



wwPDB X-ray Structure Validation Summary Report ⓘ

Dec 14, 2021 – 12:10 pm GMT

PDB ID : 7PS4
Title : Crystal structure of the receptor binding domain of SARS-CoV-2 beta variant spike glycoprotein in complex with Beta-38
Authors : Zhou, D.; Ren, J.; Stuart, D.I.
Deposited on : 2021-09-22
Resolution : 1.94 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 : 1.8.4, CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.24
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0267
CCP4 : 7.1.010 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.24

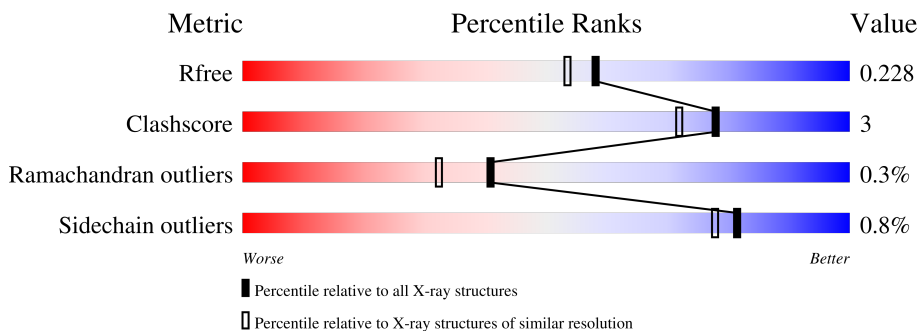
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.94 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)

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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	210	85% 8% 7%
1	E	210	84% 8% 8%
2	B	230	93% 7% .
2	H	230	91% 6% .
3	C	216	90% 8% .
3	L	216	89% 9% .
4	D	2	100%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	F	2	 50% 50%

2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 10153 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 Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	194	1544	991	257	288	8	0	1	0
1	A	195	1549	994	258	289	8	0	1	0

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	319	MET	-	initiating methionine	UNP P0DTC2
E	320	GLY	-	expression tag	UNP P0DTC2
E	321	CYS	-	expression tag	UNP P0DTC2
E	322	VAL	-	expression tag	UNP P0DTC2
E	323	ALA	-	expression tag	UNP P0DTC2
E	324	GLU	-	expression tag	UNP P0DTC2
E	325	THR	-	expression tag	UNP P0DTC2
E	326	GLY	-	expression tag	UNP P0DTC2
E	327	HIS	-	expression tag	UNP P0DTC2
E	328	HIS	-	expression tag	UNP P0DTC2
E	329	HIS	-	expression tag	UNP P0DTC2
E	330	HIS	-	expression tag	UNP P0DTC2
E	331	HIS	-	expression tag	UNP P0DTC2
E	332	HIS	-	expression tag	UNP P0DTC2
E	417	ASN	LYS	variant	UNP P0DTC2
E	484	LYS	GLU	variant	UNP P0DTC2
E	501	TYR	ASN	variant	UNP P0DTC2
E	527	LYS	-	expression tag	UNP P0DTC2
E	528	LYS	-	expression tag	UNP P0DTC2
A	319	MET	-	initiating methionine	UNP P0DTC2
A	320	GLY	-	expression tag	UNP P0DTC2
A	321	CYS	-	expression tag	UNP P0DTC2
A	322	VAL	-	expression tag	UNP P0DTC2
A	323	ALA	-	expression tag	UNP P0DTC2
A	324	GLU	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	325	THR	-	expression tag	UNP P0DTC2
A	326	GLY	-	expression tag	UNP P0DTC2
A	327	HIS	-	expression tag	UNP P0DTC2
A	328	HIS	-	expression tag	UNP P0DTC2
A	329	HIS	-	expression tag	UNP P0DTC2
A	330	HIS	-	expression tag	UNP P0DTC2
A	331	HIS	-	expression tag	UNP P0DTC2
A	332	HIS	-	expression tag	UNP P0DTC2
A	417	ASN	LYS	variant	UNP P0DTC2
A	484	LYS	GLU	variant	UNP P0DTC2
A	501	TYR	ASN	variant	UNP P0DTC2
A	527	LYS	-	expression tag	UNP P0DTC2
A	528	LYS	-	expression tag	UNP P0DTC2

- Molecule 2 is a protein called Beta-38 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	223	Total 1674	C 1063	N 275	O 328	S 8	0	0	0
2	B	228	Total 1708	C 1081	N 280	O 338	S 9	0	1	0

- Molecule 3 is a protein called Beta-38 Fab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	213	Total 1573	C 978	N 265	O 326	S 4	0	1	0
3	C	213	Total 1577	C 980	N 265	O 328	S 4	0	1	0

- Molecule 4 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			
4	D	2	Total 28	C 16	N 2	O 10	0	0	0
4	F	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	H	1	Total Cl 1 1	0	0
5	B	1	Total Cl 1 1	0	0


- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	E	81	Total O 81 81	0	0
6	H	161	Total O 161 161	0	0
6	L	79	Total O 79 79	0	0
6	A	90	Total O 90 90	0	0
6	B	35	Total O 35 35	0	0
6	C	24	Total O 24 24	0	0

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. 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: Spike protein S1

Chain E:  84% 8% 8%



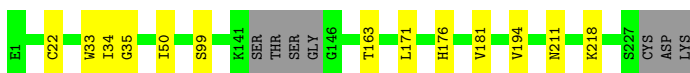
- Molecule 1: Spike protein S1

Chain A:  85% 8% 7%



- Molecule 2: Beta-38 Fab heavy chain

Chain H:  91% 6%




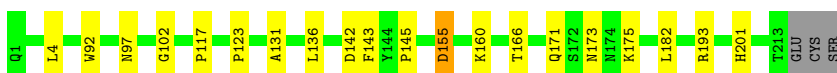
- Molecule 2: Beta-38 Fab heavy chain

Chain B:  93% 7%



- Molecule 3: Beta-38 Fab light chain

Chain L:  89% 9%



- Molecule 3: Beta-38 Fab light chain

Chain C:  90% 8%



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

Chain D:  100%



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

Chain F:  50% 50%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	55.07Å 118.49Å 109.27Å 90.00° 93.47° 90.00°	Depositor
Resolution (Å)	54.53 – 1.94 59.24 – 1.94	Depositor EDS
% Data completeness (in resolution range)	97.5 (54.53-1.94) 97.5 (59.24-1.94)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.02 (at 1.94Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.203 , 0.231 0.201 , 0.228	Depositor DCC
R_{free} test set	5035 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	43.8	Xtriage
Anisotropy	0.217	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	(Not available) , (Not available)	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	10153	wwPDB-VP
Average B, all atoms (Å ²)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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: CL, NAG

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.33	0/1596	0.56	0/2173
1	E	0.33	0/1591	0.56	0/2166
2	B	0.28	0/1756	0.53	0/2392
2	H	0.34	0/1718	0.56	0/2339
3	C	0.28	0/1618	0.51	0/2214
3	L	0.32	0/1614	0.53	0/2209
All	All	0.31	0/9893	0.54	0/13493

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 [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	1549	0	1463	9	0
1	E	1544	0	1461	9	0
2	B	1708	0	1655	9	0
2	H	1674	0	1625	8	0
3	C	1577	0	1522	9	0
3	L	1573	0	1521	14	0
4	D	28	0	25	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	F	28	0	25	1	0
5	B	1	0	0	0	0
5	H	1	0	0	0	0
6	A	90	0	0	0	0
6	B	35	0	0	0	0
6	C	24	0	0	0	0
6	E	81	0	0	0	0
6	H	161	0	0	1	0
6	L	79	0	0	2	0
All	All	10153	0	9297	54	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.

The worst 5 of 54 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)
3:L:131:ALA:O	6:L:301:HOH:O	2.09	0.71
3:L:97:ASN:OD1	6:L:302:HOH:O	2.12	0.68
1:E:388:ASN:HA	1:E:526:GLY:HA3	1.79	0.64
1:A:388:ASN:HA	1:A:526:GLY:HA3	1.79	0.64
3:C:136:LEU:HD12	3:C:182:LEU:HD23	1.82	0.61

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	194/210 (92%)	188 (97%)	6 (3%)	0	100 100
1	E	193/210 (92%)	183 (95%)	8 (4%)	2 (1%)	15 6

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	227/230 (99%)	221 (97%)	6 (3%)	0	100	100
2	H	219/230 (95%)	216 (99%)	3 (1%)	0	100	100
3	C	212/216 (98%)	206 (97%)	5 (2%)	1 (0%)	29	17
3	L	212/216 (98%)	204 (96%)	7 (3%)	1 (0%)	29	17
All	All	1257/1312 (96%)	1218 (97%)	35 (3%)	4 (0%)	41	32

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	360	ASN
3	L	155	ASP
3	C	155	ASP
1	E	362	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	168/180 (93%)	166 (99%)	2 (1%)	71	64
1	E	168/180 (93%)	167 (99%)	1 (1%)	86	85
2	B	190/193 (98%)	187 (98%)	3 (2%)	62	52
2	H	185/193 (96%)	184 (100%)	1 (0%)	88	88
3	C	180/182 (99%)	178 (99%)	2 (1%)	73	67
3	L	179/182 (98%)	179 (100%)	0	100	100
All	All	1070/1110 (96%)	1061 (99%)	9 (1%)	81	78

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	35	ASN
3	C	213	THR
1	A	518	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	109[A]	ASP
2	B	109[B]	ASP

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

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

4 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
4	NAG	D	1	4,1	14,14,15	0.45	0	17,19,21	0.59	0
4	NAG	D	2	4	14,14,15	0.21	0	17,19,21	0.52	0
4	NAG	F	1	4,1	14,14,15	0.30	0	17,19,21	0.52	0
4	NAG	F	2	4	14,14,15	0.53	0	17,19,21	0.66	1 (5%)

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	NAG	D	1	4,1	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	D	2	4	-	2/6/23/26	0/1/1/1
4	NAG	F	1	4,1	-	2/6/23/26	0/1/1/1
4	NAG	F	2	4	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	F	2	NAG	C1-O5-C5	2.32	115.33	112.19

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

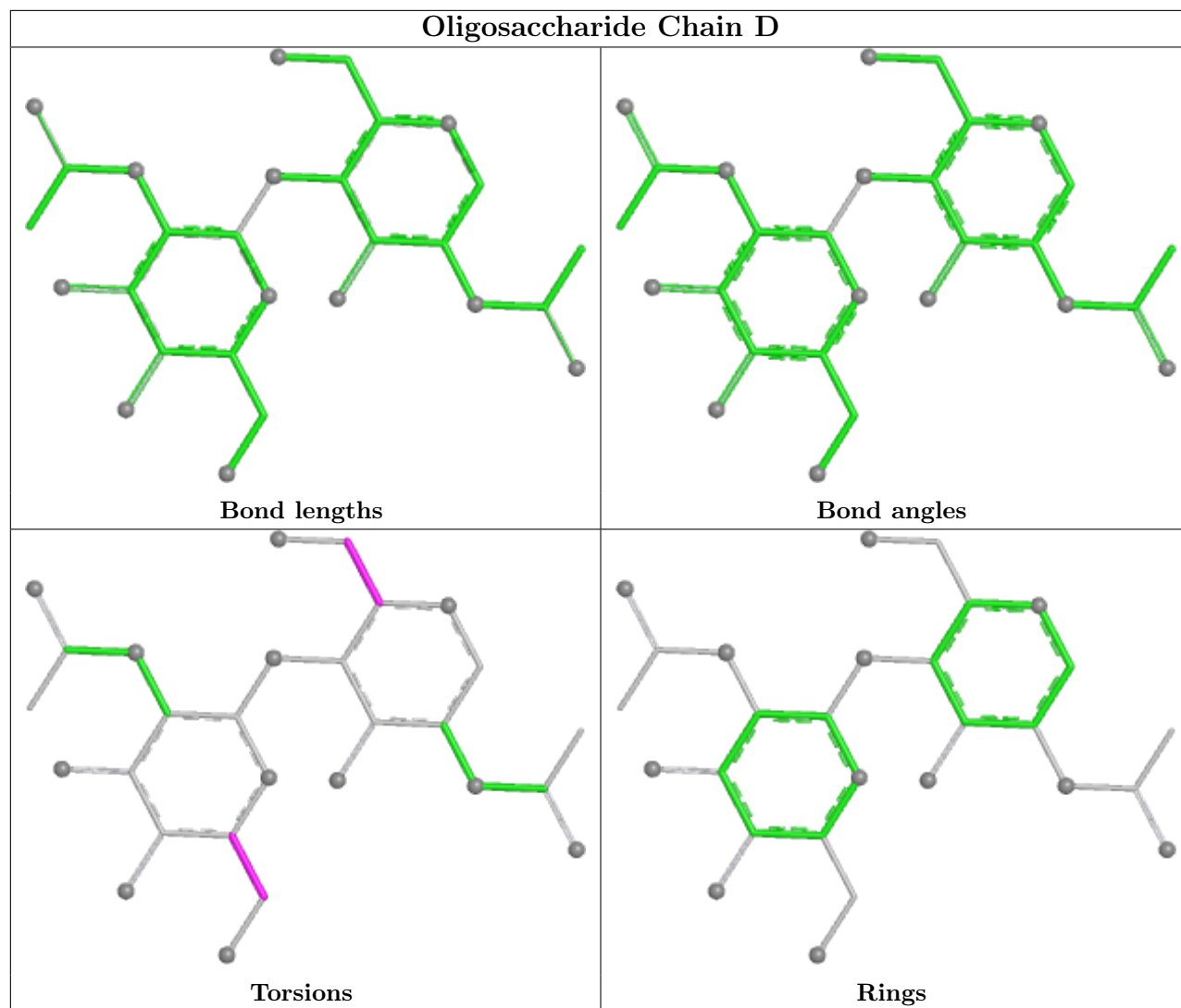
Mol	Chain	Res	Type	Atoms
4	D	2	NAG	O5-C5-C6-O6
4	F	1	NAG	O5-C5-C6-O6
4	F	1	NAG	C4-C5-C6-O6
4	D	2	NAG	C4-C5-C6-O6
4	D	1	NAG	O5-C5-C6-O6

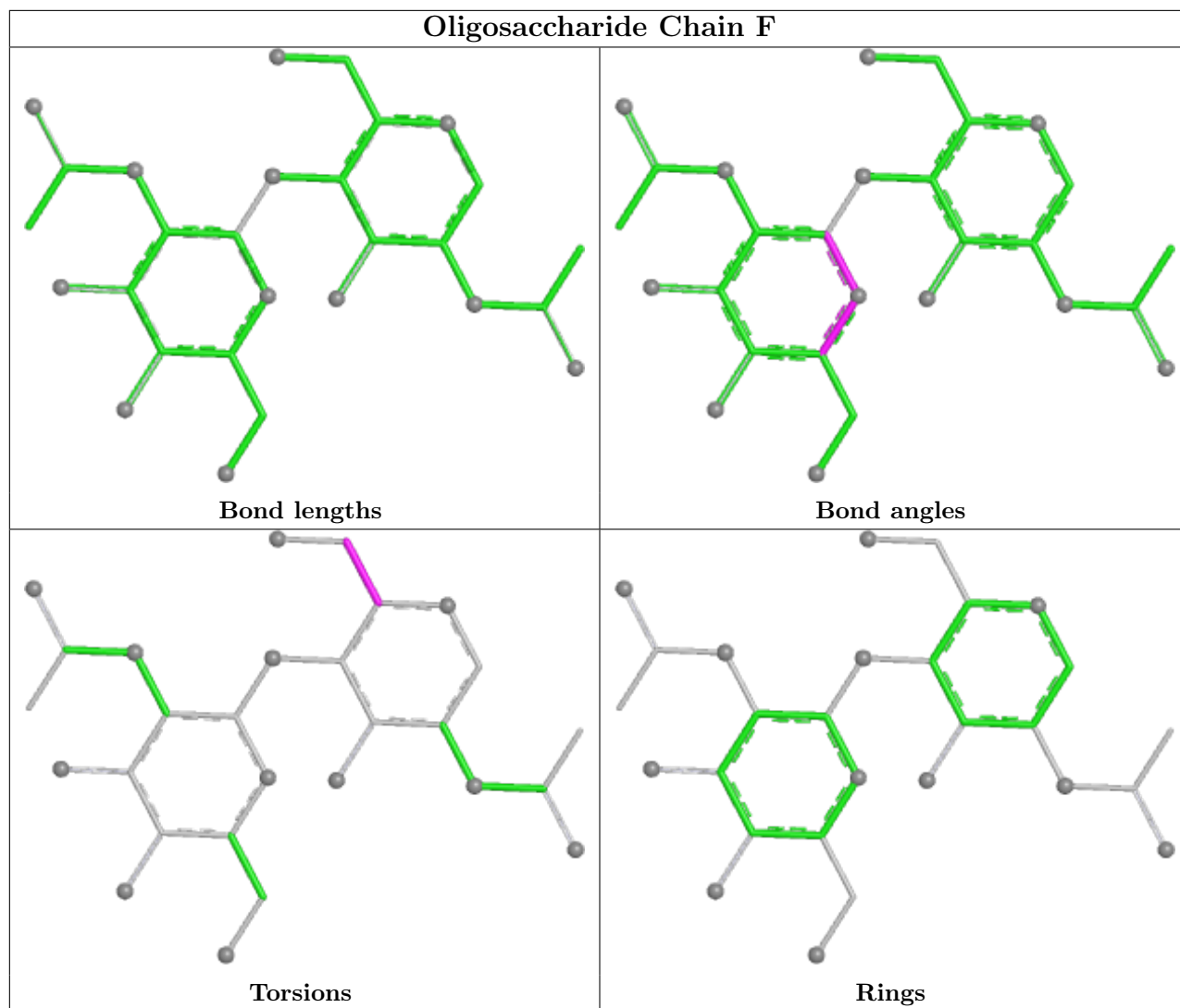
There are no ring outliers.

2 monomers are involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	2	NAG	1	0
4	F	1	NAG	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](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

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 [i](#)

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

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

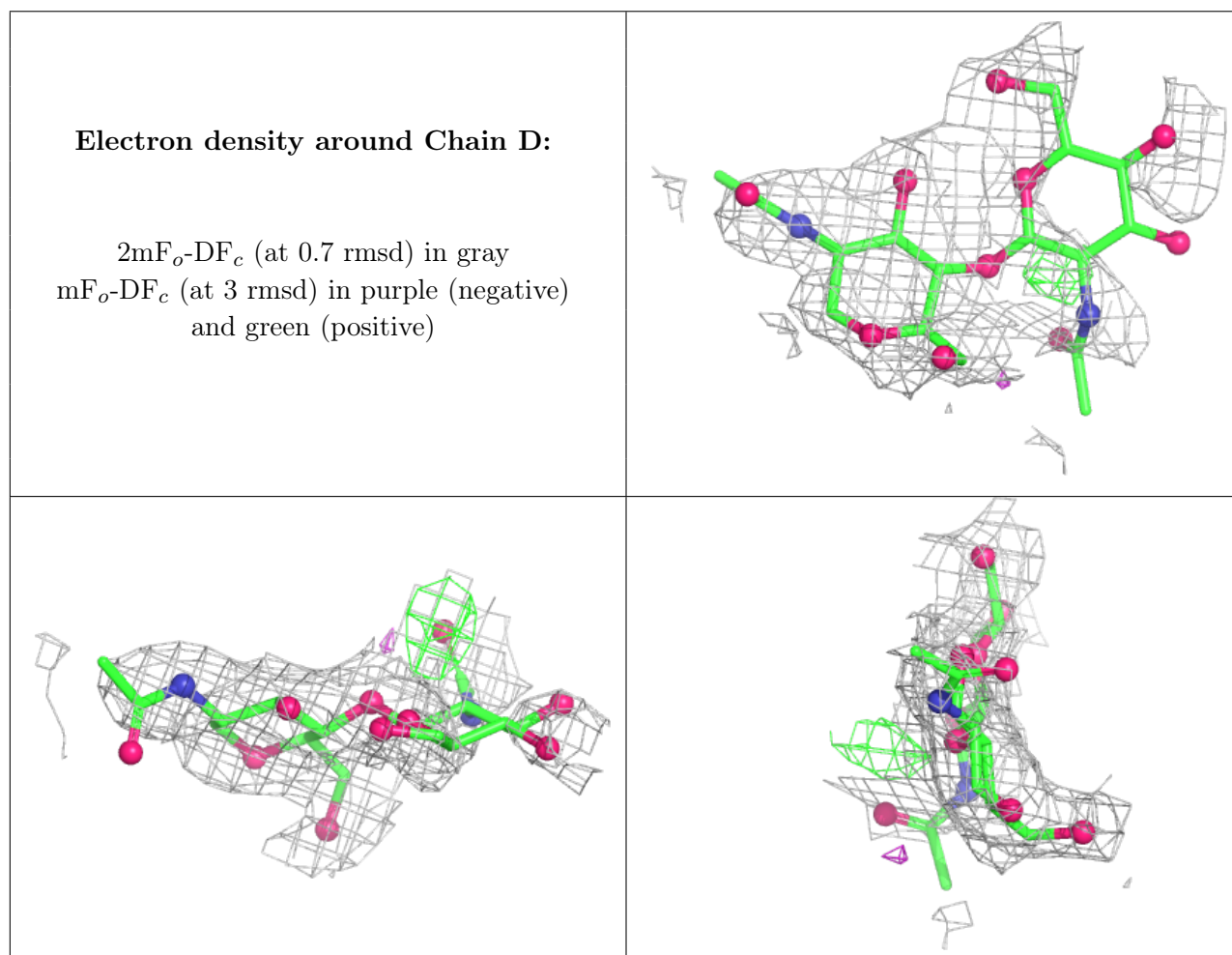
6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

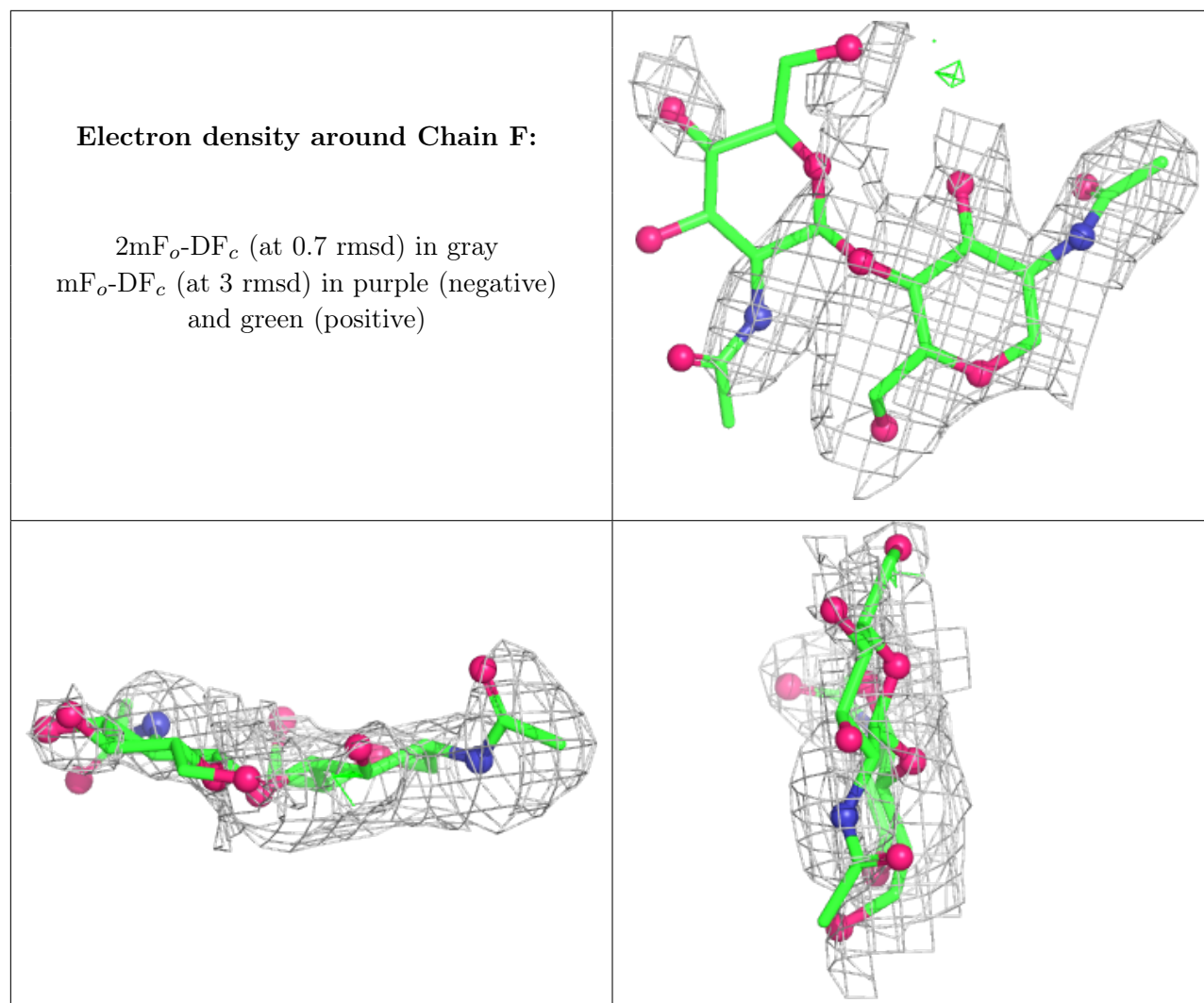
Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates [i](#)

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.





6.4 Ligands [i](#)

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

6.5 Other polymers [i](#)

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