



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2024 – 12:06 PM EST

PDB ID : 8TOS  
Title : ACE2-peptide 6 complex  
Authors : Christie, M.; Payne, R.  
Deposited on : 2023-08-03  
Resolution : 2.35 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 : 2.36  
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

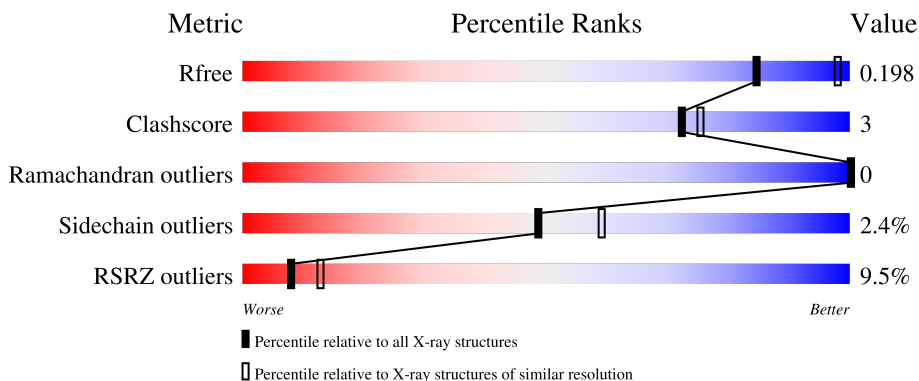
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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	1164 (2.36-2.36)
Clashscore	141614	1232 (2.36-2.36)
Ramachandran outliers	138981	1211 (2.36-2.36)
Sidechain outliers	138945	1212 (2.36-2.36)
RSRZ outliers	127900	1150 (2.36-2.36)

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	625	 9% 84% 10% 5%
2	E	17	 6% 59% 12% 29%
3	B	2	 100%
3	C	2	 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
3	FUC	C	2	-	-	-	X

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 9953 atoms, of which 4811 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 Angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	591	9436	3093	4606	797	910	30	0	1	0

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	615	GLY	-	expression tag	UNP Q9BYF1
A	616	SER	-	expression tag	UNP Q9BYF1
A	617	HIS	-	expression tag	UNP Q9BYF1
A	618	HIS	-	expression tag	UNP Q9BYF1
A	619	HIS	-	expression tag	UNP Q9BYF1
A	620	HIS	-	expression tag	UNP Q9BYF1
A	621	HIS	-	expression tag	UNP Q9BYF1
A	622	HIS	-	expression tag	UNP Q9BYF1
A	623	HIS	-	expression tag	UNP Q9BYF1
A	624	HIS	-	expression tag	UNP Q9BYF1
A	625	HIS	-	expression tag	UNP Q9BYF1
A	626	HIS	-	expression tag	UNP Q9BYF1
A	627	SER	-	expression tag	UNP Q9BYF1
A	628	GLY	-	expression tag	UNP Q9BYF1
A	629	LEU	-	expression tag	UNP Q9BYF1
A	630	ASN	-	expression tag	UNP Q9BYF1
A	631	ASP	-	expression tag	UNP Q9BYF1
A	632	ILE	-	expression tag	UNP Q9BYF1
A	633	PHE	-	expression tag	UNP Q9BYF1
A	634	GLU	-	expression tag	UNP Q9BYF1
A	635	ALA	-	expression tag	UNP Q9BYF1
A	636	GLN	-	expression tag	UNP Q9BYF1
A	637	LYS	-	expression tag	UNP Q9BYF1
A	638	ILE	-	expression tag	UNP Q9BYF1
A	639	GLU	-	expression tag	UNP Q9BYF1
A	640	TRP	-	expression tag	UNP Q9BYF1
A	641	HIS	-	expression tag	UNP Q9BYF1

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	642	GLU	-	expression tag	UNP Q9BYF1

- Molecule 2 is a protein called Peptide 6.

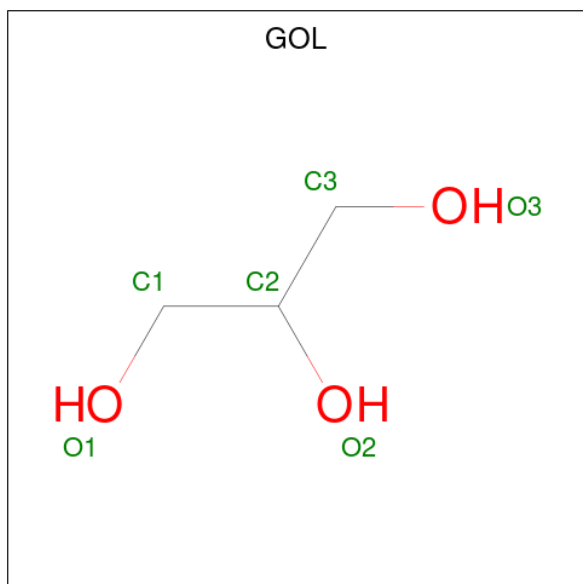
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	H	N				O
2	E	12	217	71	108	24	14	0	0	0

- Molecule 3 is an oligosaccharide called alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	H	N				O
3	B	2	46	14	22	1	9	0	0	0
3	C	2	46	14	22	1	9	0	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



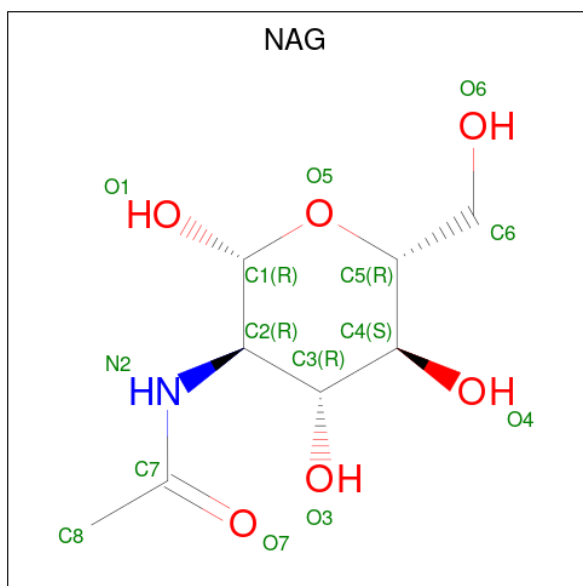
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	
			Total	C	H			O
4	A	1	14	3	8	3	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
4	A	1	Total	C	H	O	0	0
			14	3	8	3		
4	A	1	Total	C	H	O	0	0
			14	3	8	3		
4	A	1	Total	C	H	O	0	0
			14	3	8	3		
4	A	1	Total	C	H	O	0	0
			14	3	8	3		

- Molecule 5 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	H	N	O		
5	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
6	A	1	Total	Zn	0	0
			1	1		

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Cl 1 1	0	0

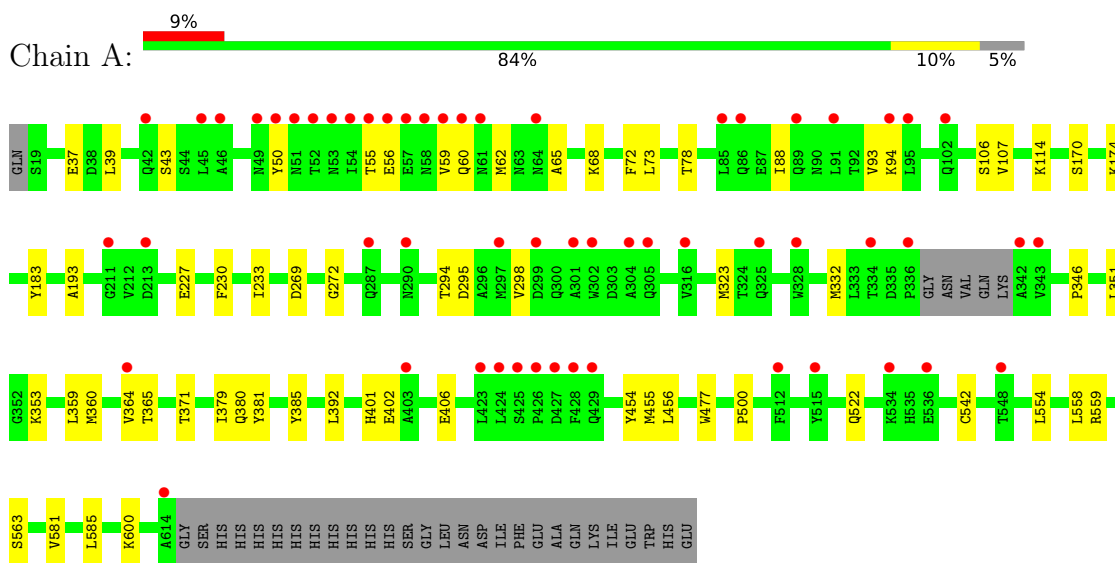
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	107	Total O 107 107	0	0
8	E	2	Total O 2 2	0	0

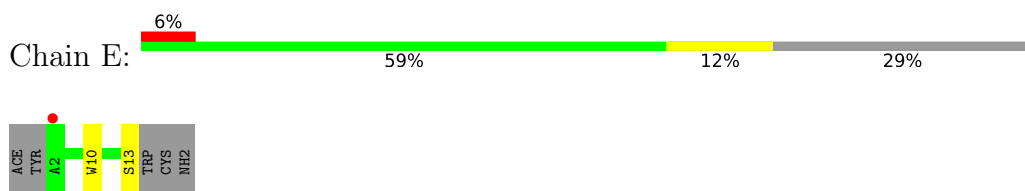
### 3 Residue-property plots [i](#)

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 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.

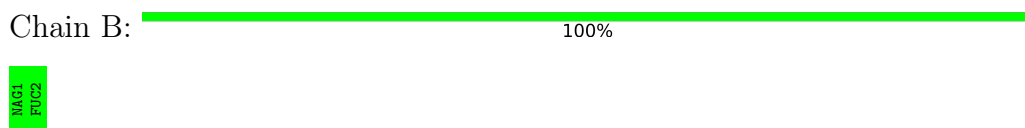
- Molecule 1: Angiotensin-converting enzyme 2



- Molecule 2: Peptide 6



- Molecule 3: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 3: alpha-L-fucopyranose-(1-6)-2-acetamido-2-deoxy-beta-D-glucopyranose





## 4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	179.57Å 179.57Å 70.25Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.08 – 2.35 45.08 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.2 (45.08-2.35) 99.2 (45.08-2.35)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.30 (at 2.34Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, $R_{free}$	0.182 , 0.200 0.180 , 0.198	Depositor DCC
$R_{free}$ test set	1704 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	39.3	Xtrriage
Anisotropy	0.655	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.42 , 41.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.013 for h,-h-k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9953	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.14% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, GOL, FUC, NAG, CL

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	0.28	0/4969	0.49	0/6750
2	E	0.30	0/114	0.81	0/155
All	All	0.28	0/5083	0.50	0/6905

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	4830	4606	4605	33	1
2	E	109	108	108	1	0
3	B	24	22	22	0	0
3	C	24	22	22	0	0
4	A	30	40	40	0	0
5	A	14	13	13	0	0
6	A	1	0	0	0	0
7	A	1	0	0	0	0
8	A	107	0	0	0	0
8	E	2	0	0	0	0
All	All	5142	4811	4810	33	1

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 33 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:56:GLU:O	1:A:60:GLN:NE2	2.28	0.67
1:A:233:ILE:CG2	1:A:581:VAL:HG21	2.26	0.65
1:A:233:ILE:HG23	1:A:581:VAL:HG21	1.79	0.62
1:A:323[B]:MET:CE	1:A:379:ILE:HG21	2.34	0.57
1:A:55:THR:O	1:A:59:VAL:HG23	2.03	0.57

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:78:THR:OG1	1:A:114:LYS:NZ[3_555]	2.08	0.12

## 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	588/625 (94%)	582 (99%)	6 (1%)	0	100	100
2	E	10/17 (59%)	9 (90%)	1 (10%)	0	100	100
All	All	598/642 (93%)	591 (99%)	7 (1%)	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	A	523/552 (95%)	511 (98%)	12 (2%)	50	61
2	E	11/14 (79%)	10 (91%)	1 (9%)	9	8
All	All	534/566 (94%)	521 (98%)	13 (2%)	49	59

5 of 13 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	455	MET
1	A	522	GLN
2	E	13	SER
1	A	559	ARG
1	A	600	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	60	GLN
1	A	117	ASN
1	A	380	GLN
1	A	522	GLN
2	E	7	HIS

### 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
3	NAG	B	1	1,3	14,14,15	0.27	0	17,19,21	0.51	0
3	FUC	B	2	3	10,10,11	0.64	0	14,14,16	0.90	0
3	NAG	C	1	1,3	14,14,15	0.31	0	17,19,21	0.58	0
3	FUC	C	2	3	10,10,11	0.66	0	14,14,16	1.16	2 (14%)

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
3	NAG	B	1	1,3	-	1/6/23/26	0/1/1/1
3	FUC	B	2	3	-	-	0/1/1/1
3	NAG	C	1	1,3	-	0/6/23/26	0/1/1/1
3	FUC	C	2	3	-	-	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	2	FUC	O5-C5-C4	2.59	114.17	109.52
3	C	2	FUC	C1-O5-C5	2.25	117.88	112.78

There are no chirality outliers.

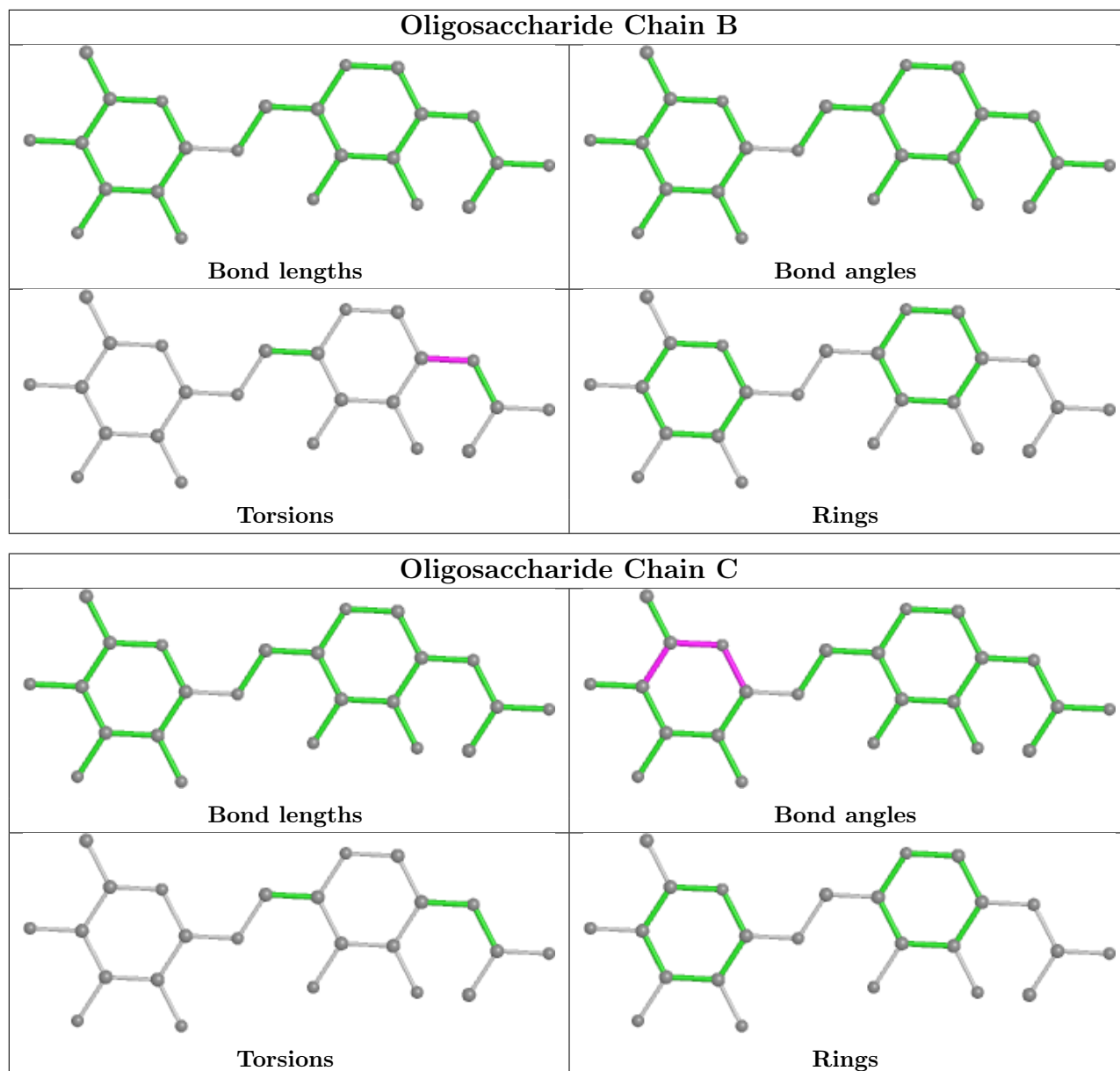
All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1	NAG	C1-C2-N2-C7

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.



## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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
4	GOL	A	702	-	5,5,5	0.89	0	5,5,5	1.07	0
4	GOL	A	705	-	5,5,5	0.77	0	5,5,5	0.94	0
4	GOL	A	701	-	5,5,5	0.85	0	5,5,5	1.04	0
4	GOL	A	706	-	5,5,5	0.82	0	5,5,5	0.97	0
5	NAG	A	703	1	14,14,15	0.17	0	17,19,21	0.64	0
4	GOL	A	704	-	5,5,5	0.84	0	5,5,5	0.86	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	GOL	A	702	-	-	0/4/4/4	-
4	GOL	A	705	-	-	2/4/4/4	-
4	GOL	A	701	-	-	0/4/4/4	-
4	GOL	A	706	-	-	2/4/4/4	-
5	NAG	A	703	1	-	2/6/23/26	0/1/1/1
4	GOL	A	704	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	705	GOL	O1-C1-C2-C3
5	A	703	NAG	O5-C5-C6-O6
5	A	703	NAG	C4-C5-C6-O6
4	A	706	GOL	O1-C1-C2-C3
4	A	705	GOL	O1-C1-C2-O2

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](#)

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	591/625 (94%)	0.62	56 (9%) 8 13	32, 48, 89, 122	0
2	E	12/17 (70%)	0.77	1 (8%) 11 16	52, 60, 74, 77	0
All	All	603/642 (93%)	0.62	57 (9%) 8 13	32, 48, 89, 122	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	56	GLU	6.9
1	A	54	ILE	6.4
1	A	91	LEU	5.7
1	A	55	THR	5.2
1	A	342	ALA	5.0

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

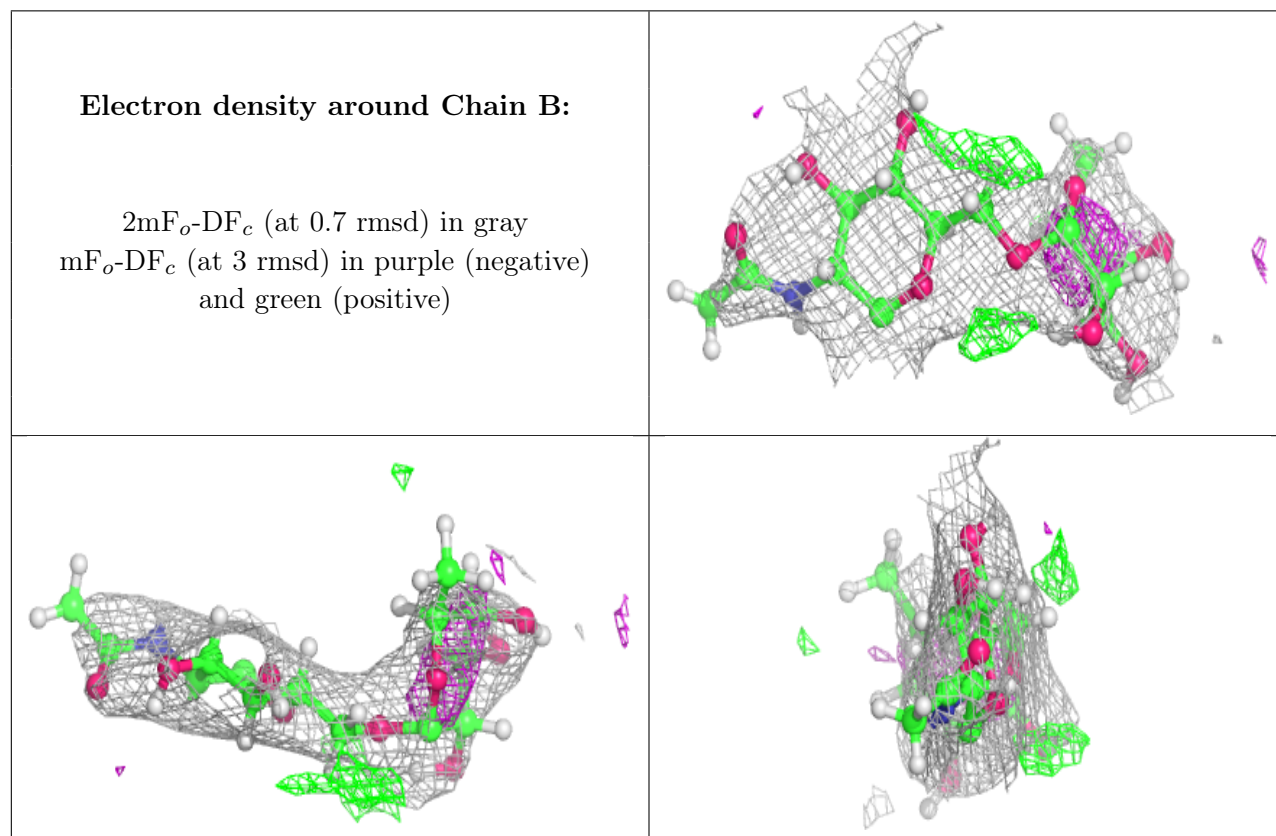
There are no non-standard protein/DNA/RNA residues in this entry.

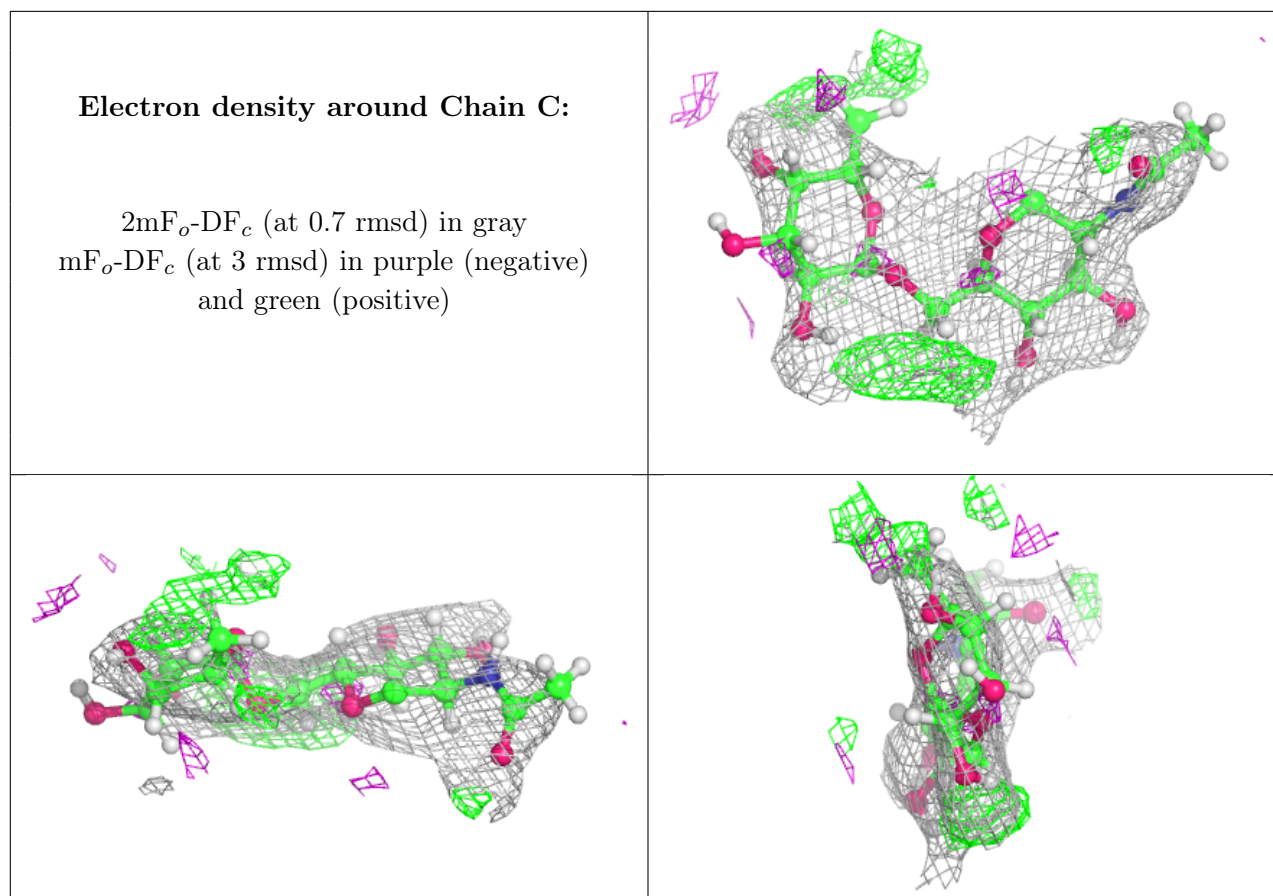
### 6.3 Carbohydrates [i](#)

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	FUC	C	2	10/11	0.72	0.41	84,93,106,111	0
3	NAG	C	1	14/15	0.79	0.23	76,87,103,104	0
3	FUC	B	2	10/11	0.80	0.47	75,88,97,100	0
3	NAG	B	1	14/15	0.86	0.26	79,90,105,109	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.





## 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	GOL	A	704	6/6	0.76	0.19	48,59,71,71	0
6	ZN	A	707	1/1	0.76	0.32	98,98,98,98	0
4	GOL	A	705	6/6	0.80	0.28	54,65,71,74	0
5	NAG	A	703	14/15	0.88	0.34	82,97,116,121	0
4	GOL	A	702	6/6	0.90	0.14	46,56,62,62	0
4	GOL	A	701	6/6	0.92	0.15	34,43,54,65	0
4	GOL	A	706	6/6	0.95	0.10	59,71,79,80	0
7	CL	A	708	1/1	0.99	0.26	37,37,37,37	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.