



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 25, 2023 – 01:39 AM EDT

PDB ID : 5WB0  
Title : Crystal structure of human metapneumovirus fusion glycoprotein stabilized in the prefusion state  
Authors : Battles, M.B.; McLellan, J.S.  
Deposited on : 2017-06-27  
Resolution : 2.60 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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35.1  
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) : 2.35.1

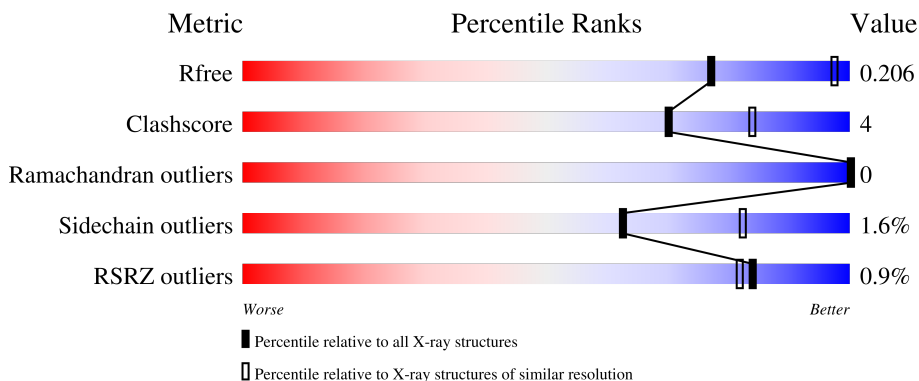
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	F	542	 75% 6% 18%
2	A	2	 100%
3	B	4	 50% 50%

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
5	SO4	F	610	-	-	-	X

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3602 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 Fusion glycoprotein F0.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	F	442	3355	2104	575	655	21	0	0	0

There are 56 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	100	ARG	GLN	engineered mutation	UNP Q1A2Z0
F	101	ARG	SER	engineered mutation	UNP Q1A2Z0
F	185	PRO	ALA	engineered mutation	UNP Q1A2Z0
F	490	SER	-	expression tag	UNP Q1A2Z0
F	491	GLY	-	expression tag	UNP Q1A2Z0
F	492	ARG	-	expression tag	UNP Q1A2Z0
F	493	GLU	-	expression tag	UNP Q1A2Z0
F	494	ASN	-	expression tag	UNP Q1A2Z0
F	495	LEU	-	expression tag	UNP Q1A2Z0
F	496	TYR	-	expression tag	UNP Q1A2Z0
F	497	PHE	-	expression tag	UNP Q1A2Z0
F	498	GLN	-	expression tag	UNP Q1A2Z0
F	499	GLY	-	expression tag	UNP Q1A2Z0
F	500	GLY	-	expression tag	UNP Q1A2Z0
F	501	GLY	-	expression tag	UNP Q1A2Z0
F	502	GLY	-	expression tag	UNP Q1A2Z0
F	503	SER	-	expression tag	UNP Q1A2Z0
F	504	GLY	-	expression tag	UNP Q1A2Z0
F	505	TYR	-	expression tag	UNP Q1A2Z0
F	506	ILE	-	expression tag	UNP Q1A2Z0
F	507	PRO	-	expression tag	UNP Q1A2Z0
F	508	GLU	-	expression tag	UNP Q1A2Z0
F	509	ALA	-	expression tag	UNP Q1A2Z0
F	510	PRO	-	expression tag	UNP Q1A2Z0
F	511	ARG	-	expression tag	UNP Q1A2Z0
F	512	ASP	-	expression tag	UNP Q1A2Z0
F	513	GLY	-	expression tag	UNP Q1A2Z0

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	514	GLN	-	expression tag	UNP Q1A2Z0
F	515	ALA	-	expression tag	UNP Q1A2Z0
F	516	TYR	-	expression tag	UNP Q1A2Z0
F	517	VAL	-	expression tag	UNP Q1A2Z0
F	518	ARG	-	expression tag	UNP Q1A2Z0
F	519	LYS	-	expression tag	UNP Q1A2Z0
F	520	ASP	-	expression tag	UNP Q1A2Z0
F	521	GLY	-	expression tag	UNP Q1A2Z0
F	522	GLU	-	expression tag	UNP Q1A2Z0
F	523	TRP	-	expression tag	UNP Q1A2Z0
F	524	VAL	-	expression tag	UNP Q1A2Z0
F	525	LEU	-	expression tag	UNP Q1A2Z0
F	526	LEU	-	expression tag	UNP Q1A2Z0
F	527	SER	-	expression tag	UNP Q1A2Z0
F	528	THR	-	expression tag	UNP Q1A2Z0
F	529	PHE	-	expression tag	UNP Q1A2Z0
F	530	LEU	-	expression tag	UNP Q1A2Z0
F	531	GLY	-	expression tag	UNP Q1A2Z0
F	532	GLY	-	expression tag	UNP Q1A2Z0
F	533	ILE	-	expression tag	UNP Q1A2Z0
F	534	GLU	-	expression tag	UNP Q1A2Z0
F	535	GLY	-	expression tag	UNP Q1A2Z0
F	536	ARG	-	expression tag	UNP Q1A2Z0
F	537	HIS	-	expression tag	UNP Q1A2Z0
F	538	HIS	-	expression tag	UNP Q1A2Z0
F	539	HIS	-	expression tag	UNP Q1A2Z0
F	540	HIS	-	expression tag	UNP Q1A2Z0
F	541	HIS	-	expression tag	UNP Q1A2Z0
F	542	HIS	-	expression tag	UNP Q1A2Z0

- Molecule 2 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			
2	A	2	28	16	2	10	0	0	0

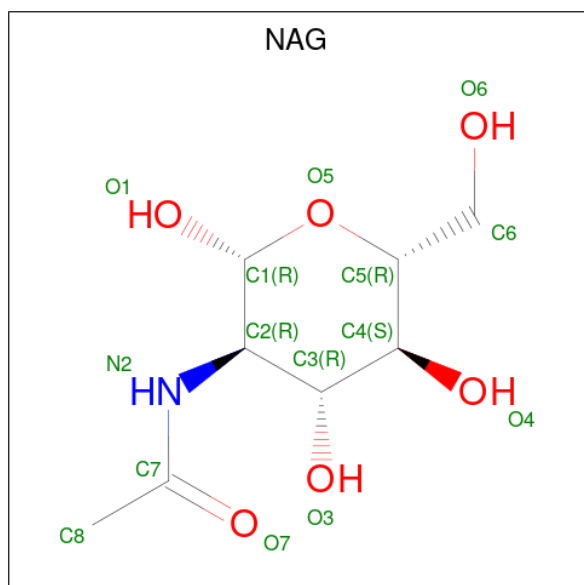
- Molecule 3 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranos

e-(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			
3	B	4	50	28	2	20	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	F	1	14	8	1	5	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula:  $O_4S$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		
5	F	1	Total	O	S	0	0
			5	4	1		

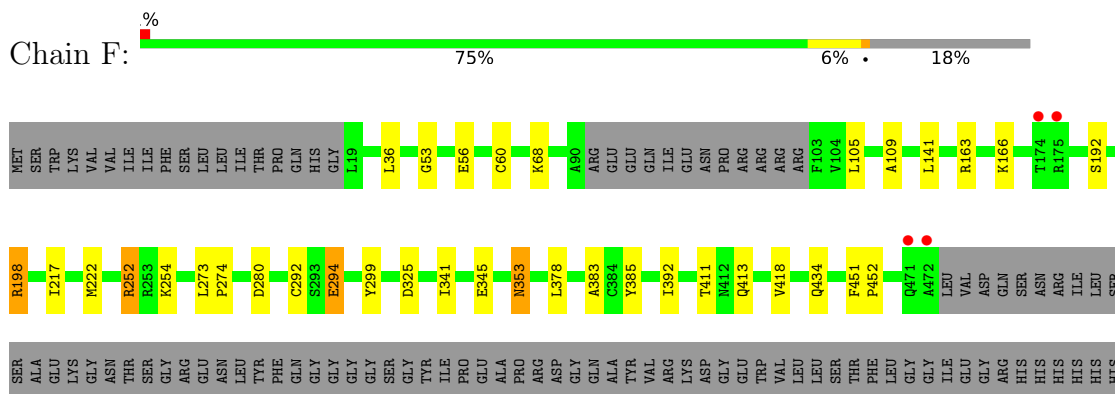
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	F	135	Total	O	0	0
			135	135		

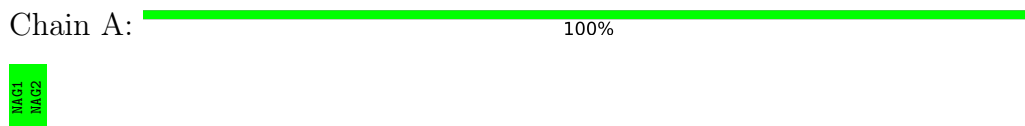
### 3 Residue-property plots [i](#)

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: Fusion glycoprotein F0



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



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





## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	177.70Å 177.70Å 177.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	51.30 – 2.60 51.30 – 2.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (51.30-2.60) 100.0 (51.30-2.60)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.61Å)	Xtrriage
Refinement program	PHENIX 1.11.1_2575	Depositor
R, $R_{free}$	0.170 , 0.207 0.170 , 0.206	Depositor DCC
$R_{free}$ test set	1455 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.4	Xtrriage
Anisotropy	0.000	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 38.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.033 for -l,-k,-h	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	3602	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.60% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: NAG, MAN, BMA, SO4

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	F	0.26	0/3403	0.46	0/4613

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	F	3355	0	3369	24	0
2	A	28	0	25	0	0
3	B	50	0	43	4	0
4	F	14	0	13	1	0
5	F	20	0	0	0	0
6	F	135	0	0	2	0
All	All	3602	0	3450	25	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 4.

The worst 5 of 25 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:F:411:THR:HG22	1:F:413:GLN:H	1.35	0.89
1:F:105:LEU:HD22	1:F:341:ILE:HD13	1.63	0.81
6:F:701:HOH:O	3:B:3:BMA:O2	2.08	0.70
1:F:325:ASP:OD2	6:F:702:HOH:O	2.15	0.65
1:F:392:ILE:HG22	1:F:418:VAL:HG22	1.79	0.63

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	F	438/542 (81%)	427 (98%)	11 (2%)	0	100 100

There are no Ramachandran outliers to report.

### 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	F	373/457 (82%)	367 (98%)	6 (2%)	62 82

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	294	GLU
1	F	345	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	F	353	ASN
1	F	198	ARG
1	F	192	SER

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](#)

6 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	A	1	1,2	14,14,15	0.29	0	17,19,21	0.45	0
2	NAG	A	2	2	14,14,15	0.36	0	17,19,21	0.43	0
3	NAG	B	1	1,3	14,14,15	0.32	0	17,19,21	0.61	0
3	NAG	B	2	3	14,14,15	0.33	0	17,19,21	0.62	0
3	BMA	B	3	3	11,11,12	0.19	0	15,15,17	0.89	0
3	MAN	B	4	3	11,11,12	0.22	0	15,15,17	0.63	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. '2' means no outliers of that kind were identified.

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

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

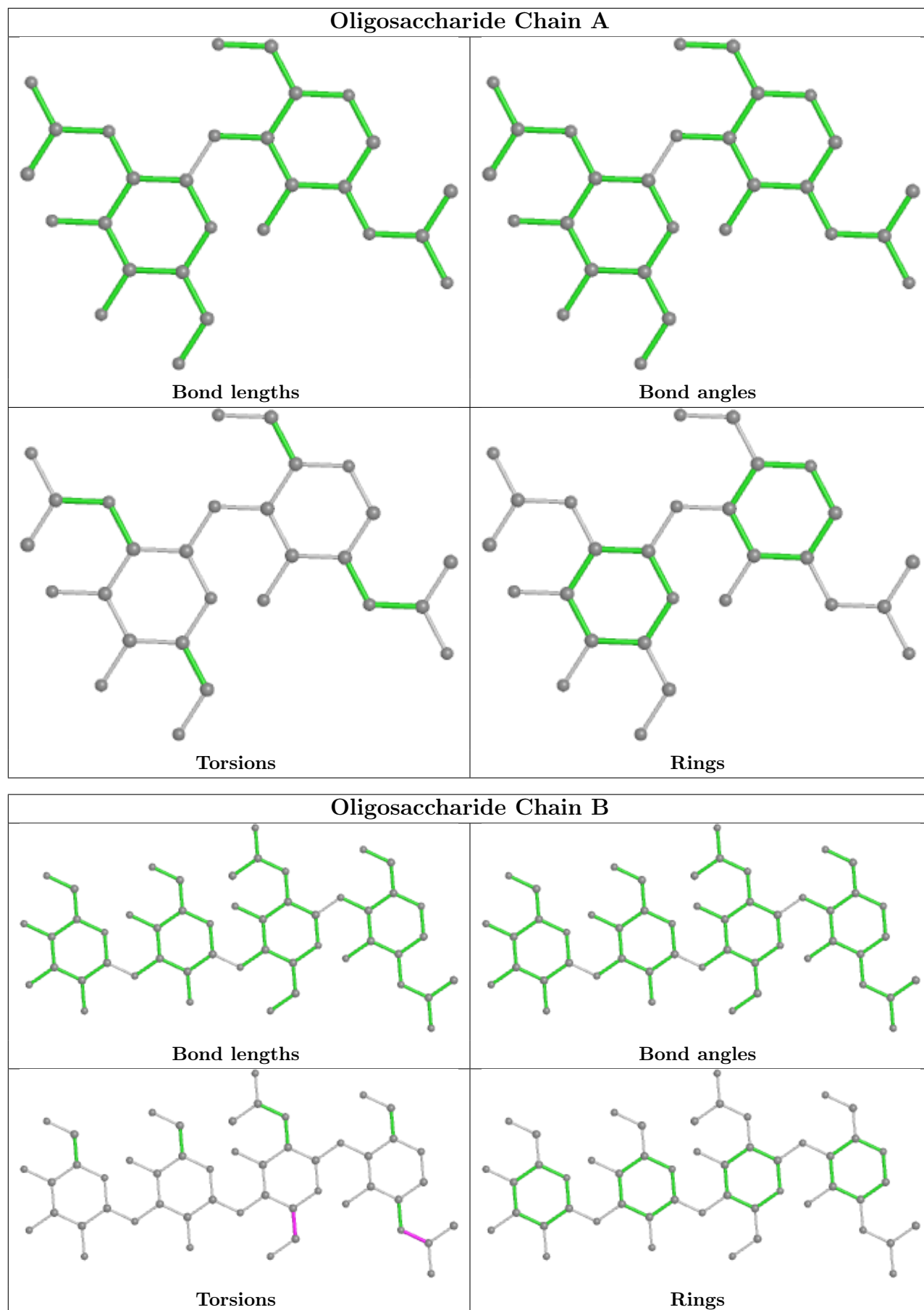
Mol	Chain	Res	Type	Atoms
3	B	1	NAG	C8-C7-N2-C2
3	B	1	NAG	O7-C7-N2-C2
3	B	2	NAG	O5-C5-C6-O6

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1	NAG	3	0
3	B	3	BMA	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](#)

5 ligands are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	F	608	-	4,4,4	0.14	0	6,6,6	0.11	0
4	NAG	F	601	1	14,14,15	0.49	0	17,19,21	0.54	0
5	SO4	F	609	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SO4	F	611	-	4,4,4	0.14	0	6,6,6	0.05	0
5	SO4	F	610	-	4,4,4	0.15	0	6,6,6	0.05	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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	F	601	1	-	0/6/23/26	0/1/1/1

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	601	NAG	1	0

## 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](#)

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	F	442/542 (81%)	-0.34	4 (0%) 84 82	35, 52, 91, 131	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	472	ALA	5.0
1	F	471	GLN	2.6
1	F	174	THR	2.3
1	F	175	ARG	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.

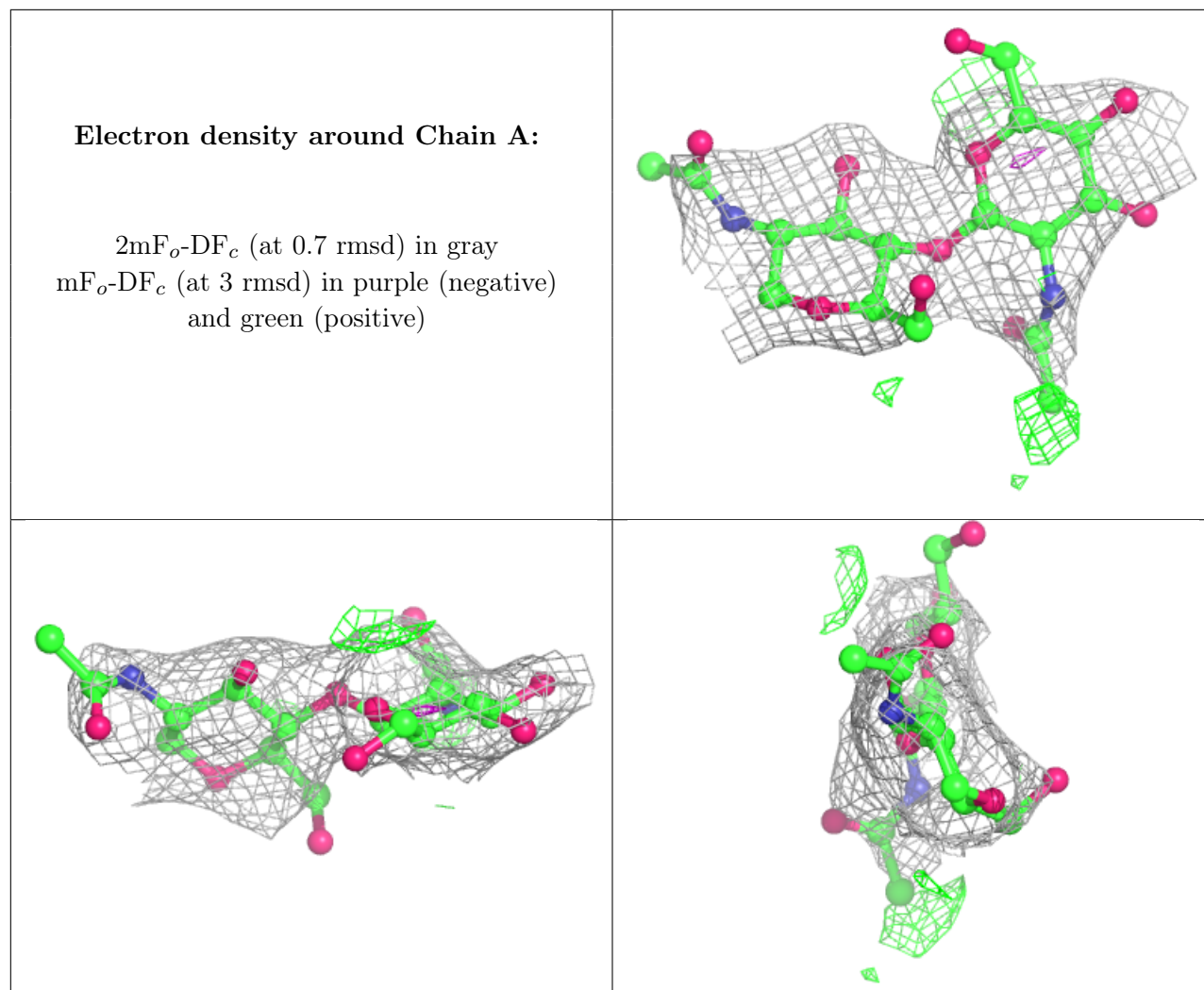
### 6.3 Carbohydrates [i](#)

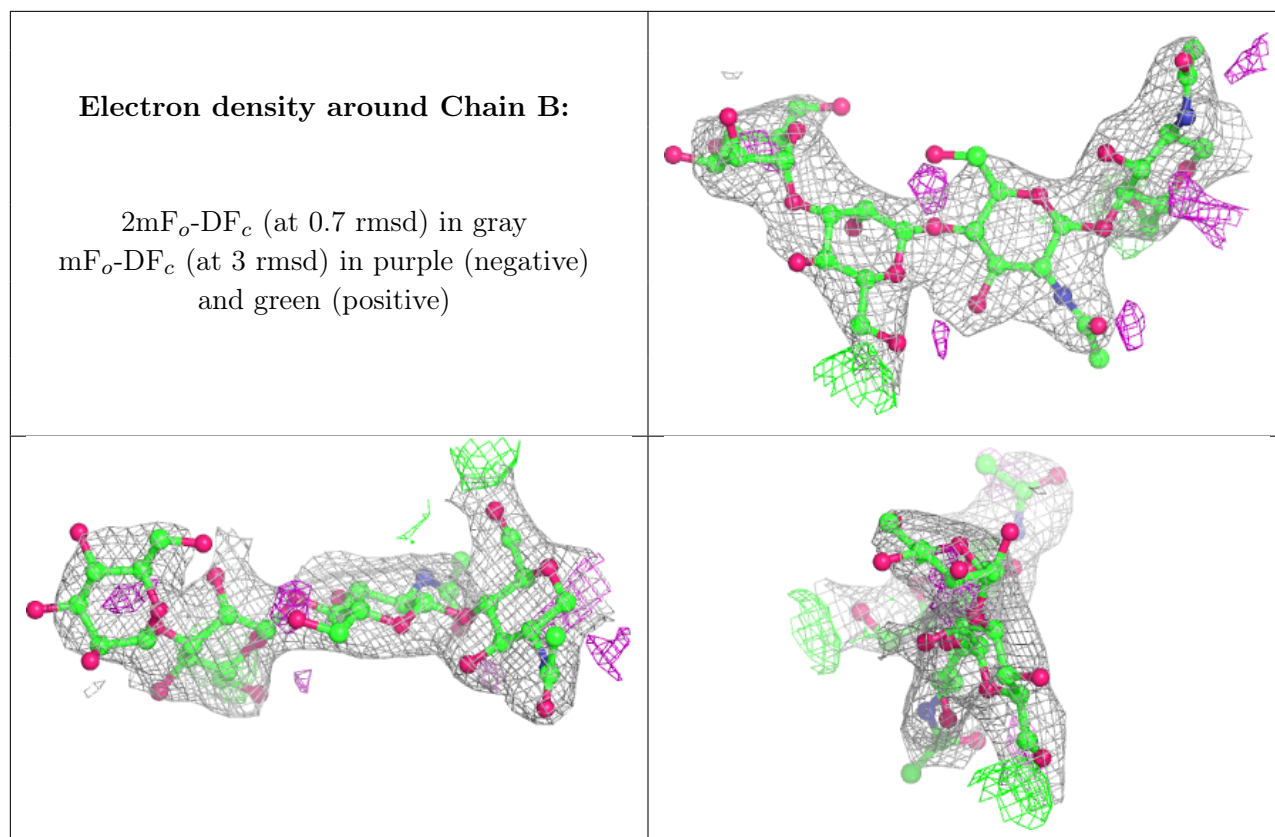
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 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
2	NAG	A	2	14/15	0.83	0.36	106,129,134,135	0
3	BMA	B	3	11/12	0.84	0.27	103,126,133,137	0
3	NAG	B	2	14/15	0.87	0.31	100,119,125,132	0
2	NAG	A	1	14/15	0.91	0.29	82,98,114,122	0
3	NAG	B	1	14/15	0.91	0.29	84,91,103,118	0
3	MAN	B	4	11/12	0.94	0.37	123,139,143,144	0

The following is a graphical depiction of the model fit to experimental electron density for oligosac-

charide. Each fit is shown from different orientation to approximate a three-dimensional view.





## 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(Å <sup>2</sup> )	Q<0.9
4	NAG	F	601	14/15	0.72	0.40	95,111,119,119	0
5	SO4	F	611	5/5	0.72	0.34	149,150,151,151	0
5	SO4	F	610	5/5	0.78	0.72	158,162,163,164	0
5	SO4	F	609	5/5	0.88	0.12	133,133,135,138	0
5	SO4	F	608	5/5	0.89	0.22	120,126,130,132	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.