.76)
Clashscore	141614	2466 (1.76-1.76)
Ramachandran outliers	138981	2437 (1.76-1.76)
Sidechain outliers	138945	2437 (1.76-1.76)
RSRZ outliers	127900	2298 (1.76-1.76)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	388	
1	B	388	
1	C	388	
1	D	388	
2	E	5	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	2	 100%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	MAN	E	4	-	-	-	X
3	NAG	F	2	X	-	-	-
6	NAG	D	504	-	-	-	X

## 2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 13687 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.

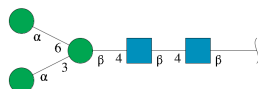
- Molecule 1 is a protein called Neuraminidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	387	Total 3005	C 1887	N 516	O 581	S 21	0	3	0
1	B	387	Total 2995	C 1881	N 514	O 579	S 21	0	2	0
1	C	387	Total 3000	C 1885	N 515	O 579	S 21	0	2	0
1	D	387	Total 3001	C 1884	N 515	O 581	S 21	0	1	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	223	VAL	ILE	conflict	UNP W5R8B8
A	275	TYR	HIS	conflict	UNP W5R8B8
B	223	VAL	ILE	conflict	UNP W5R8B8
B	275	TYR	HIS	conflict	UNP W5R8B8
C	223	VAL	ILE	conflict	UNP W5R8B8
C	275	TYR	HIS	conflict	UNP W5R8B8
D	223	VAL	ILE	conflict	UNP W5R8B8
D	275	TYR	HIS	conflict	UNP W5R8B8

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	5	Total 61	C 34	N 2	O 25	0	0	0

- Molecule 3 is an oligosaccharide called 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.

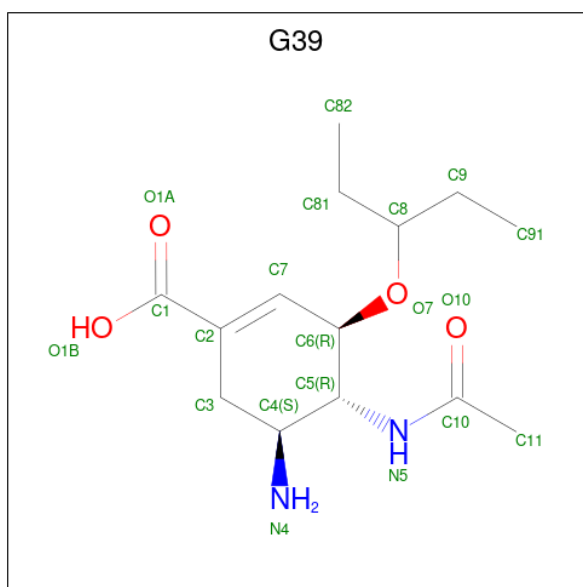


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace	
			Total	C	N				O
3	F	2	28	16	2	10	0	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

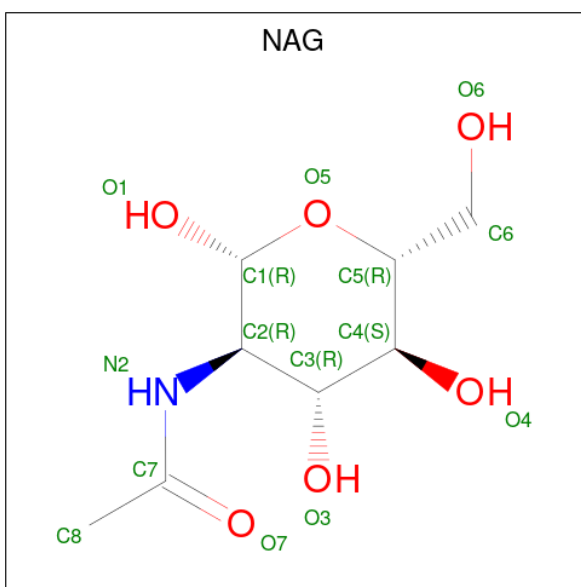
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	3	Total	Ca	0	0
			3	3		
4	B	2	Total	Ca	0	0
			2	2		
4	C	2	Total	Ca	0	0
			2	2		
4	D	2	Total	Ca	0	0
			2	2		

- Molecule 5 is (3R,4R,5S)-4-(acetylamino)-5-amino-3-(pentan-3-yloxy)cyclohex-1-ene-1-carboxylic acid (three-letter code: G39) (formula: C<sub>14</sub>H<sub>24</sub>N<sub>2</sub>O<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			20	14	2	4		
5	B	1	Total	C	N	O	0	0
			20	14	2	4		
5	C	1	Total	C	N	O	0	0
			20	14	2	4		
5	D	1	Total	C	N	O	0	0
			20	14	2	4		

- Molecule 6 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



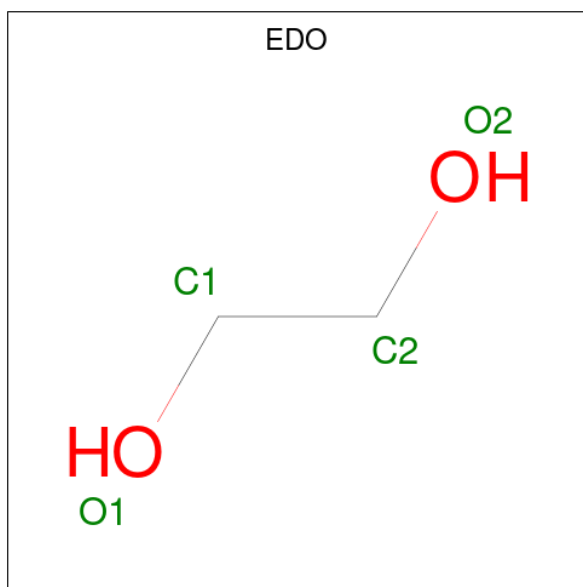
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	A	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	B	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		
6	C	1	Total	C	N	O	0	0
			14	8	1	5		

*Continued on next page...*

Continued from previous page...

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

- Molecule 7 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

Continued on next page...

*Continued from previous page...*

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

- Molecule 8 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	1	Total	Na	0	0
			1	1		

- Molecule 9 is water.

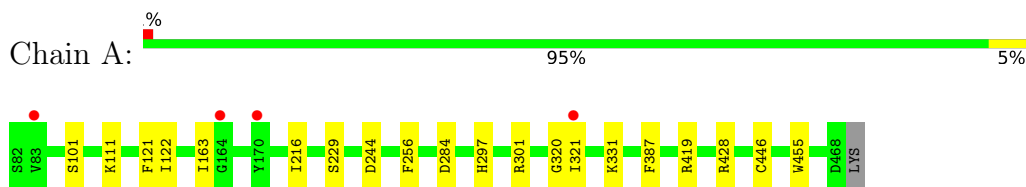
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	336	Total	O	0	2
			338	338		
9	B	270	Total	O	0	1
			271	271		
9	C	351	Total	O	0	2
			353	353		
9	D	368	Total	O	0	8
			375	375		



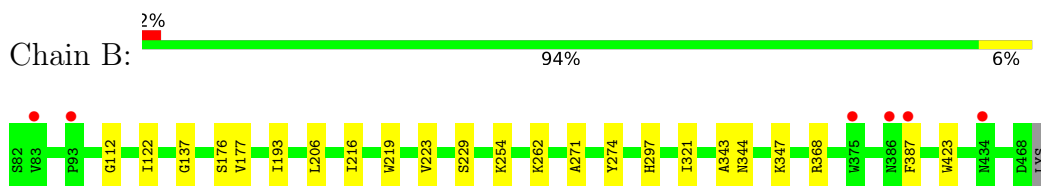
### 3 Residue-property plots [i](#)

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.

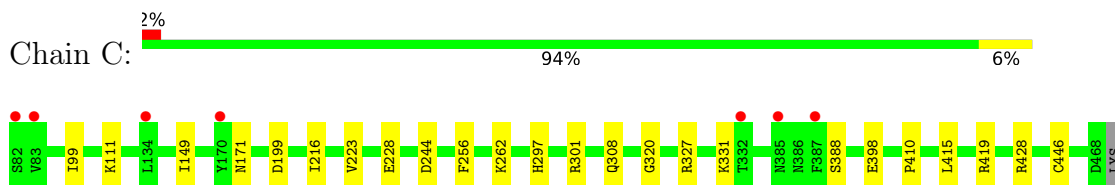
- Molecule 1: Neuraminidase



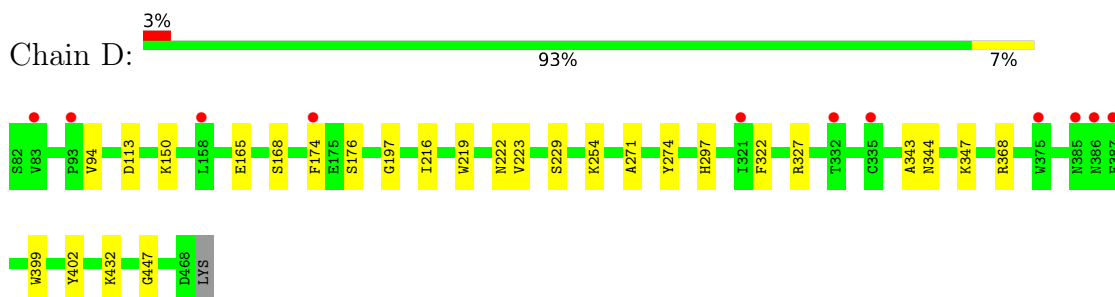
- Molecule 1: Neuraminidase



- Molecule 1: Neuraminidase



- Molecule 1: Neuraminidase



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





- Molecule 3: 2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose

Chain F:

100%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	85.65Å 127.22Å 96.79Å 90.00° 92.93° 90.00°	Depositor
Resolution (Å)	45.18 – 1.75 45.18 – 1.75	Depositor EDS
% Data completeness (in resolution range)	99.2 (45.18-1.75) 99.2 (45.18-1.75)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 1.75Å)	Xtrriage
Refinement program	REFMAC 5.8.0155	Depositor
R, $R_{free}$	0.214 , 0.255 0.222 , 0.262	Depositor DCC
$R_{free}$ test set	2183 reflections (1.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	16.6	Xtrriage
Anisotropy	0.090	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 45.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.008 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13687	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	19.0	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 21.73 % 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 6.6035e-03. 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: G39, MAN, NA, NAG, CA, EDO, BMA

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	0.84	0/3096	0.88	4/4209 (0.1%)
1	B	0.82	0/3088	0.85	1/4199 (0.0%)
1	C	0.85	1/3089 (0.0%)	0.90	6/4198 (0.1%)
1	D	0.87	0/3083	0.88	1/4191 (0.0%)
All	All	0.85	1/12356 (0.0%)	0.88	12/16797 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	228	GLU	CD-OE1	5.53	1.31	1.25

All (12) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	368	ARG	NE-CZ-NH2	-6.79	116.91	120.30
1	D	368	ARG	NE-CZ-NH2	-6.69	116.96	120.30
1	C	199	ASP	CB-CG-OD1	6.62	124.26	118.30
1	A	244	ASP	CB-CG-OD1	6.39	124.05	118.30
1	A	284	ASP	CB-CG-OD1	6.11	123.80	118.30
1	C	301	ARG	NE-CZ-NH1	6.03	123.31	120.30
1	A	301	ARG	NE-CZ-NH1	5.79	123.20	120.30
1	C	327	ARG	NE-CZ-NH2	-5.67	117.46	120.30
1	C	428	ARG	NE-CZ-NH1	5.64	123.12	120.30
1	C	244	ASP	CB-CG-OD1	5.53	123.28	118.30
1	A	428	ARG	NE-CZ-NH2	-5.37	117.62	120.30
1	C	327	ARG	NE-CZ-NH1	5.09	122.84	120.30

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	A	3005	0	2833	11	0
1	B	2995	0	2817	11	0
1	C	3000	0	2833	11	0
1	D	3001	0	2824	18	0
2	E	61	0	52	0	0
3	F	28	0	25	2	0
4	A	3	0	0	0	0
4	B	2	0	0	0	0
4	C	2	0	0	0	0
4	D	2	0	0	0	0
5	A	20	0	23	0	0
5	B	20	0	23	0	0
5	C	20	0	23	0	0
5	D	20	0	23	0	0
6	A	28	0	26	0	0
6	B	42	0	39	0	0
6	C	28	0	26	0	0
6	D	28	0	26	0	0
7	B	12	0	18	1	0
7	C	8	0	12	3	0
7	D	24	0	36	3	0
8	D	1	0	0	0	0
9	A	338	0	0	2	0
9	B	271	0	0	1	0
9	C	353	0	0	4	0
9	D	375	0	0	2	0
All	All	13687	0	11659	49	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.

All (49) 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:D:222:ASN:OD1	7:D:511:EDO:H21	2.01	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:C:511:EDO:H12	1:D:165:GLU:OE2	2.06	0.55
1:C:410:PRO:HB3	1:C:415:LEU:O	2.06	0.55
3:F:1:NAG:H62	3:F:2:NAG:H3	1.93	0.51
1:B:112:GLY:O	1:C:111:LYS:NZ	2.39	0.50
1:C:149:ILE:HD11	7:C:512:EDO:H21	1.94	0.50
1:B:219:TRP:CZ2	1:B:254:LYS:HG3	2.46	0.50
1:A:321:ILE:HD11	1:A:387:PHE:HB3	1.94	0.49
1:D:344:ASN:HB3	9:D:828:HOH:O	2.13	0.49
1:B:122:ILE:HD12	1:B:423:TRP:HB3	1.95	0.48
1:D:222:ASN:CG	7:D:511:EDO:H21	2.34	0.48
1:C:398:GLU:OE1	9:C:601:HOH:O	2.20	0.48
1:A:320:GLY:HA3	1:A:331:LYS:O	2.13	0.47
1:D:113:ASP:O	1:D:168:SER:HB2	2.15	0.47
1:A:216:ILE:HD13	1:A:256:PHE:CE2	2.50	0.47
1:A:455:TRP:CE3	1:D:197:GLY:HA2	2.51	0.46
1:B:193:ILE:HG12	1:B:206:LEU:HG	1.97	0.46
1:C:419:ARG:HD3	1:C:446:CYS:HB3	1.97	0.46
1:D:347:LYS:HB3	1:D:402:TYR:CG	2.51	0.45
1:B:344:ASN:HB3	9:B:803:HOH:O	2.16	0.45
1:B:229:SER:HB3	1:B:347:LYS:HE2	1.99	0.45
3:F:1:NAG:H61	3:F:2:NAG:H5	1.98	0.44
1:A:419:ARG:HD3	1:A:446:CYS:HB3	2.00	0.44
1:D:399:TRP:CH2	1:D:432:LYS:HB3	2.52	0.44
1:D:150:LYS:HA	7:D:513:EDO:H21	1.99	0.44
1:B:137:GLY:HA2	9:C:859:HOH:O	2.18	0.44
1:B:321:ILE:HD11	1:B:387:PHE:HB3	2.01	0.43
1:A:331:LYS:HE2	9:A:696:HOH:O	2.18	0.43
1:A:111:LYS:HE2	1:D:113:ASP:OD1	2.19	0.43
1:C:320:GLY:HA3	1:C:331:LYS:O	2.19	0.43
1:D:344:ASN:ND2	9:D:614:HOH:O	2.48	0.42
1:D:322:PHE:HB2	1:D:327:ARG:HD2	2.01	0.42
1:A:101:SER:HA	1:D:174:PHE:CZ	2.55	0.42
1:D:229:SER:HB3	1:D:347:LYS:HE2	2.02	0.42
1:C:171:ASN:C	7:C:511:EDO:H22	2.40	0.42
1:C:308:GLN:HG3	9:C:771:HOH:O	2.20	0.41
1:A:321:ILE:CD1	1:A:387:PHE:HB3	2.51	0.41
1:D:271:ALA:HB1	1:D:274:TYR:HB2	2.01	0.41
1:C:216:ILE:HD13	1:C:256:PHE:CE2	2.55	0.41
1:D:219:TRP:CZ2	1:D:254:LYS:HG3	2.55	0.41
1:C:262:LYS:HD2	9:C:935:HOH:O	2.20	0.41
1:D:343:ALA:O	1:D:344:ASN:HB2	2.21	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:94:VAL:HG11	1:D:447:GLY:HA3	2.04	0.40
1:A:121:PHE:CG	1:A:229:SER:HA	2.57	0.40
1:A:122:ILE:HD13	1:A:163:ILE:CD1	2.51	0.40
1:B:177:VAL:HG21	1:C:99:ILE:HG12	2.03	0.40
1:B:343:ALA:O	1:B:344:ASN:HB2	2.21	0.40
9:A:755:HOH:O	7:B:507:EDO:H12	2.22	0.40
1:B:271:ALA:HB1	1:B:274:TYR:HB2	2.04	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	388/388 (100%)	374 (96%)	14 (4%)	0	100	100
1	B	387/388 (100%)	373 (96%)	13 (3%)	1 (0%)	41	22
1	C	387/388 (100%)	372 (96%)	14 (4%)	1 (0%)	41	22
1	D	386/388 (100%)	371 (96%)	14 (4%)	1 (0%)	41	22
All	All	1548/1552 (100%)	1490 (96%)	55 (4%)	3 (0%)	47	29

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	223	VAL
1	C	223	VAL
1	D	223	VAL

### 5.3.2 Protein sidechains [i](#)

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

resolution.

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	337/335 (101%)	336 (100%)	1 (0%)	92	89
1	B	336/335 (100%)	332 (99%)	4 (1%)	71	56
1	C	336/335 (100%)	334 (99%)	2 (1%)	86	79
1	D	335/335 (100%)	332 (99%)	3 (1%)	78	67
All	All	1344/1340 (100%)	1334 (99%)	10 (1%)	84	75

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	297	HIS
1	B	176	SER
1	B	216	ILE
1	B	262	LYS
1	B	297	HIS
1	C	297	HIS
1	C	388	SER
1	D	176	SER
1	D	216	ILE
1	D	297	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	221	ASN
1	D	344	ASN

### 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

7 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	NAG	E	1	2,1	14,14,15	0.56	0	17,19,21	1.04	0
2	NAG	E	2	2	14,14,15	0.50	0	17,19,21	1.03	1 (5%)
2	BMA	E	3	2	11,11,12	0.58	0	15,15,17	0.77	0
2	MAN	E	4	2	11,11,12	0.53	0	15,15,17	1.40	1 (6%)
2	MAN	E	5	2	11,11,12	0.94	0	15,15,17	0.96	0
3	NAG	F	1	3,1	14,14,15	1.16	2 (14%)	17,19,21	2.97	8 (47%)
3	NAG	F	2	3	14,14,15	0.89	0	17,19,21	2.62	6 (35%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	E	1	2,1	-	0/6/23/26	0/1/1/1
2	NAG	E	2	2	-	0/6/23/26	0/1/1/1
2	BMA	E	3	2	-	0/2/19/22	0/1/1/1
2	MAN	E	4	2	-	2/2/19/22	0/1/1/1
2	MAN	E	5	2	-	1/2/19/22	0/1/1/1
3	NAG	F	1	3,1	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	1/1/5/7	5/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	1	NAG	O5-C1	-2.72	1.39	1.43
3	F	1	NAG	C1-C2	2.09	1.55	1.52

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	1	NAG	C1-C2-N2	-8.91	95.27	110.49
3	F	2	NAG	C2-N2-C7	6.56	132.25	122.90
3	F	2	NAG	C8-C7-N2	4.79	124.20	116.10
3	F	2	NAG	C1-O5-C5	4.08	117.72	112.19
2	E	4	MAN	O5-C5-C6	3.88	113.29	107.20
3	F	1	NAG	C6-C5-C4	3.58	121.40	113.00
3	F	1	NAG	O5-C5-C4	-3.48	102.36	110.83
3	F	1	NAG	C3-C4-C5	-2.86	105.13	110.24
3	F	2	NAG	O3-C3-C2	2.50	114.64	109.47
3	F	1	NAG	O5-C1-C2	2.48	115.20	111.29
3	F	1	NAG	O4-C4-C5	2.45	115.37	109.30
3	F	1	NAG	C4-C3-C2	-2.43	107.45	111.02
3	F	1	NAG	C2-N2-C7	2.25	126.11	122.90
3	F	2	NAG	O7-C7-N2	-2.23	117.85	121.95
3	F	2	NAG	O7-C7-C8	-2.21	117.94	122.06
2	E	2	NAG	O4-C4-C5	-2.08	104.13	109.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	F	2	NAG	C1

All (8) torsion outliers are listed below:

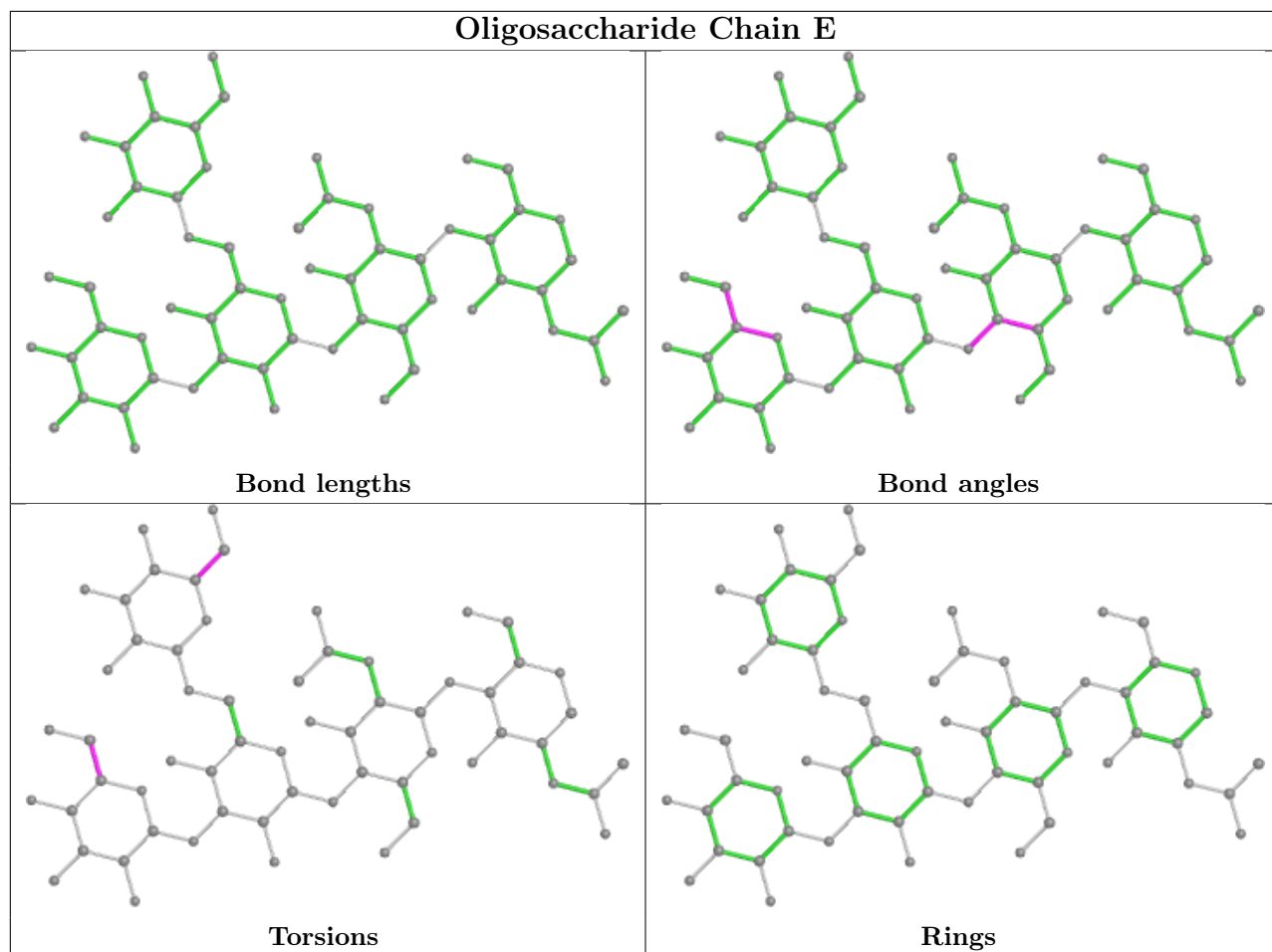
Mol	Chain	Res	Type	Atoms
3	F	2	NAG	O5-C5-C6-O6
3	F	2	NAG	C8-C7-N2-C2
3	F	2	NAG	O7-C7-N2-C2
2	E	4	MAN	O5-C5-C6-O6
3	F	2	NAG	C4-C5-C6-O6
3	F	2	NAG	C3-C2-N2-C7
2	E	5	MAN	C4-C5-C6-O6
2	E	4	MAN	C4-C5-C6-O6

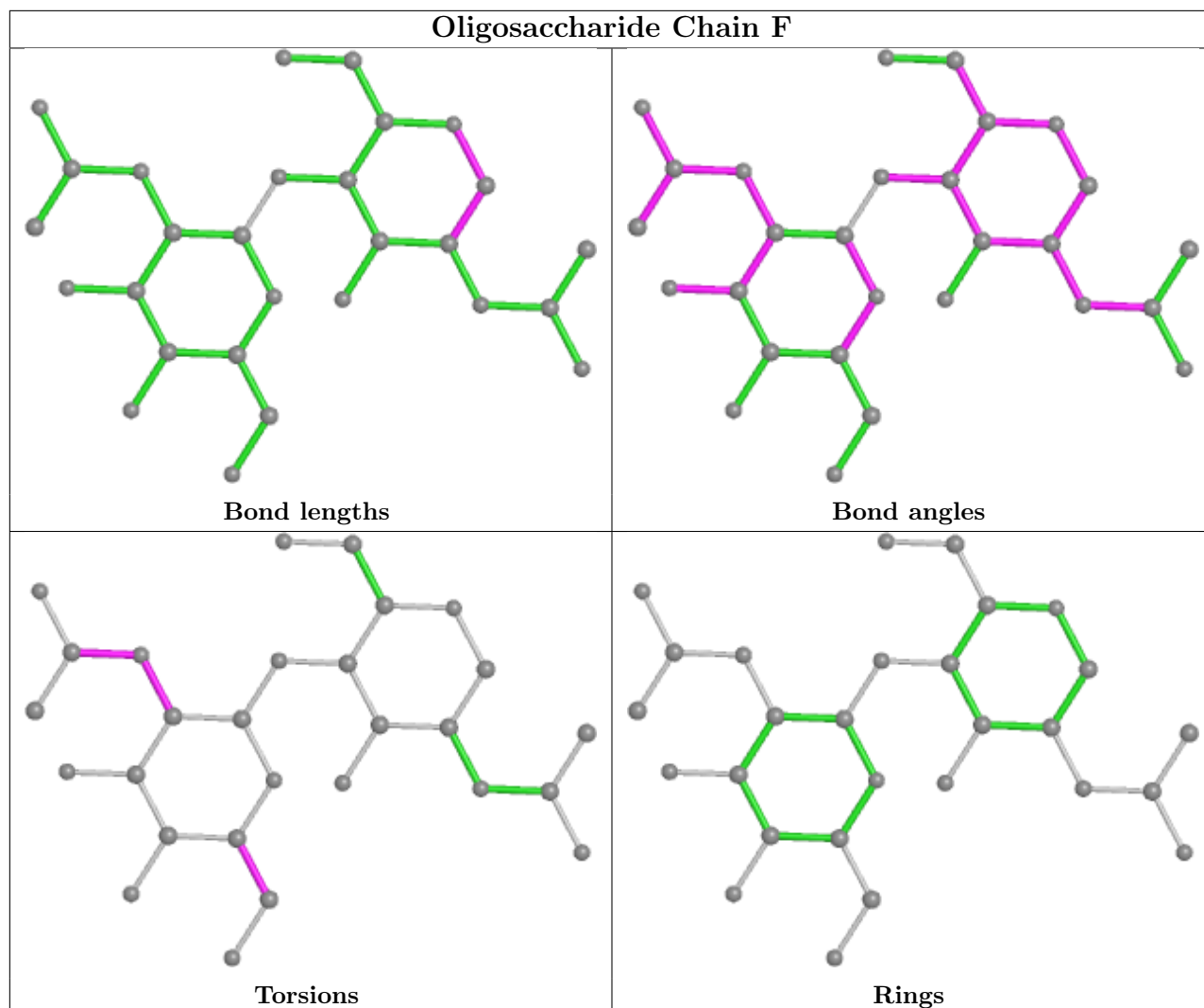
There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	1	NAG	2	0
3	F	2	NAG	2	0

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





## 5.6 Ligand geometry [i](#)

Of 34 ligands modelled in this entry, 10 are monoatomic - leaving 24 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
5	G39	A	503	-	20,20,20	1.00	2 (10%)	19,27,27	1.12	2 (10%)
7	EDO	C	512	-	3,3,3	0.44	0	2,2,2	0.67	0
6	NAG	B	506	1	14,14,15	0.35	0	17,19,21	0.85	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	NAG	D	504	1	14,14,15	0.50	0	17,19,21	1.18	2 (11%)
7	EDO	B	509	-	3,3,3	0.29	0	2,2,2	0.81	0
6	NAG	B	504	1	14,14,15	0.75	0	17,19,21	2.27	3 (17%)
7	EDO	B	507	-	3,3,3	0.60	0	2,2,2	0.36	0
7	EDO	D	509	-	3,3,3	0.45	0	2,2,2	0.41	0
6	NAG	A	504	1	14,14,15	0.57	0	17,19,21	1.89	5 (29%)
6	NAG	A	505	1	14,14,15	0.48	0	17,19,21	0.85	0
5	G39	D	503	-	20,20,20	0.99	1 (5%)	19,27,27	1.13	1 (5%)
7	EDO	D	510	-	3,3,3	0.68	0	2,2,2	0.28	0
6	NAG	B	505	1	14,14,15	0.53	0	17,19,21	1.36	2 (11%)
5	G39	B	503	-	20,20,20	1.09	2 (10%)	19,27,27	1.54	2 (10%)
6	NAG	C	510	1	14,14,15	0.51	0	17,19,21	0.80	1 (5%)
7	EDO	D	513	-	3,3,3	0.83	0	2,2,2	0.25	0
6	NAG	D	507	1	14,14,15	0.46	0	17,19,21	0.57	0
7	EDO	D	512	-	3,3,3	0.33	0	2,2,2	0.84	0
7	EDO	D	511	-	3,3,3	0.24	0	2,2,2	1.08	0
6	NAG	C	504	1	14,14,15	0.67	0	17,19,21	2.14	4 (23%)
5	G39	C	503	-	20,20,20	0.82	1 (5%)	19,27,27	1.49	6 (31%)
7	EDO	D	508	-	3,3,3	0.60	0	2,2,2	0.26	0
7	EDO	C	511	-	3,3,3	0.53	0	2,2,2	0.42	0
7	EDO	B	508	-	3,3,3	0.42	0	2,2,2	1.09	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
5	G39	A	503	-	-	0/16/32/32	0/1/1/1
7	EDO	C	512	-	-	0/1/1/1	-
6	NAG	B	506	1	-	0/6/23/26	0/1/1/1
6	NAG	D	504	1	-	0/6/23/26	0/1/1/1
7	EDO	B	509	-	-	1/1/1/1	-
6	NAG	B	504	1	-	1/6/23/26	0/1/1/1
7	EDO	B	507	-	-	1/1/1/1	-
7	EDO	D	509	-	-	0/1/1/1	-
6	NAG	A	504	1	-	2/6/23/26	0/1/1/1
6	NAG	A	505	1	-	0/6/23/26	0/1/1/1
5	G39	D	503	-	-	0/16/32/32	0/1/1/1
7	EDO	D	510	-	-	1/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	NAG	B	505	1	-	2/6/23/26	0/1/1/1
5	G39	B	503	-	-	0/16/32/32	0/1/1/1
6	NAG	C	510	1	-	0/6/23/26	0/1/1/1
7	EDO	D	513	-	-	1/1/1/1	-
6	NAG	D	507	1	-	0/6/23/26	0/1/1/1
7	EDO	D	512	-	-	0/1/1/1	-
7	EDO	D	511	-	-	1/1/1/1	-
6	NAG	C	504	1	-	0/6/23/26	0/1/1/1
5	G39	C	503	-	-	0/16/32/32	0/1/1/1
7	EDO	D	508	-	-	0/1/1/1	-
7	EDO	C	511	-	-	0/1/1/1	-
7	EDO	B	508	-	-	1/1/1/1	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	503	G39	C6-C7	2.91	1.54	1.50
5	A	503	G39	C6-C5	-2.42	1.50	1.53
5	A	503	G39	O1A-C1	2.03	1.28	1.22
5	B	503	G39	C3-C2	2.00	1.54	1.51
5	D	503	G39	O10-C10	-2.00	1.18	1.23
5	C	503	G39	O1A-C1	2.00	1.28	1.22

All (29) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	504	NAG	C1-C2-N2	-6.88	98.73	110.49
6	C	504	NAG	C1-C2-N2	-6.56	99.27	110.49
6	A	504	NAG	C1-C2-N2	-4.86	102.18	110.49
6	B	504	NAG	C2-N2-C7	4.10	128.74	122.90
5	B	503	G39	C4-C3-C2	4.08	114.59	109.75
5	A	503	G39	C4-C3-C2	3.18	113.52	109.75
6	A	504	NAG	C4-C3-C2	3.17	115.66	111.02
6	B	505	NAG	C1-C2-N2	2.87	115.39	110.49
5	B	503	G39	O1B-C1-C2	2.82	121.47	115.49
6	C	504	NAG	C4-C3-C2	2.79	115.11	111.02
6	B	504	NAG	O5-C5-C4	-2.67	104.34	110.83
5	C	503	G39	O10-C10-N5	2.62	126.77	121.95
5	C	503	G39	O10-C10-C11	-2.59	117.24	122.06
5	C	503	G39	O1A-C1-C2	-2.59	117.14	121.59
6	A	504	NAG	O5-C5-C6	2.58	111.24	107.20
5	D	503	G39	C4-C3-C2	2.57	112.79	109.75

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	C	503	G39	C4-C3-C2	2.55	112.77	109.75
6	D	504	NAG	O5-C5-C6	2.55	111.21	107.20
6	D	504	NAG	O5-C5-C4	-2.51	104.72	110.83
6	A	504	NAG	O5-C1-C2	2.48	115.21	111.29
6	C	504	NAG	C2-N2-C7	2.31	126.19	122.90
6	B	506	NAG	O5-C1-C2	-2.31	107.65	111.29
5	A	503	G39	C3-C4-N4	-2.27	106.27	110.88
6	A	504	NAG	O5-C5-C4	-2.14	105.61	110.83
5	C	503	G39	C3-C4-N4	-2.08	106.65	110.88
6	C	510	NAG	O5-C1-C2	-2.04	108.07	111.29
6	C	504	NAG	O5-C5-C4	-2.02	105.92	110.83
6	B	505	NAG	O5-C1-C2	-2.01	108.12	111.29
5	C	503	G39	O1B-C1-C2	2.00	119.74	115.49

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	504	NAG	C3-C2-N2-C7
7	D	513	EDO	O1-C1-C2-O2
7	D	510	EDO	O1-C1-C2-O2
6	B	505	NAG	C4-C5-C6-O6
7	B	509	EDO	O1-C1-C2-O2
7	B	508	EDO	O1-C1-C2-O2
6	B	505	NAG	O5-C5-C6-O6
7	B	507	EDO	O1-C1-C2-O2
7	D	511	EDO	O1-C1-C2-O2
6	A	504	NAG	O5-C5-C6-O6
6	A	504	NAG	C4-C5-C6-O6

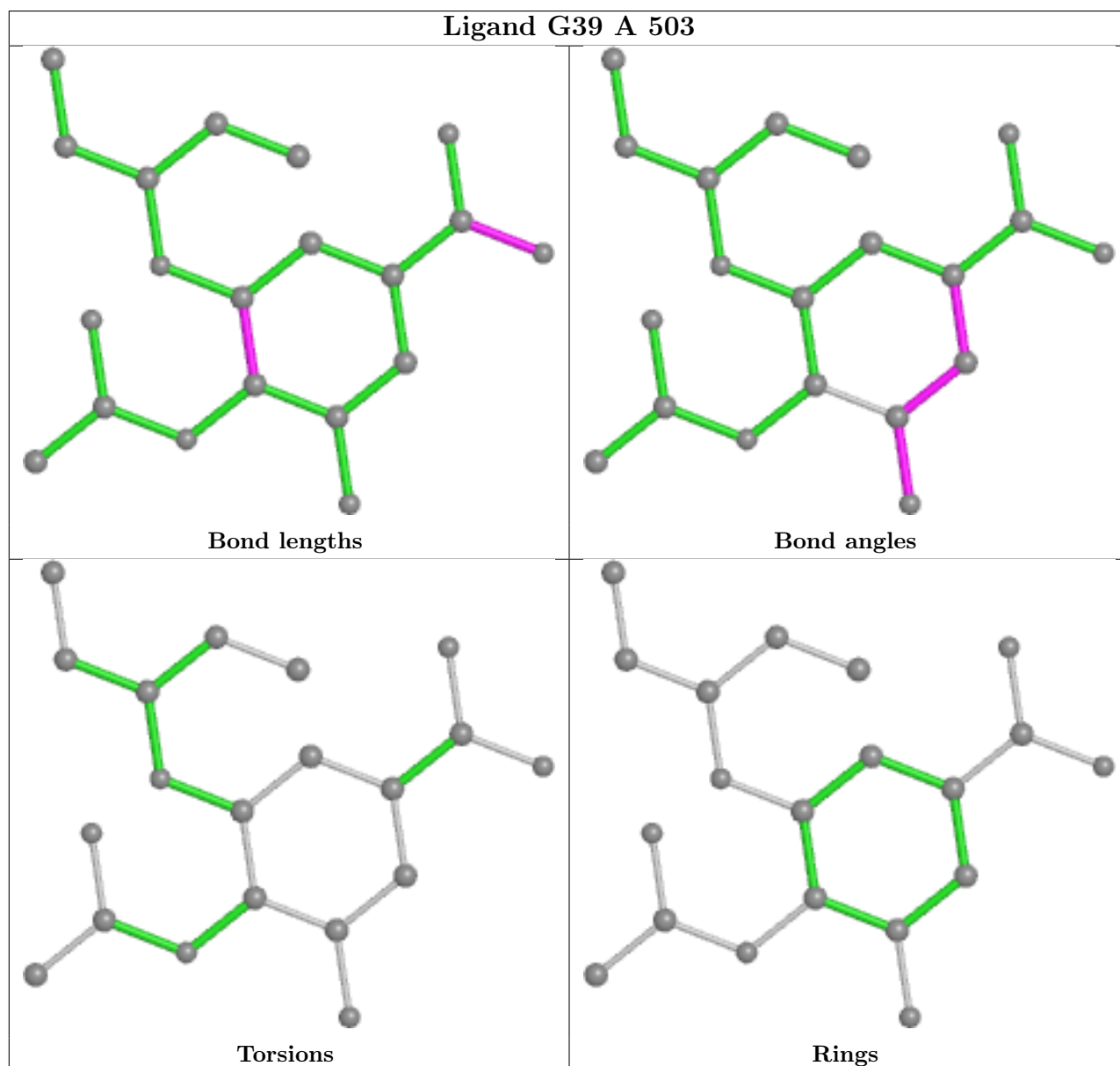
There are no ring outliers.

5 monomers are involved in 7 short contacts:

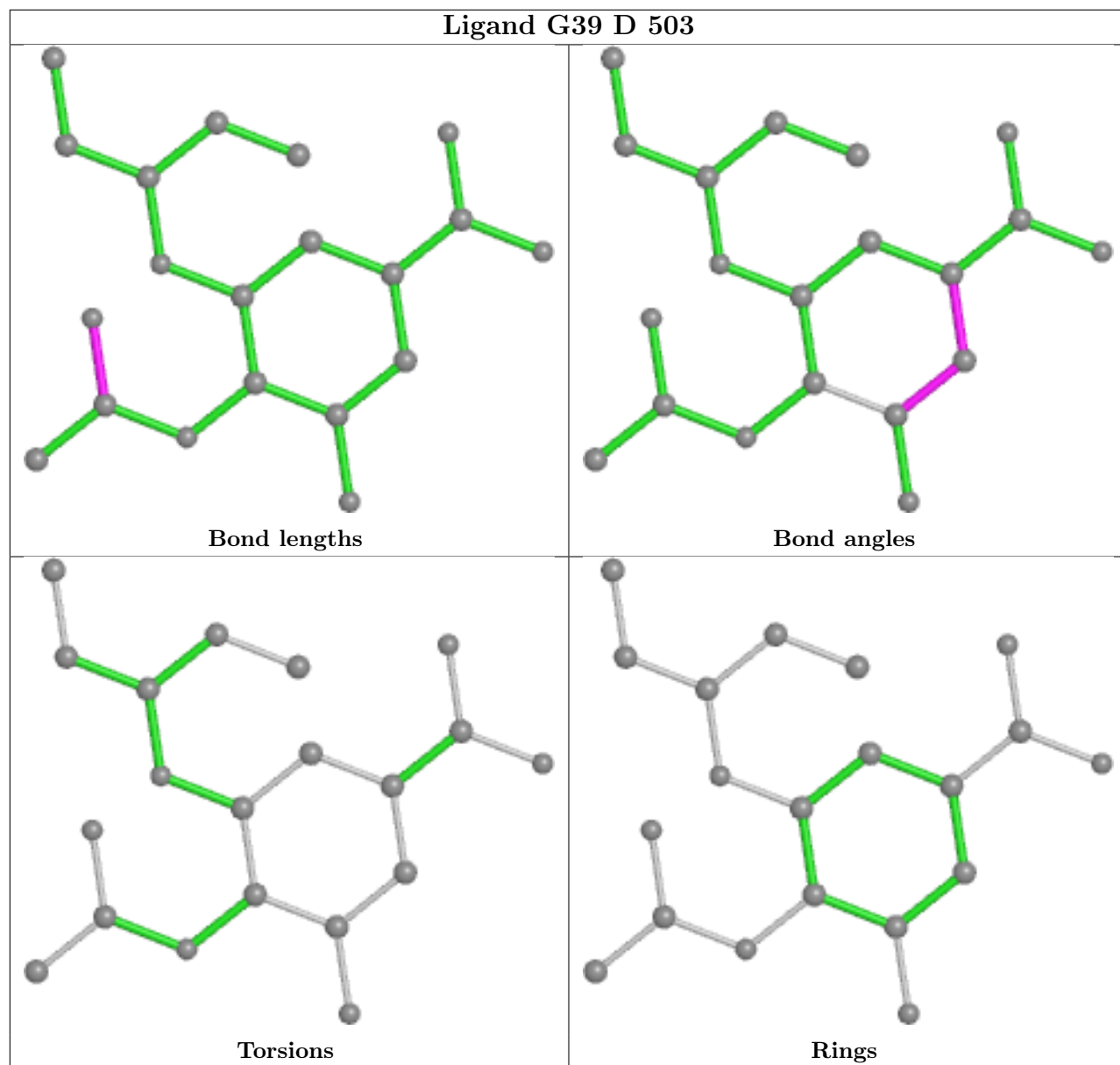
Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	C	512	EDO	1	0
7	B	507	EDO	1	0
7	D	513	EDO	1	0
7	D	511	EDO	2	0
7	C	511	EDO	2	0

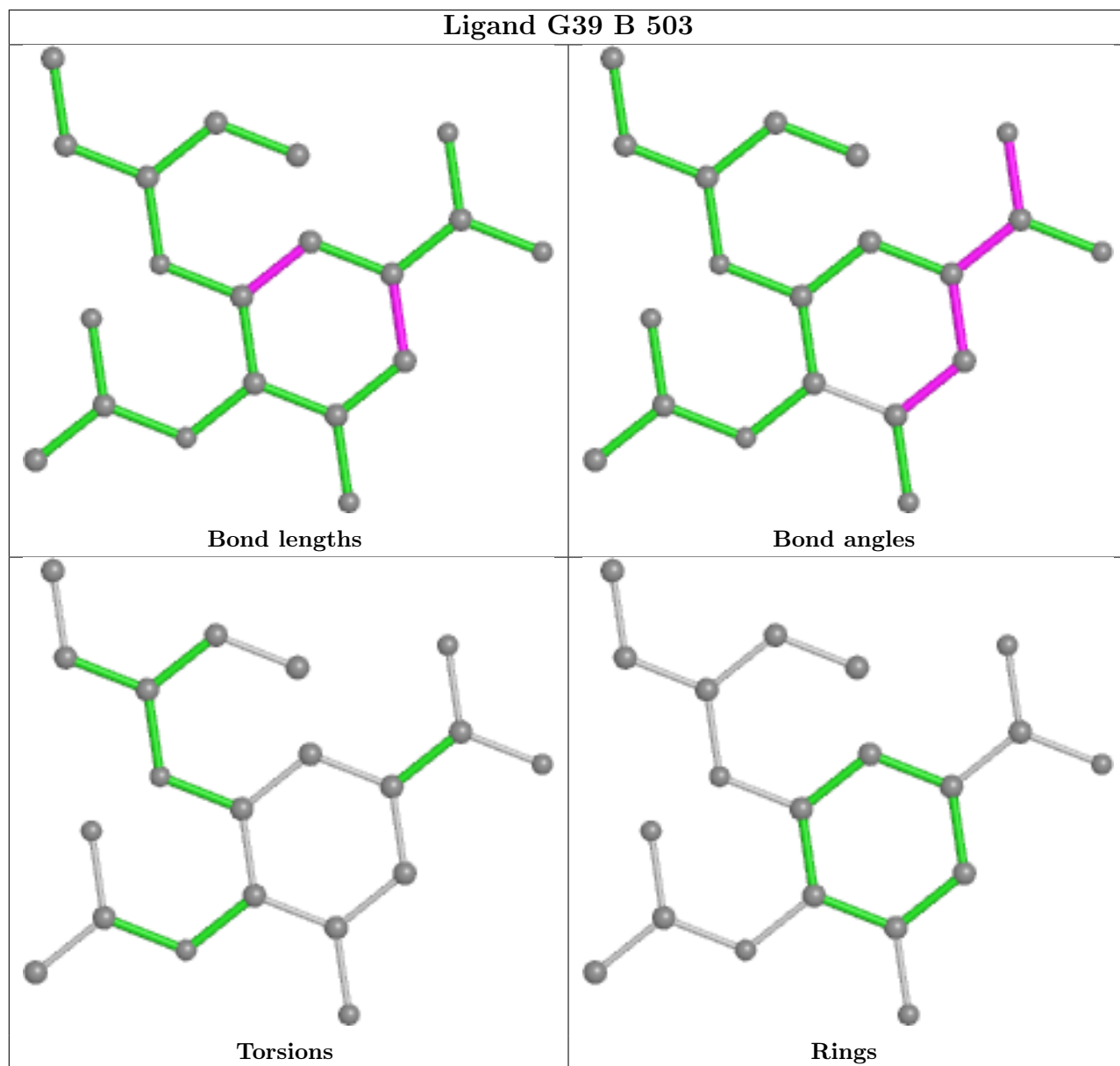
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths,

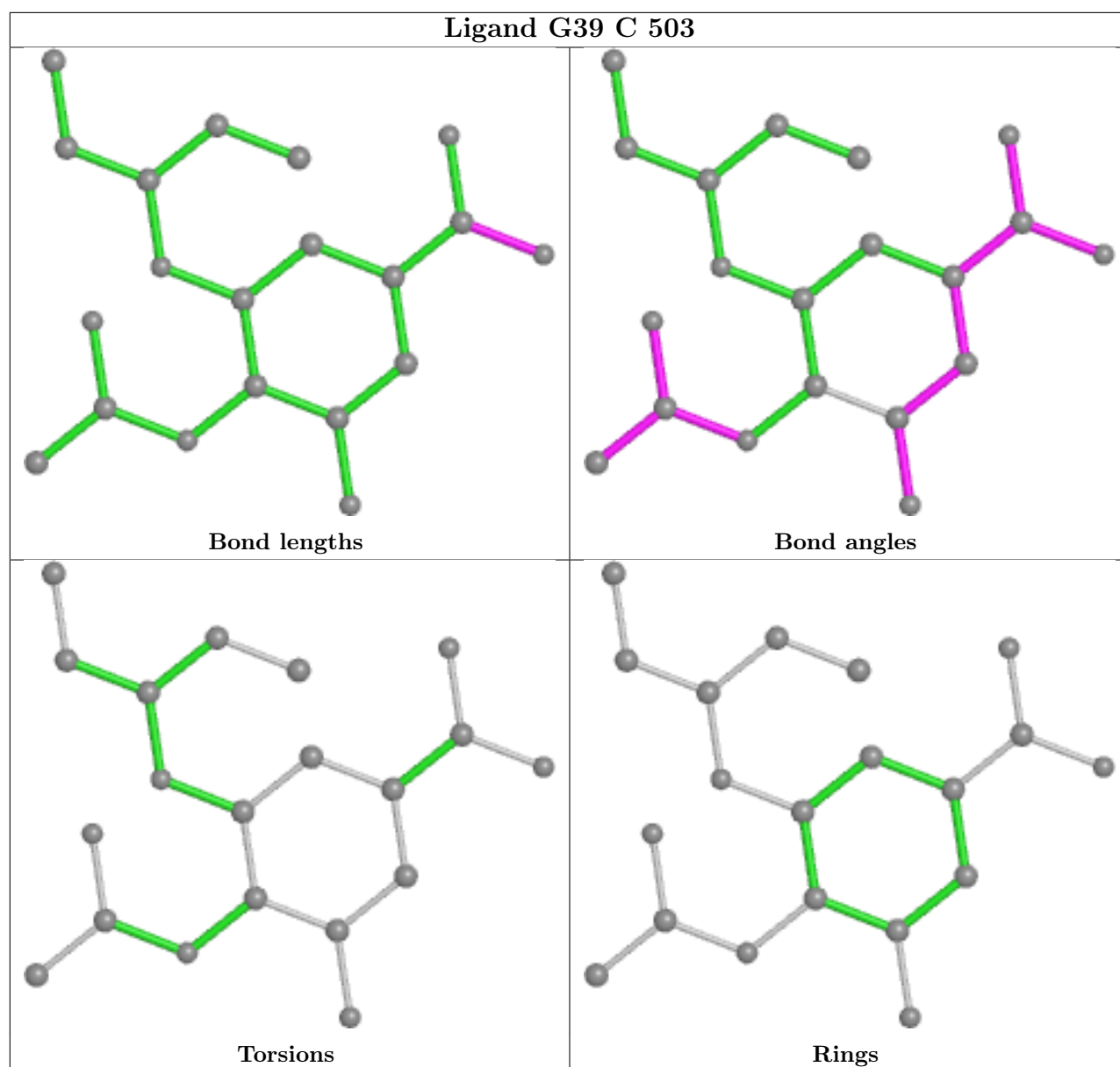
bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within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











## 5.7 Other polymers [i](#)

There are no such residues in this entry.

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

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	387/388 (99%)	0.20	4 (1%) 82 87	10, 16, 25, 36	0
1	B	387/388 (99%)	0.29	6 (1%) 72 79	11, 18, 29, 42	0
1	C	387/388 (99%)	0.28	7 (1%) 68 76	11, 16, 26, 38	0
1	D	387/388 (99%)	0.48	11 (2%) 53 58	11, 18, 29, 43	0
All	All	1548/1552 (99%)	0.31	28 (1%) 68 76	10, 17, 28, 43	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	93	PRO	3.5
1	A	83	VAL	3.3
1	D	387	PHE	3.2
1	D	93	PRO	3.2
1	C	83	VAL	3.1
1	D	83	VAL	2.8
1	B	83	VAL	2.7
1	D	386	ASN	2.7
1	D	158	LEU	2.6
1	B	434	ASN	2.5
1	C	134	LEU	2.5
1	B	386	ASN	2.5
1	D	375	TRP	2.4
1	D	385	ASN	2.4
1	D	335	CYS	2.3
1	D	332	THR	2.3
1	B	387	PHE	2.3
1	C	385	ASN	2.3
1	C	387	PHE	2.2
1	D	321	ILE	2.2
1	A	321	ILE	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	170	TYR	2.2
1	D	174	PHE	2.1
1	C	170	TYR	2.1
1	A	164	GLY	2.0
1	B	375	TRP	2.0
1	C	332	THR	2.0
1	C	82	SER	2.0

## 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 [i](#)

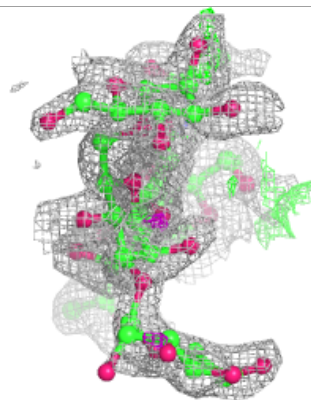
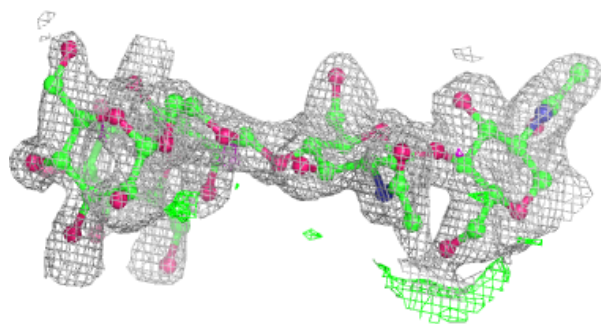
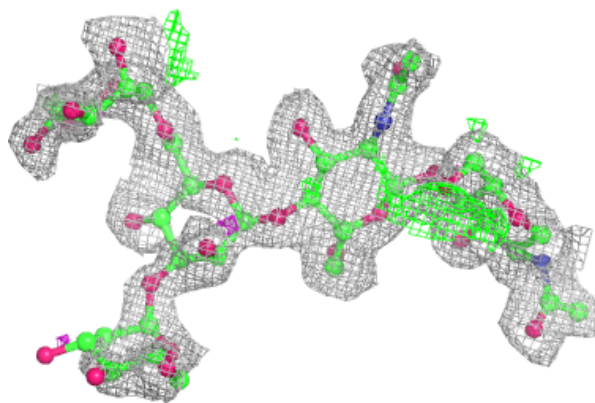
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

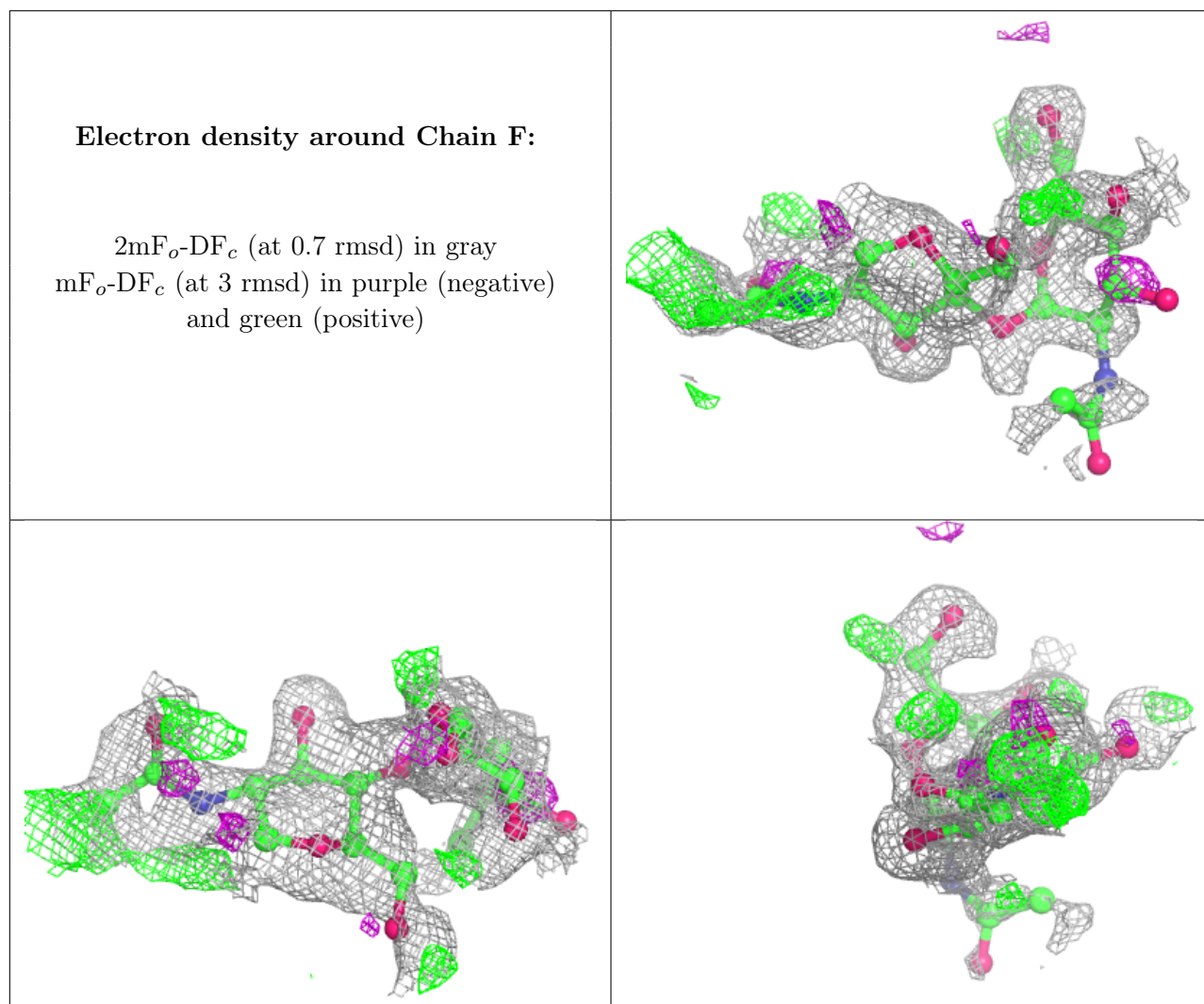
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	F	2	14/15	0.49	0.37	39,50,58,61	0
3	NAG	F	1	14/15	0.66	0.24	38,42,48,49	0
2	MAN	E	4	11/12	0.72	0.50	50,52,55,59	0
2	MAN	E	5	11/12	0.72	0.32	36,40,43,43	0
2	BMA	E	3	11/12	0.73	0.31	36,41,44,48	0
2	NAG	E	1	14/15	0.89	0.14	24,28,33,34	0
2	NAG	E	2	14/15	0.91	0.15	25,29,31,35	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 E:**

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





## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	EDO	D	513	4/4	0.63	0.16	29,29,29,30	0
6	NAG	B	504	14/15	0.65	0.38	45,50,56,57	0
6	NAG	D	504	14/15	0.67	0.40	43,47,57,58	0
6	NAG	A	504	14/15	0.71	0.32	38,44,49,53	0
6	NAG	C	504	14/15	0.71	0.36	41,46,54,56	0
7	EDO	D	510	4/4	0.80	0.15	34,34,34,38	0
6	NAG	B	506	14/15	0.81	0.34	39,45,48,50	0
7	EDO	B	507	4/4	0.82	0.16	28,30,30,34	0

*Continued on next page...*

*Continued from previous page...*

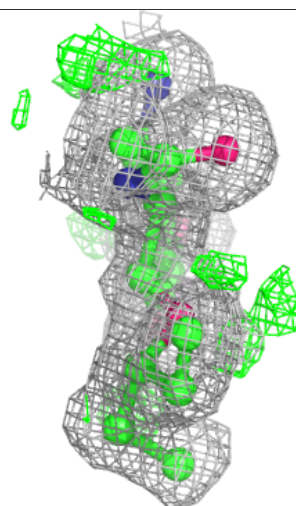
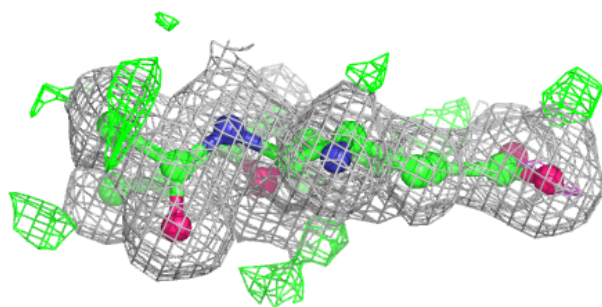
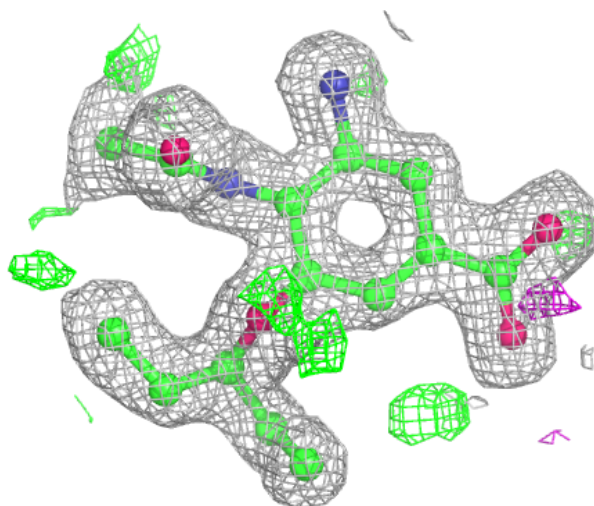
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
7	EDO	C	512	4/4	0.82	0.12	26,27,27,30	0
6	NAG	A	505	14/15	0.83	0.19	31,39,43,46	0
6	NAG	B	505	14/15	0.84	0.23	31,39,39,45	0
6	NAG	D	507	14/15	0.84	0.38	36,42,45,47	0
6	NAG	C	510	14/15	0.84	0.42	38,42,47,47	0
7	EDO	C	511	4/4	0.85	0.16	27,27,28,29	0
7	EDO	B	509	4/4	0.85	0.17	33,33,35,39	0
7	EDO	D	511	4/4	0.86	0.25	31,31,32,35	0
7	EDO	D	508	4/4	0.87	0.19	27,29,29,31	0
7	EDO	B	508	4/4	0.88	0.20	26,26,27,30	0
7	EDO	D	512	4/4	0.91	0.21	25,26,27,28	0
7	EDO	D	509	4/4	0.91	0.11	30,30,30,32	0
5	G39	C	503	20/20	0.92	0.12	11,14,19,20	0
5	G39	B	503	20/20	0.93	0.10	14,16,18,18	0
5	G39	D	503	20/20	0.93	0.10	13,15,18,18	0
5	G39	A	503	20/20	0.95	0.09	10,14,16,17	0
8	NA	D	514	1/1	0.95	0.12	21,21,21,21	0
4	CA	B	502	1/1	0.97	0.14	28,28,28,28	0
4	CA	C	502	1/1	0.98	0.10	25,25,25,25	0
4	CA	D	502	1/1	0.98	0.16	29,29,29,29	0
4	CA	A	502	1/1	0.98	0.10	24,24,24,24	0
4	CA	A	501	1/1	0.99	0.04	14,14,14,14	0
4	CA	C	501	1/1	0.99	0.05	16,16,16,16	0
4	CA	A	506	1/1	0.99	0.05	21,21,21,21	0
4	CA	D	501	1/1	0.99	0.05	22,22,22,22	0
4	CA	B	501	1/1	0.99	0.04	22,22,22,22	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.



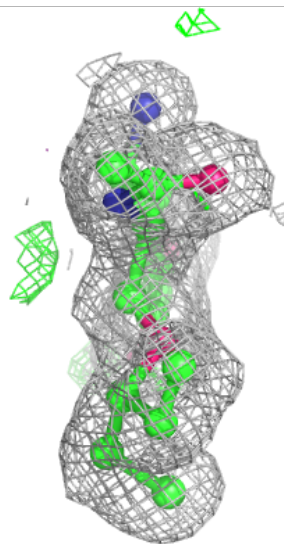
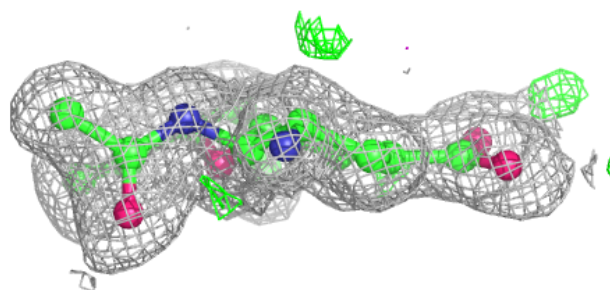
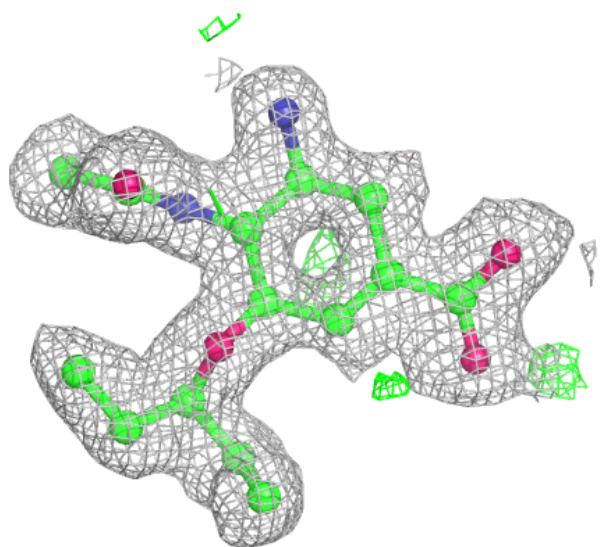
**Electron density around G39 C 503:**

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



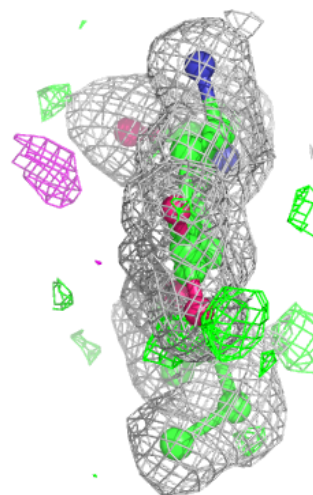
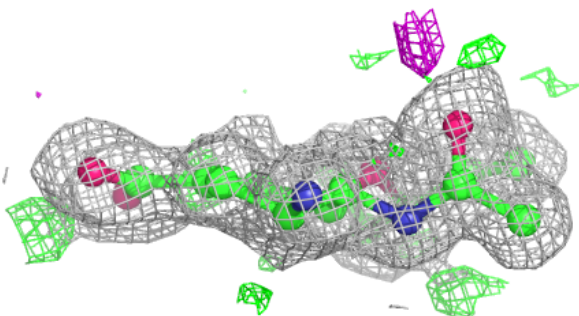
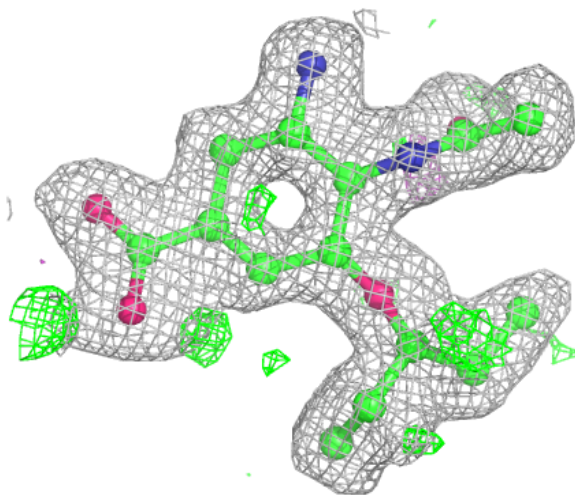
**Electron density around G39 B 503:**

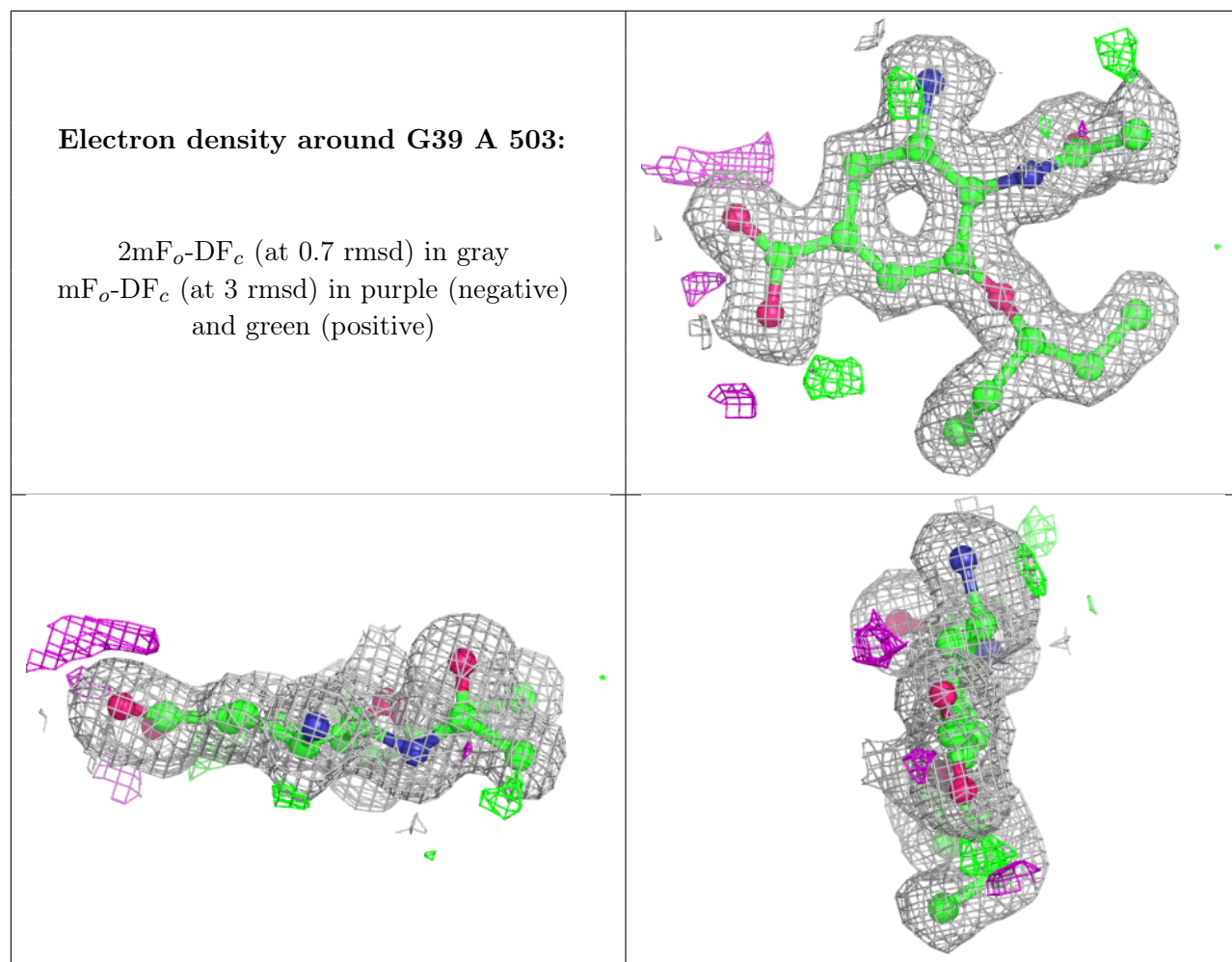
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around G39 D 503:**

$2mF_o-DF_c$  (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.