WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 18, 2024 – 10:33 PM EST

PDB ID	:	4GU3
Title	:	Crystal structure of the T1L reovirus attachment protein sigma1 in complex with the GM2 glycan
Authors	:	Reiss, K.; Stehle, T.
Deposited on	:	2012-08-29
Resolution	:	3.60 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

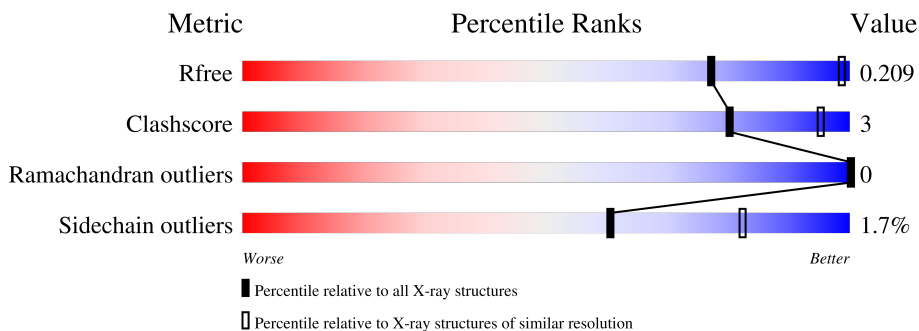|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 3.60 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	Percentile Ranks	Value
Rfree		0.209
Clashscore		3
Ramachandran outliers		0
Sidechain outliers		1.7%

Worse   *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	219	86% 8% 6%
1	B	219	85% 8% 6%
1	C	219	86% 7% 7%
2	D	4	75% 25%
2	E	4	50% 50%
2	F	4	25% 75%

WORLDWIDE PDB
PROTEIN DATA BANK

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 4947 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 Outer capsid protein sigma-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	205	Total	C	N	O	S	0	0	0
			1589	1013	265	304	7			
1	B	205	Total	C	N	O	S	0	0	0
			1587	1013	260	307	7			
1	C	204	Total	C	N	O	S	0	0	0
			1600	1019	268	306	7			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	252	GLY	-	expression tag	UNP P04506
A	253	VAL	-	expression tag	UNP P04506
A	254	LEU	-	expression tag	UNP P04506
A	255	ASN	-	expression tag	UNP P04506
A	256	GLN	-	expression tag	UNP P04506
A	257	GLY	-	expression tag	UNP P04506
A	258	VAL	-	expression tag	UNP P04506
A	259	THR	-	expression tag	UNP P04506
A	260	SER	-	expression tag	UNP P04506
B	252	GLY	-	expression tag	UNP P04506
B	253	VAL	-	expression tag	UNP P04506
B	254	LEU	-	expression tag	UNP P04506
B	255	ASN	-	expression tag	UNP P04506
B	256	GLN	-	expression tag	UNP P04506
B	257	GLY	-	expression tag	UNP P04506
B	258	VAL	-	expression tag	UNP P04506
B	259	THR	-	expression tag	UNP P04506
B	260	SER	-	expression tag	UNP P04506
C	252	GLY	-	expression tag	UNP P04506
C	253	VAL	-	expression tag	UNP P04506
C	254	LEU	-	expression tag	UNP P04506
C	255	ASN	-	expression tag	UNP P04506
C	256	GLN	-	expression tag	UNP P04506

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	257	GLY	-	expression tag	UNP P04506
C	258	VAL	-	expression tag	UNP P04506
C	259	THR	-	expression tag	UNP P04506
C	260	SER	-	expression tag	UNP P04506

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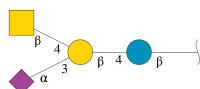

| Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	D	4	Total	C	N	O	0	0	0
			57	31	2	24			
2	E	4	Total	C	N	O	0	0	0
			57	31	2	24			
2	F	4	Total	C	N	O	0	0	0
			57	31	2	24			

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.
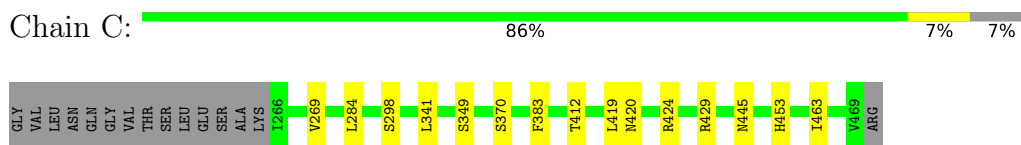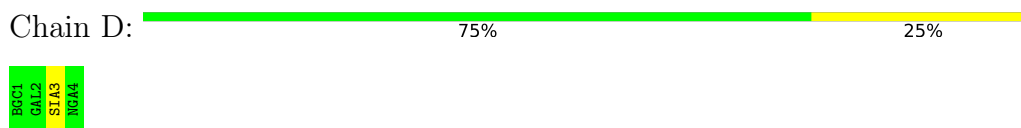
• Molecule 1: Outer capsid protein sigma-1
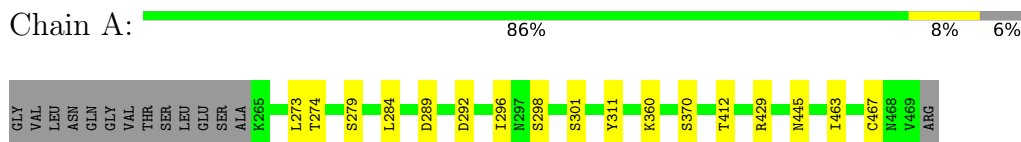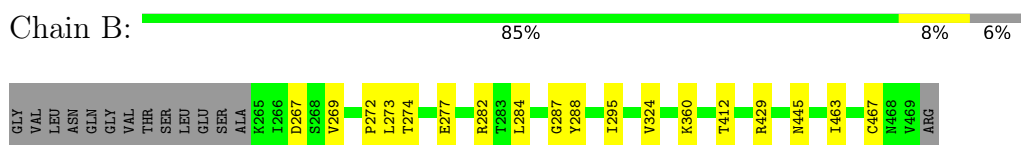
Chain A:

• Molecule 1: Outer capsid protein sigma-1

Chain B:

• Molecule 1: Outer capsid protein sigma-1

Chain C:

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

Chain D:

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

Chain E:

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

Chain F:

25%

75%

BGC1 GAL2 SIA3 NGA4

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	147.46Å 147.46Å 164.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	46.31 – 3.60 46.31 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (46.31-3.60) 98.3 (46.31-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.30 (at 3.01Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.185 , 0.204 0.193 , 0.209	Depositor DCC
R_{free} test set	4121 reflections (10.00%)	wwPDB-VP
Wilson B-factor (Å^2)	73.7	Xtriage
Anisotropy	0.429	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 55.9	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4947	wwPDB-VP
Average B, all atoms (Å^2)	90.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality ⓘ

5.1   Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SIA, BGC, NGA, GAL

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	0.55	0/1630	0.68	0/2230
1	B	0.54	0/1628	0.70	0/2230
1	C	0.48	0/1641	0.65	0/2244
All	All	0.53	0/4899	0.68	0/6704

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts ⓘ

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	1589	0	1508	12	0
1	B	1587	0	1503	11	0
1	C	1600	0	1534	10	0
2	D	57	0	49	0	0
2	E	57	0	49	0	0
2	F	57	0	49	0	0
All	All	4947	0	4692	26	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 3.

All (26) 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:429:ARG:HH11	1:A:445:ASN:HD21	1.22	0.84
1:A:284:LEU:HB3	1:C:269:VAL:HG12	1.69	0.74
1:C:429:ARG:HH11	1:C:445:ASN:HD21	1.40	0.69
1:B:429:ARG:HH11	1:B:445:ASN:HD21	1.44	0.64
1:A:273:LEU:HD13	1:B:284:LEU:HD23	1.86	0.56
1:B:277:GLU:HA	1:B:282:ARG:HA	1.90	0.54
1:C:420:ASN:HB2	1:C:453:HIS:HA	1.93	0.51
1:B:360:LYS:HE3	1:B:467:CYS:SG	2.51	0.51
1:C:419:LEU:CD2	1:C:424:ARG:HG2	2.42	0.50
1:A:273:LEU:HD13	1:B:284:LEU:CD2	2.41	0.50
1:B:269:VAL:HG12	1:C:284:LEU:HB3	1.93	0.50
1:A:429:ARG:NH1	1:A:445:ASN:HD21	2.02	0.49
1:A:360:LYS:HE3	1:A:467:CYS:SG	2.52	0.49
1:A:284:LEU:HB3	1:C:269:VAL:CG1	2.42	0.47
1:A:412:THR:HG22	1:A:463:ILE:HD12	1.97	0.47
1:C:419:LEU:HD23	1:C:424:ARG:HG2	1.97	0.46
1:C:412:THR:HG22	1:C:463:ILE:HD12	2.00	0.44
1:B:273:LEU:HD13	1:C:284:LEU:HG	1.99	0.44
1:A:289:ASP:HB3	1:A:292:ASP:HB2	2.00	0.44
1:B:412:THR:HG22	1:B:463:ILE:HD12	1.99	0.43
1:A:296:ILE:HG12	1:A:301:SER:OG	2.18	0.43
1:B:288:TYR:HE1	1:B:295:ILE:HD11	1.83	0.43
1:A:273:LEU:HB3	1:A:284:LEU:HD11	2.01	0.43
1:C:341:LEU:HD11	1:C:383:PHE:HD2	1.85	0.42
1:A:311:TYR:CD1	1:B:324:VAL:HB	2.56	0.41
1:B:272:PRO:HB3	1:B:287:GLY:O	2.21	0.41

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	203/219 (93%)	195 (96%)	8 (4%)	0	100	100
1	B	203/219 (93%)	197 (97%)	6 (3%)	0	100	100
1	C	202/219 (92%)	198 (98%)	4 (2%)	0	100	100
All	All	608/657 (92%)	590 (97%)	18 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	171/192 (89%)	167 (98%)	4 (2%)	50	76
1	B	172/192 (90%)	170 (99%)	2 (1%)	71	87
1	C	175/192 (91%)	172 (98%)	3 (2%)	60	82
All	All	518/576 (90%)	509 (98%)	9 (2%)	60	82

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	274	THR
1	A	279	SER
1	A	298	SER
1	A	370	SER
1	B	267	ASP
1	B	274	THR
1	C	298	SER
1	C	349	SER
1	C	370	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	297	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	445	ASN
1	B	445	ASN
1	C	445	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 ⓘ

12 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	D	1	2	12,12,12	0.32	0	17,17,17	0.51	0
2	GAL	D	2	2	11,11,12	0.31	0	15,15,17	0.53	0
2	SIA	D	3	2	20,20,21	1.32	1 (5%)	24,28,31	2.17	8 (33%)
2	NGA	D	4	2	14,14,15	0.67	0	17,19,21	0.65	0
2	BGC	E	1	2	12,12,12	0.34	0	17,17,17	0.73	0
2	GAL	E	2	2	11,11,12	0.41	0	15,15,17	0.94	1 (6%)
2	SIA	E	3	2	20,20,21	0.89	1 (5%)	24,28,31	1.72	5 (20%)
2	NGA	E	4	2	14,14,15	0.48	0	17,19,21	0.63	0
2	BGC	F	1	2	12,12,12	0.29	0	17,17,17	0.90	1 (5%)
2	GAL	F	2	2	11,11,12	0.31	0	15,15,17	0.83	1 (6%)
2	SIA	F	3	2	20,20,21	1.03	1 (5%)	24,28,31	1.84	8 (33%)
2	NGA	F	4	2	14,14,15	0.58	0	17,19,21	0.68	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	BGC	D	1	2	-	0/2/22/22	0/1/1/1
2	GAL	D	2	2	-	1/2/19/22	0/1/1/1
2	SIA	D	3	2	-	1/18/34/38	0/1/1/1
2	NGA	D	4	2	-	2/6/23/26	0/1/1/1
2	BGC	E	1	2	-	0/2/22/22	0/1/1/1
2	GAL	E	2	2	-	0/2/19/22	0/1/1/1
2	SIA	E	3	2	-	2/18/34/38	0/1/1/1
2	NGA	E	4	2	-	0/6/23/26	0/1/1/1
2	BGC	F	1	2	-	0/2/22/22	0/1/1/1
2	GAL	F	2	2	-	1/2/19/22	0/1/1/1
2	SIA	F	3	2	-	3/18/34/38	0/1/1/1
2	NGA	F	4	2	-	0/6/23/26	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	3	SIA	C2-C1	4.54	1.56	1.52
2	F	3	SIA	C2-C1	3.35	1.55	1.52
2	E	3	SIA	C2-C1	2.20	1.54	1.52

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	3	SIA	C3-C4-C5	5.55	118.17	111.46
2	E	3	SIA	C3-C4-C5	4.19	116.53	111.46
2	D	3	SIA	C6-O6-C2	4.09	120.08	111.34
2	E	3	SIA	C6-O6-C2	4.05	120.00	111.34
2	F	3	SIA	O1B-C1-C2	3.55	123.15	113.03
2	D	3	SIA	O6-C2-C1	3.50	114.57	107.70
2	F	3	SIA	O6-C2-C1	2.90	113.38	107.70
2	F	3	SIA	C4-C3-C2	2.89	114.98	109.81
2	D	3	SIA	O8-C8-C7	2.82	115.95	109.10
2	D	3	SIA	C4-C5-N5	-2.77	104.89	110.38
2	E	3	SIA	C6-C5-N5	-2.70	106.42	110.91
2	D	3	SIA	C11-C10-N5	-2.62	111.67	116.10
2	F	3	SIA	C3-C4-C5	2.60	114.60	111.46
2	F	3	SIA	C6-O6-C2	2.57	116.84	111.34
2	E	2	GAL	C1-C2-C3	2.54	112.78	109.67
2	F	3	SIA	C9-C8-C7	-2.52	106.96	112.41

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	F	2	GAL	C1-C2-C3	2.46	112.69	109.67
2	D	3	SIA	O1B-C1-O1A	-2.45	118.53	124.09
2	E	3	SIA	C4-C3-C2	2.39	114.10	109.81
2	F	3	SIA	O1A-C1-C2	-2.33	117.06	122.57
2	D	3	SIA	C4-C3-C2	2.20	113.75	109.81
2	E	3	SIA	O6-C2-C1	2.14	111.90	107.70
2	F	3	SIA	C6-C5-N5	-2.06	107.49	110.91
2	F	1	BGC	O4-C4-C3	-2.00	105.72	110.35

There are no chirality outliers.

All (10) torsion outliers are listed below:

| Mol	Chain	Res	Type	Atoms
2	D	4	NGA	C8-C7-N2-C2
2	D	4	NGA	O7-C7-N2-C2
2	F	3	SIA	O6-C6-C7-O7
2	E	3	SIA	O8-C8-C9-O9
2	F	3	SIA	C6-C7-C8-O8
2	F	3	SIA	C6-C7-C8-C9
2	F	2	GAL	O5-C5-C6-O6
2	D	2	GAL	O5-C5-C6-O6
2	E	3	SIA	C7-C8-C9-O9
2	D	3	SIA	O7-C7-C8-O8

There are no ring outliers.

No monomer is involved in short contacts.

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

Oligosaccharide Chain D

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain E

Bond lengths

Bond angles

Torsions

Rings

Oligosaccharide Chain F

Bond lengths

Bond angles

Torsions

Rings
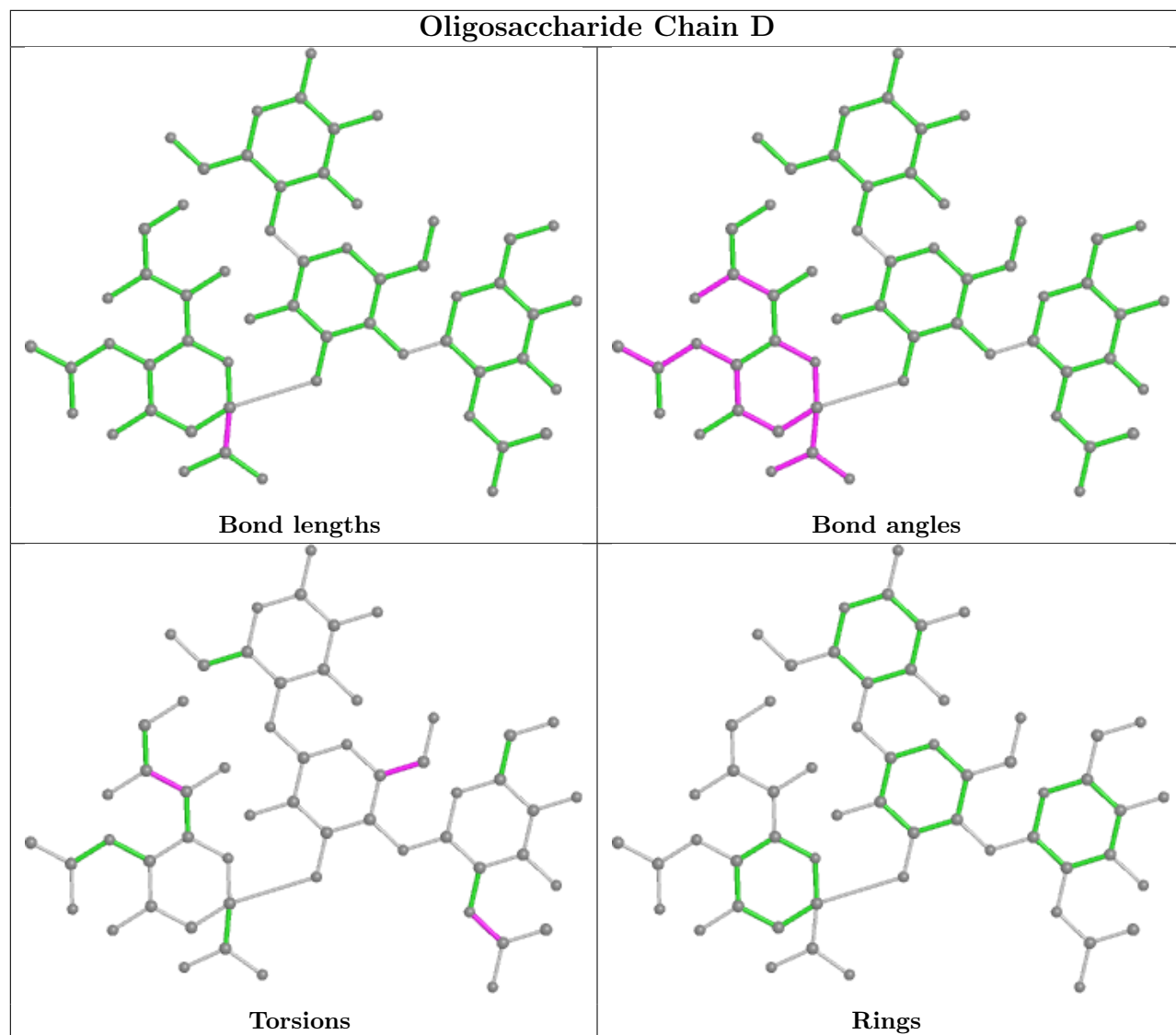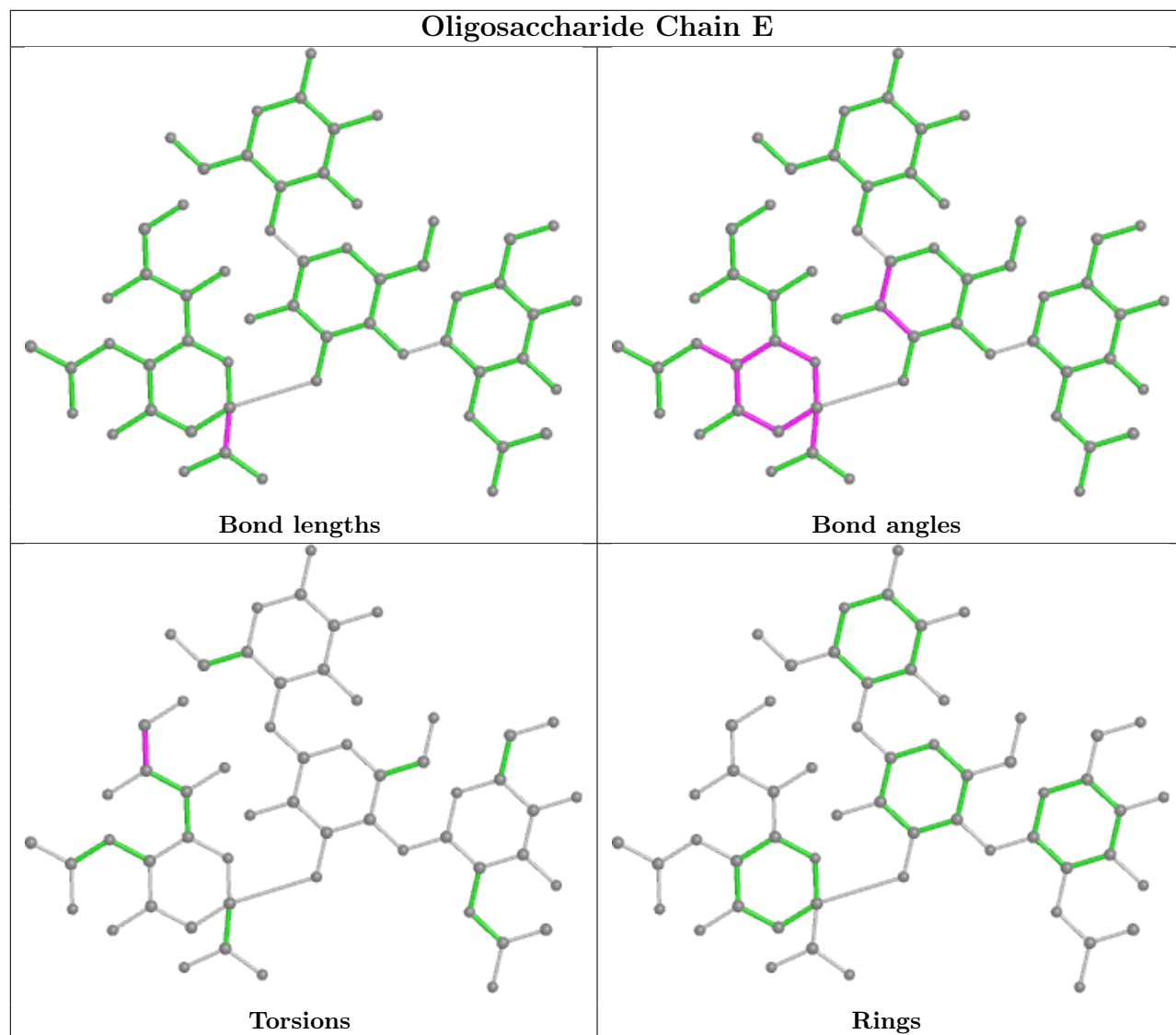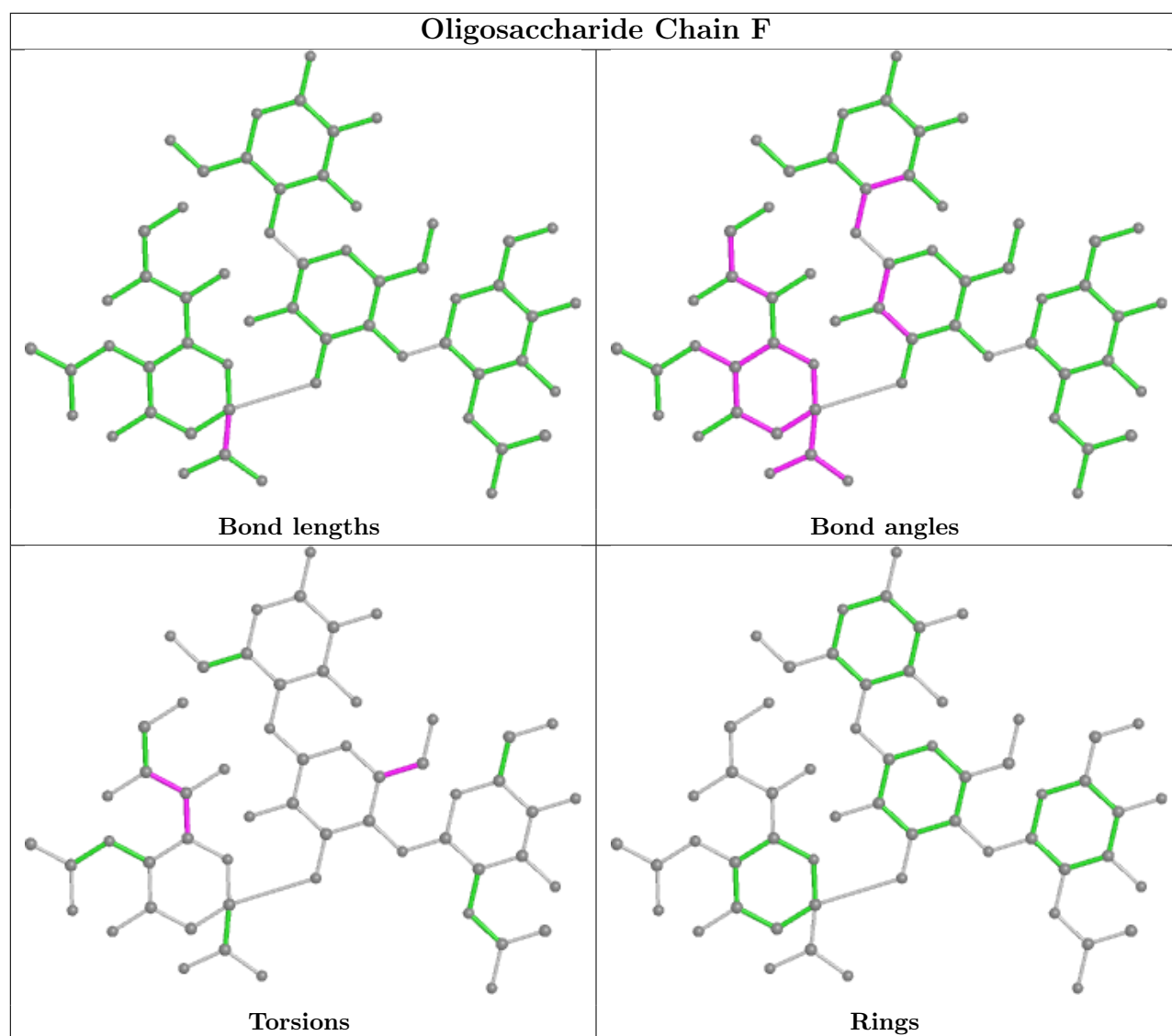
5.6 Ligand geometry ⓘ

There are no ligands in this entry.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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 D:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)
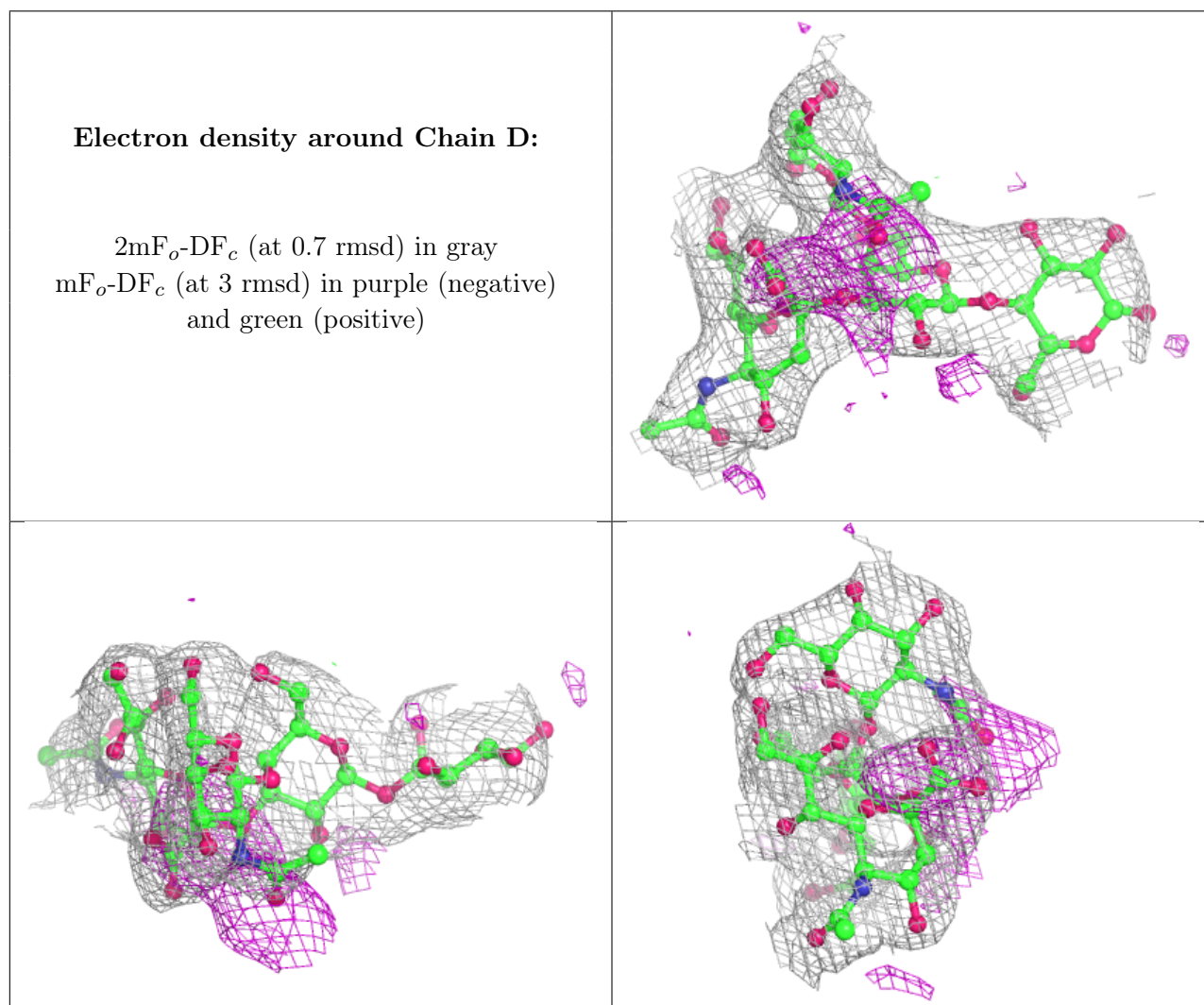
Electron density around Chain E:

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

Electron density around Chain F:

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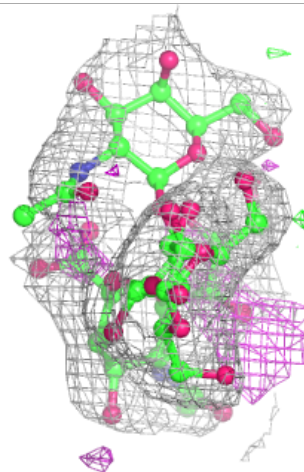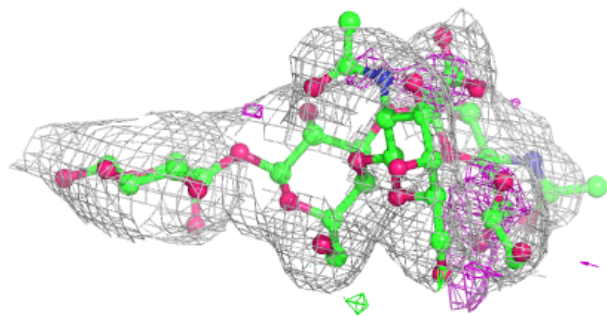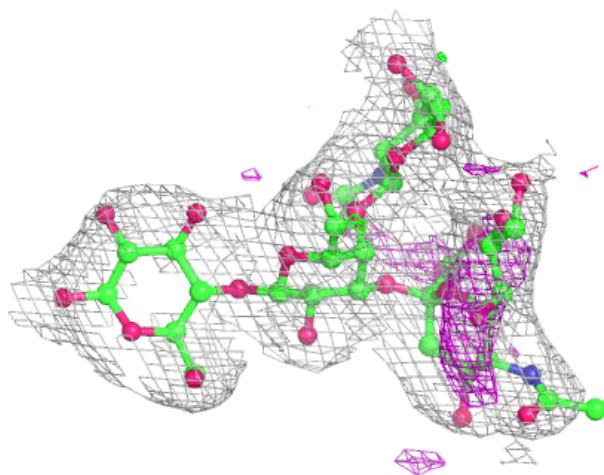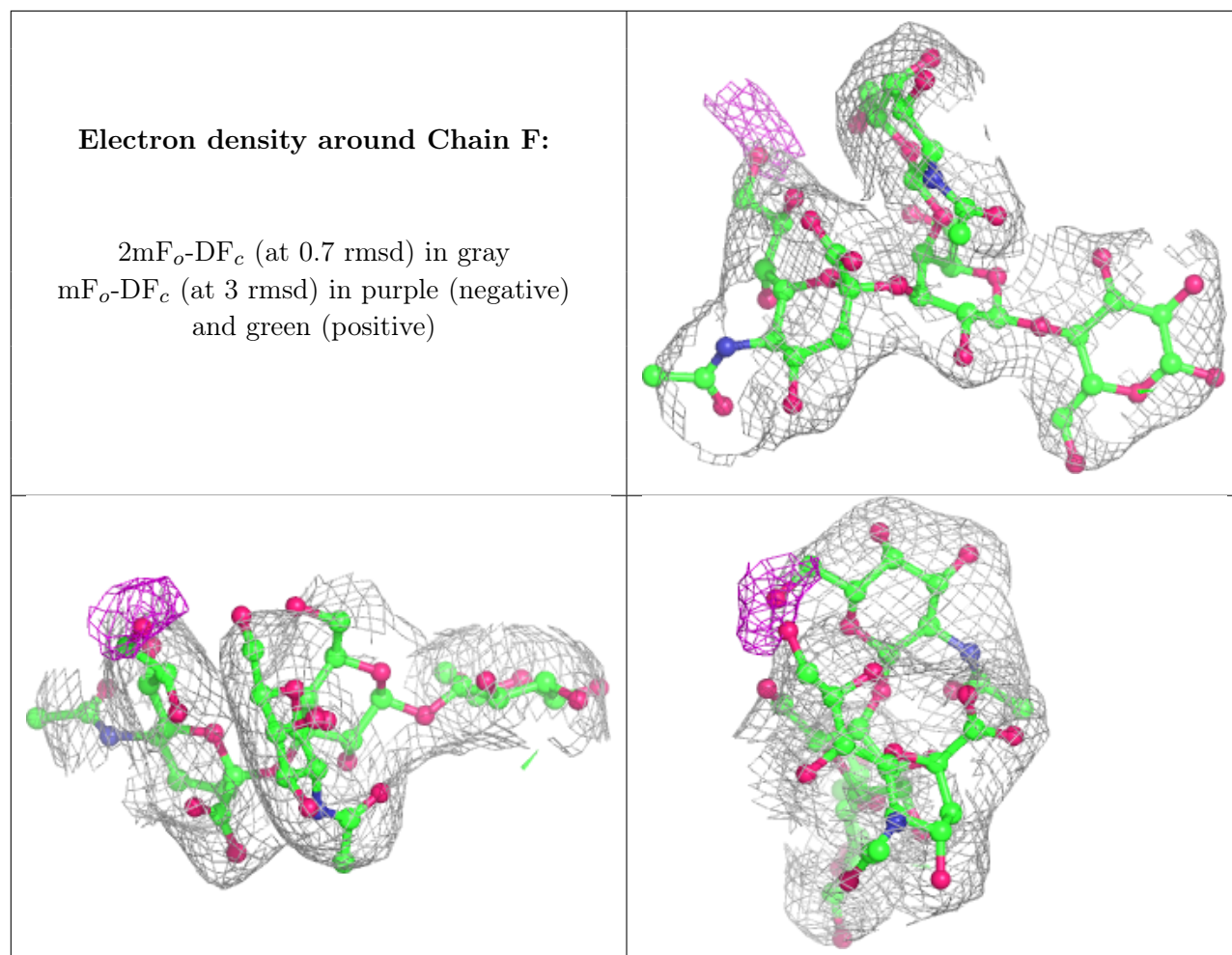
6.4　Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5　Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.