



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 13, 2025 – 02:51 PM JST

PDB ID : 8XYE  
Title : Crystal structure of SARS-CoV-2 BA.4 RBD and human ACE2  
Authors : Lan, J.; Wang, C.H.  
Deposited on : 2024-01-19  
Resolution : 3.32 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:

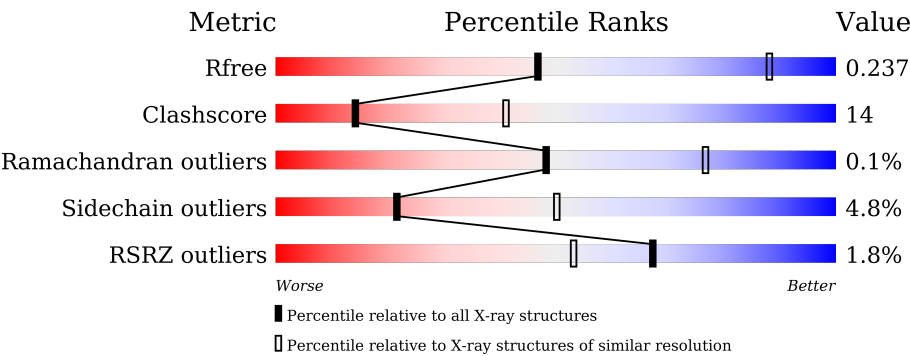
MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.21  
EDS : 3.0  
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)  
CCP4 : 9.0.004 (Gargrove)  
Density-Fitness : 1.0.11  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.40

# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:  
*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.32 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))
$R_{free}$	164625	1066 (3.34-3.30)
Clashscore	180529	1111 (3.34-3.30)
Ramachandran outliers	177936	1109 (3.34-3.30)
Sidechain outliers	177891	1108 (3.34-3.30)
RSRZ outliers	164620	1066 (3.34-3.30)

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	598	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>78%20%.</div></div>
1	B	598	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>72%27%.</div></div>
2	C	194	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>5%46%47%7%.</div></div>
2	D	194	<div><div></div><div><div></div><div></div><div></div><div></div></div><div>5%52%41%7%.</div></div>
3	E	3	<div><div></div><div><div></div><div></div></div><div>33%67%</div></div>
3	F	3	<div><div></div><div><div></div><div></div></div><div>33%67%</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	G	3	 33%67%
4	H	2	 100%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13101 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 Processed angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	598	Total	C	N	O	S	0	1	0
			4884	3125	809	921	29			
1	A	597	Total	C	N	O	S	0	1	0
			4876	3121	808	918	29			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	17	ASP	-	expression tag	UNP Q9BYF1
B	18	PRO	-	expression tag	UNP Q9BYF1
A	17	ASP	-	expression tag	UNP Q9BYF1
A	18	PRO	-	expression tag	UNP Q9BYF1

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	194	Total	C	N	O	S	0	0	0
			1549	997	262	282	8			
2	D	194	Total	C	N	O	S	0	0	0
			1549	997	262	282	8			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	339	ASP	GLY	variant	UNP P0DTC2
C	371	PHE	SER	variant	UNP P0DTC2
C	373	PRO	SER	variant	UNP P0DTC2
C	375	PHE	SER	variant	UNP P0DTC2
C	405	ASN	ASP	variant	UNP P0DTC2
C	408	SER	ARG	variant	UNP P0DTC2
C	417	ASN	LYS	variant	UNP P0DTC2
C	440	LYS	ASN	variant	UNP P0DTC2

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	452	ARG	LEU	variant	UNP P0DTC2
C	477	ASN	SER	variant	UNP P0DTC2
C	478	LYS	THR	variant	UNP P0DTC2
C	484	ALA	GLU	variant	UNP P0DTC2
C	486	VAL	PHE	variant	UNP P0DTC2
C	498	ARG	GLN	variant	UNP P0DTC2
C	501	TYR	ASN	variant	UNP P0DTC2
C	505	HIS	TYR	variant	UNP P0DTC2
D	339	ASP	GLY	variant	UNP P0DTC2
D	371	PHE	SER	variant	UNP P0DTC2
D	373	PRO	SER	variant	UNP P0DTC2
D	375	PHE	SER	variant	UNP P0DTC2
D	405	ASN	ASP	variant	UNP P0DTC2
D	408	SER	ARG	variant	UNP P0DTC2
D	417	ASN	LYS	variant	UNP P0DTC2
D	440	LYS	ASN	variant	UNP P0DTC2
D	452	ARG	LEU	variant	UNP P0DTC2
D	477	ASN	SER	variant	UNP P0DTC2
D	478	LYS	THR	variant	UNP P0DTC2
D	484	ALA	GLU	variant	UNP P0DTC2
D	486	VAL	PHE	variant	UNP P0DTC2
D	498	ARG	GLN	variant	UNP P0DTC2
D	501	TYR	ASN	variant	UNP P0DTC2
D	505	HIS	TYR	variant	UNP P0DTC2

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



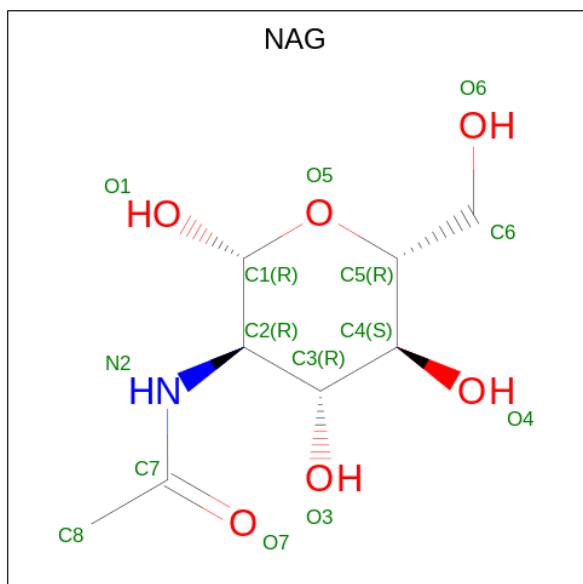
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	3	Total	C	N	O	0	0	0
			39	22	2	15			
3	F	3	Total	C	N	O	0	0	0
			39	22	2	15			
3	G	3	Total	C	N	O	0	0	0
			39	22	2	15			

- 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
4	H	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

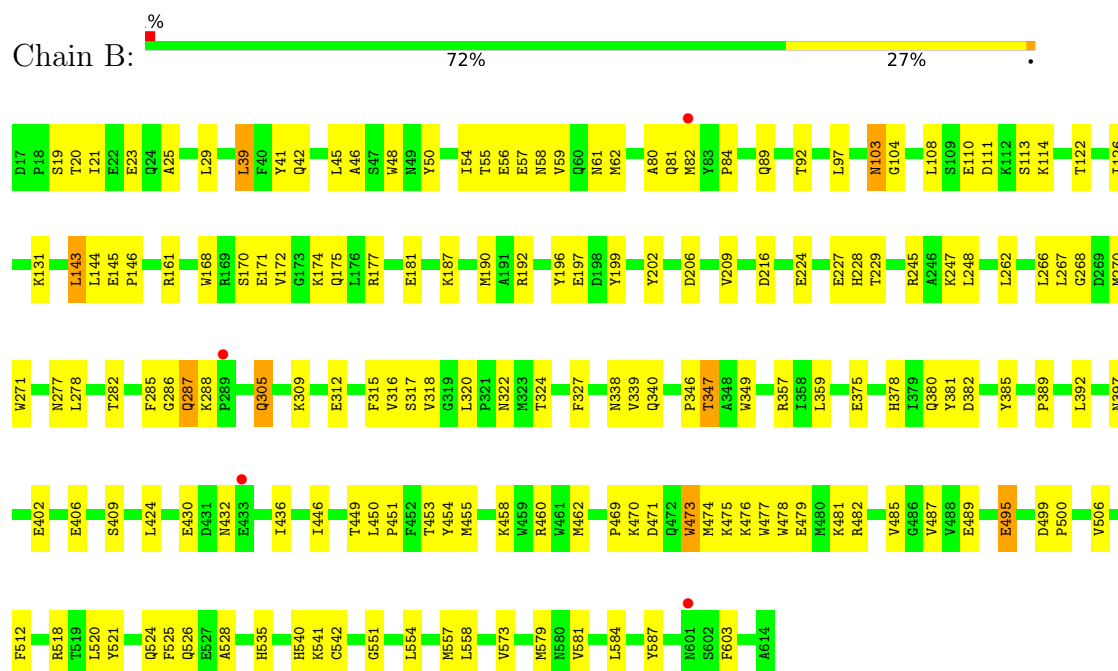


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	B	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		

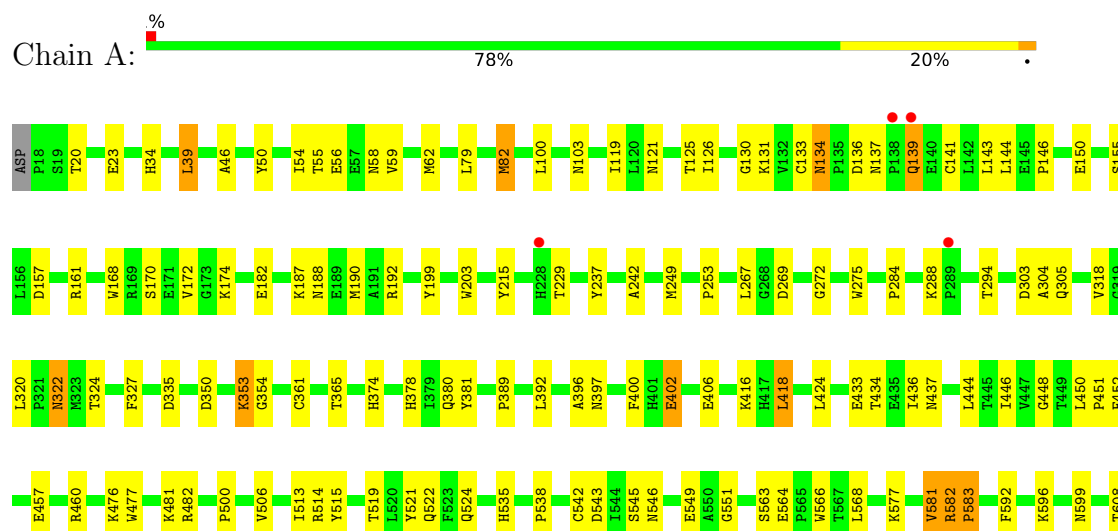
### 3 Residue-property plots

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 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: Processed angiotensin-converting enzyme 2



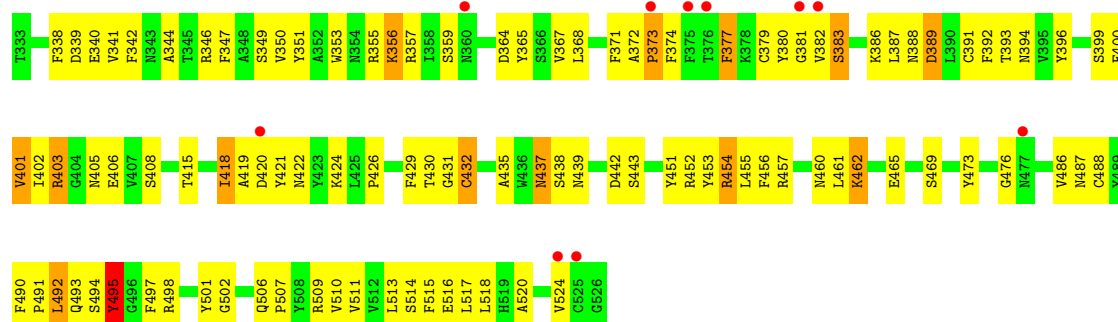
- Molecule 1: Processed angiotensin-converting enzyme 2





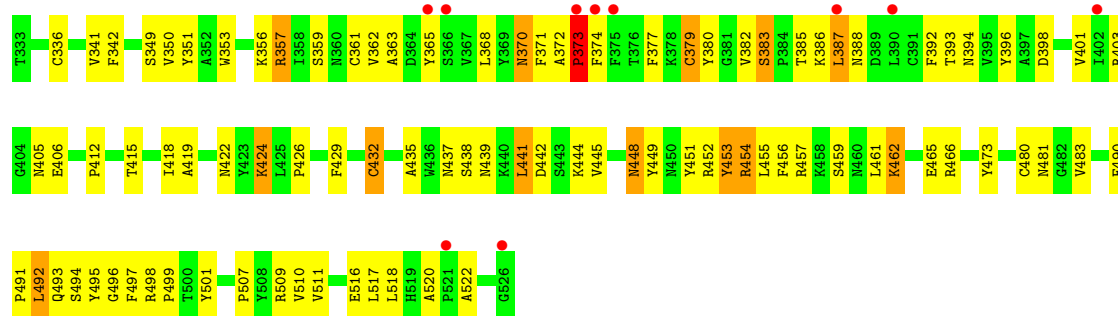
- Molecule 2: Spike protein S1

Chain C: 5% 46% 47% 7%



- Molecule 2: Spike protein S1

Chain D: 5% 52% 41% 7%



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

Chain E: 33% 67%




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

Chain F: 33% 67%



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



Chain G:  33% 67%

MAG1  
MAG2  
BGA3

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

Chain H:  100%

MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	84.47Å 165.82Å 98.98Å 90.00° 102.18° 90.00°	Depositor
Resolution (Å)	31.37 – 3.32 31.37 – 3.32	Depositor EDS
% Data completeness (in resolution range)	98.3 (31.37-3.32) 98.2 (31.37-3.32)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 3.31Å)	Xtriage
Refinement program	PHENIX (1.19.1_4122: ???)	Depositor
R, $R_{free}$	0.191 , 0.242 0.197 , 0.237	Depositor DCC
$R_{free}$ test set	1973 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.4	Xtriage
Anisotropy	0.340	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 48.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	13101	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.96% 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, 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	0.61	3/5018 (0.1%)	0.79	6/6818 (0.1%)
1	B	0.59	0/5026	0.79	6/6830 (0.1%)
2	C	0.74	1/1595 (0.1%)	0.94	4/2171 (0.2%)
2	D	0.63	0/1595	0.88	5/2171 (0.2%)
All	All	0.62	4/13234 (0.0%)	0.82	21/17990 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	C	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	402	GLU	CB-CG	7.80	1.67	1.52
1	A	203	TRP	CB-CG	-6.03	1.39	1.50
1	A	402	GLU	CG-CD	5.83	1.60	1.51
2	C	495	TYR	CE1-CZ	-5.39	1.31	1.38

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	492	LEU	CB-CG-CD2	-8.19	97.08	111.00
2	D	455	LEU	CA-CB-CG	7.41	132.35	115.30
1	A	402	GLU	OE1-CD-OE2	-7.38	114.44	123.30
1	B	584	LEU	CB-CG-CD2	-6.12	100.59	111.00
2	C	373	PRO	N-CA-C	-6.10	96.24	112.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	495	TYR	Sidechain

## 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	4876	0	4649	87	0
1	B	4884	0	4650	99	0
2	C	1549	0	1472	98	0
2	D	1549	0	1472	76	0
3	E	39	0	34	0	0
3	F	39	0	34	3	0
3	G	39	0	34	0	0
4	H	28	0	25	1	0
5	A	42	0	39	1	0
5	B	56	0	52	0	0
All	All	13101	0	12461	359	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 14.

The worst 5 of 359 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)
2:D:453:TYR:CD2	2:D:495:TYR:HE1	1.85	0.95
1:A:378:HIS:NE2	1:A:402:GLU:OE2	2.02	0.92
2:C:393:THR:HG21	2:C:518:LEU:HD23	1.53	0.89
2:D:442:ASP:OD1	2:D:451:TYR:OH	1.93	0.86
2:D:453:TYR:CD2	2:D:495:TYR:CE1	2.65	0.85

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	596/598 (100%)	568 (95%)	28 (5%)	0	100	100
1	B	597/598 (100%)	563 (94%)	34 (6%)	0	100	100
2	C	192/194 (99%)	172 (90%)	20 (10%)	0	100	100
2	D	192/194 (99%)	166 (86%)	25 (13%)	1 (0%)	25	57
All	All	1577/1584 (100%)	1469 (93%)	107 (7%)	1 (0%)	48	77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	373	PRO

### 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	528/528 (100%)	515 (98%)	13 (2%)	42	67
1	B	529/528 (100%)	510 (96%)	19 (4%)	30	57
2	C	167/167 (100%)	152 (91%)	15 (9%)	8	28
2	D	167/167 (100%)	147 (88%)	20 (12%)	4	17
All	All	1391/1390 (100%)	1324 (95%)	67 (5%)	21	50

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	139	GLN
1	A	305	GLN
1	A	582	ARG
2	C	389	ASP
2	C	383	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	460	ASN
1	A	380	GLN
1	A	535	HIS
1	B	287	GLN
1	B	24	GLN

### 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 ⓘ

11 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	E	1	1,3	14,14,15	0.72	1 (7%)	17,19,21	0.58	0
3	NAG	E	2	3	14,14,15	0.53	0	17,19,21	0.67	0
3	BMA	E	3	3	11,11,12	3.30	8 (72%)	15,15,17	1.78	5 (33%)
3	NAG	F	1	1,3	14,14,15	1.07	1 (7%)	17,19,21	1.08	2 (11%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	NAG	F	2	3	14,14,15	0.80	1 (7%)	17,19,21	1.18	1 (5%)
3	BMA	F	3	3	11,11,12	3.81	9 (81%)	15,15,17	2.21	6 (40%)
3	NAG	G	1	1,3	14,14,15	1.03	1 (7%)	17,19,21	0.55	0
3	NAG	G	2	3	14,14,15	0.62	0	17,19,21	0.94	0
3	BMA	G	3	3	11,11,12	2.80	8 (72%)	15,15,17	1.58	3 (20%)
4	NAG	H	1	1,4	14,14,15	1.17	1 (7%)	17,19,21	0.92	1 (5%)
4	NAG	H	2	4	14,14,15	0.93	1 (7%)	17,19,21	0.72	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
3	NAG	E	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	E	2	3	-	0/6/23/26	0/1/1/1
3	BMA	E	3	3	-	1/2/19/22	0/1/1/1
3	NAG	F	1	1,3	-	0/6/23/26	0/1/1/1
3	NAG	F	2	3	-	0/6/23/26	0/1/1/1
3	BMA	F	3	3	-	1/2/19/22	0/1/1/1
3	NAG	G	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	G	2	3	-	3/6/23/26	0/1/1/1
3	BMA	G	3	3	-	1/2/19/22	0/1/1/1
4	NAG	H	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	H	2	4	-	4/6/23/26	0/1/1/1

The worst 5 of 31 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	3	BMA	O5-C1	7.06	1.55	1.43
3	E	3	BMA	C4-C3	5.06	1.65	1.52
3	F	3	BMA	C1-C2	4.72	1.63	1.52
3	F	3	BMA	O5-C5	4.71	1.53	1.43
3	F	3	BMA	C4-C5	4.58	1.62	1.53

The worst 5 of 18 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	3	BMA	O5-C5-C6	4.87	114.84	107.20
3	E	3	BMA	C1-O5-C5	3.56	117.02	112.19

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	2	NAG	O4-C4-C3	-3.20	102.95	110.35
3	F	3	BMA	O5-C1-C2	3.11	115.56	110.77
3	E	3	BMA	O3-C3-C4	3.10	117.52	110.35

There are no chirality outliers.

5 of 14 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	G	2	NAG	C8-C7-N2-C2
3	G	2	NAG	O7-C7-N2-C2
4	H	1	NAG	C8-C7-N2-C2
4	H	1	NAG	O7-C7-N2-C2
4	H	2	NAG	C8-C7-N2-C2

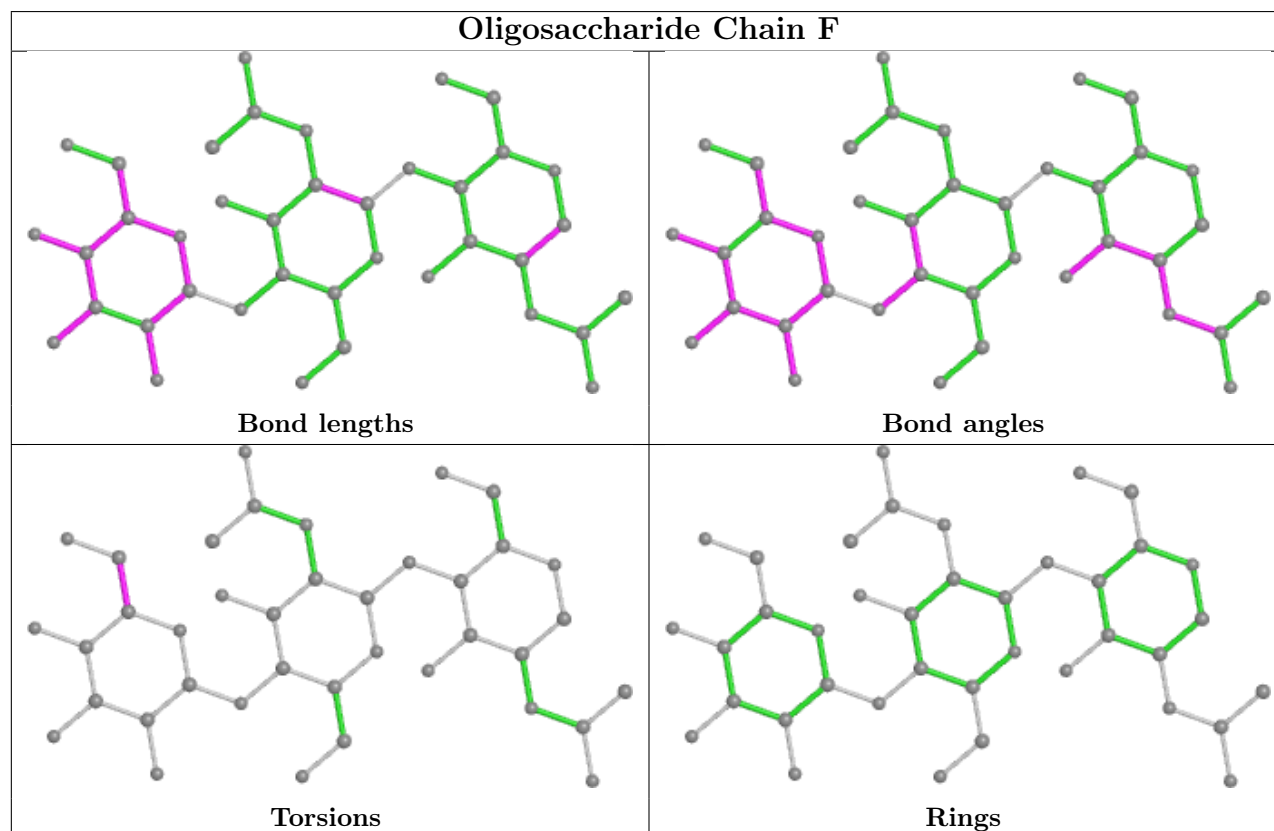
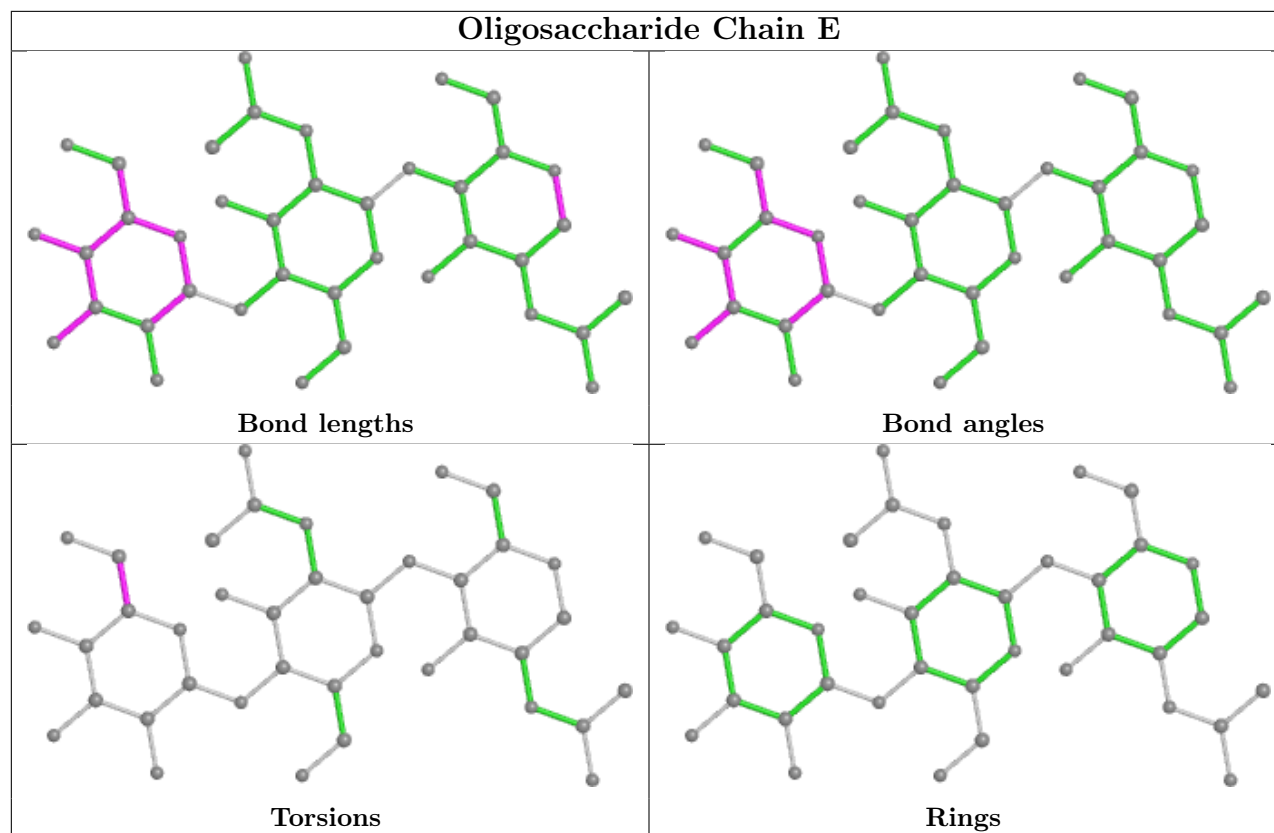
There are no ring outliers.

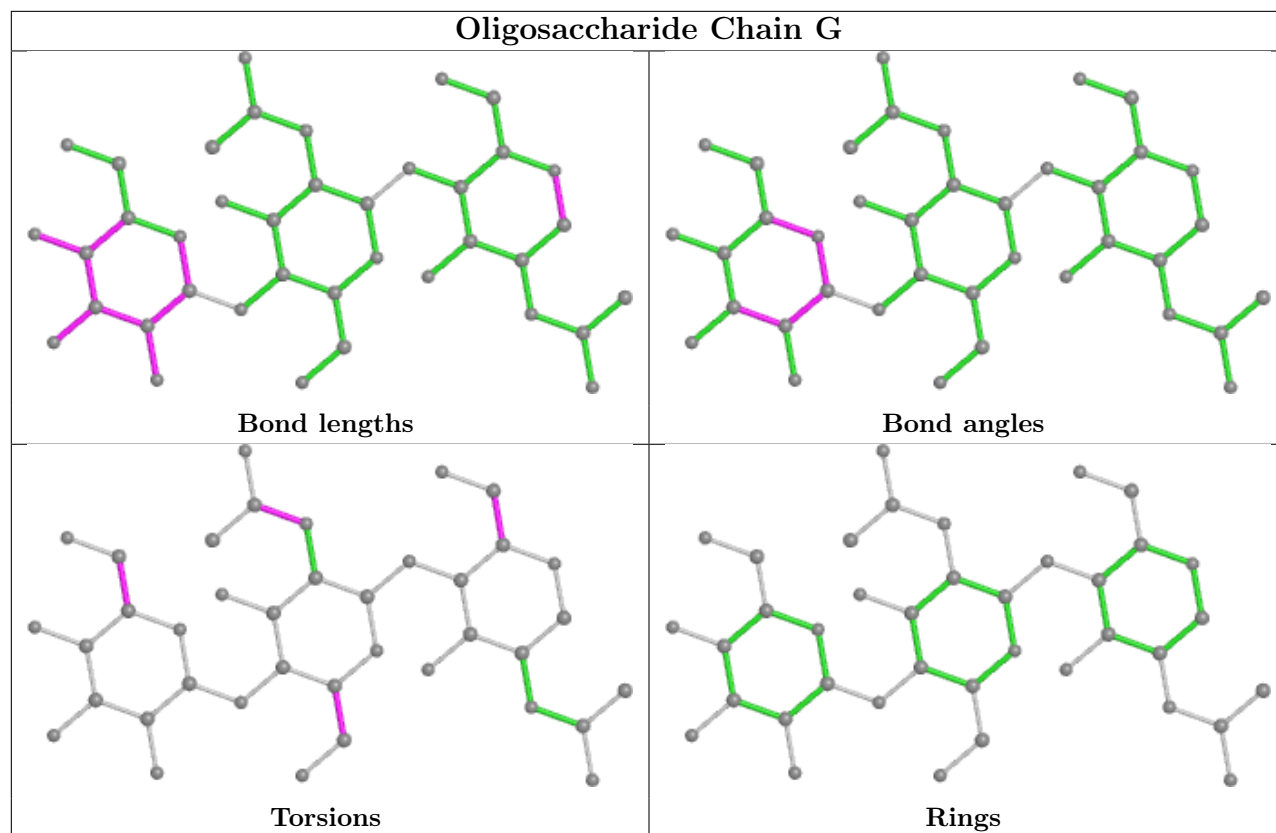
4 monomers are involved in 4 short contacts:

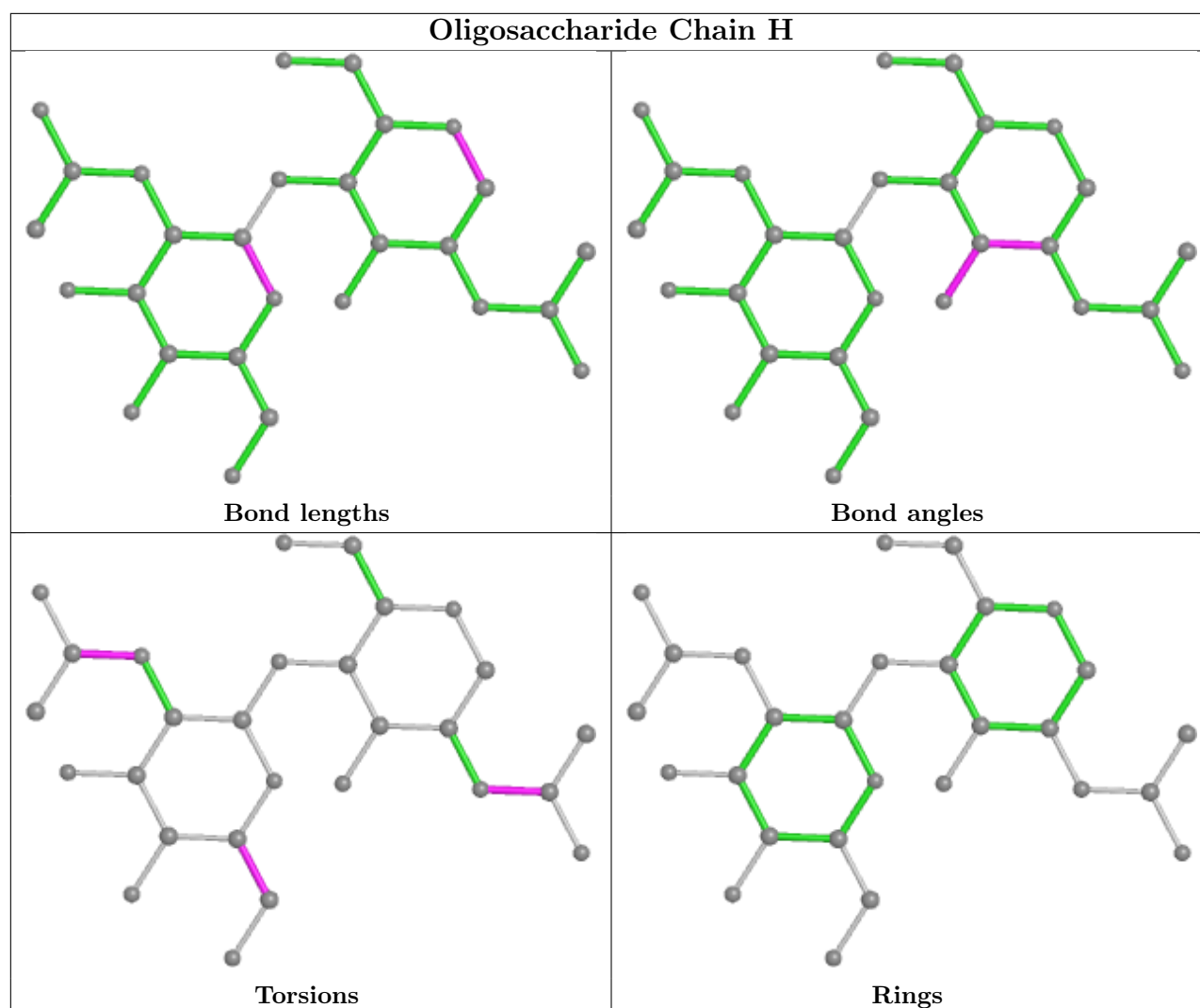
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	2	NAG	3	0
4	H	1	NAG	1	0
4	H	2	NAG	1	0
3	F	3	BMA	1	0

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](#)

7 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	NAG	B	701	1	14,14,15	1.60	1 (7%)	17,19,21	1.73	1 (5%)
5	NAG	A	702	1	14,14,15	1.40	2 (14%)	17,19,21	1.10	1 (5%)
5	NAG	A	701	1	14,14,15	2.18	1 (7%)	17,19,21	1.18	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	B	702	1	14,14,15	1.67	1 (7%)	17,19,21	1.36	3 (17%)
5	NAG	B	703	1	14,14,15	1.43	1 (7%)	17,19,21	0.74	0
5	NAG	A	703	1	14,14,15	1.96	2 (14%)	17,19,21	0.82	0
5	NAG	B	704	1	14,14,15	1.28	1 (7%)	17,19,21	0.89	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	NAG	B	701	1	-	1/6/23/26	0/1/1/1
5	NAG	A	702	1	-	4/6/23/26	0/1/1/1
5	NAG	A	701	1	-	2/6/23/26	0/1/1/1
5	NAG	B	702	1	-	3/6/23/26	0/1/1/1
5	NAG	B	703	1	-	1/6/23/26	0/1/1/1
5	NAG	A	703	1	-	2/6/23/26	0/1/1/1
5	NAG	B	704	1	-	2/6/23/26	0/1/1/1

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	701	NAG	O5-C1	7.86	1.56	1.43
5	A	703	NAG	O5-C1	6.39	1.53	1.43
5	B	702	NAG	O5-C1	5.74	1.52	1.43
5	B	701	NAG	O5-C1	5.65	1.52	1.43
5	B	703	NAG	O5-C1	5.04	1.51	1.43

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	701	NAG	C1-O5-C5	6.33	120.76	112.19
5	A	701	NAG	C1-O5-C5	3.75	117.27	112.19
5	A	702	NAG	C1-O5-C5	3.68	117.18	112.19
5	B	702	NAG	C1-O5-C5	3.48	116.91	112.19
5	B	702	NAG	C2-N2-C7	2.25	126.11	122.90

There are no chirality outliers.

5 of 15 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	702	NAG	O5-C5-C6-O6
5	A	703	NAG	O5-C5-C6-O6
5	B	704	NAG	C8-C7-N2-C2
5	B	704	NAG	O7-C7-N2-C2
5	A	701	NAG	C8-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	703	NAG	1	0

## 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 [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	597/598 (99%)	-0.46	4 (0%)	84 75	43, 68, 114, 247	1 (0%)
1	B	598/598 (100%)	-0.28	4 (0%)	84 75	47, 80, 134, 184	1 (0%)
2	C	194/194 (100%)	0.22	10 (5%)	34 26	47, 76, 162, 237	0
2	D	194/194 (100%)	0.29	10 (5%)	34 26	61, 90, 189, 260	0
All	All	1583/1584 (99%)	-0.22	28 (1%)	67 53	43, 76, 150, 260	2 (0%)

The worst 5 of 28 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	365	TYR	4.6
1	A	138	PRO	3.8
1	A	228[A]	HIS	3.6
2	C	375	PHE	3.1
2	D	366	SER	3.0

### 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.

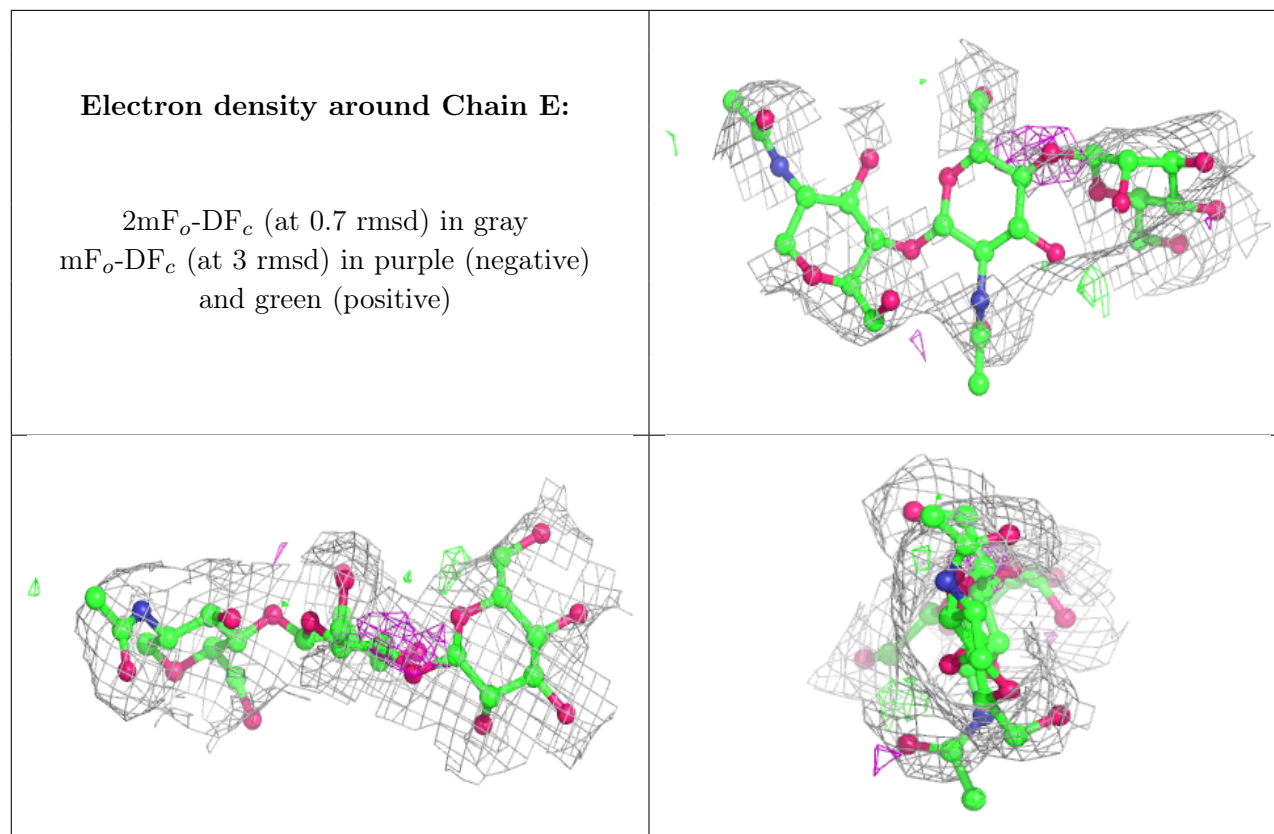
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	BMA	F	3	11/12	0.47	0.14	88,95,106,107	0
3	NAG	G	2	14/15	0.59	0.14	102,108,114,114	0

*Continued on next page...*

*Continued from previous page...*

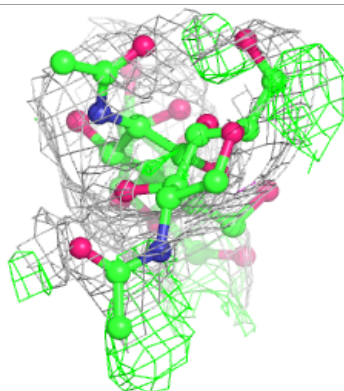
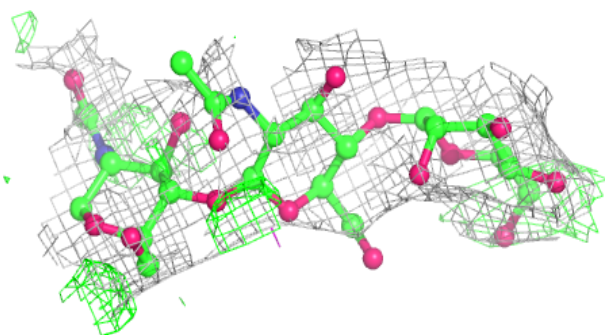
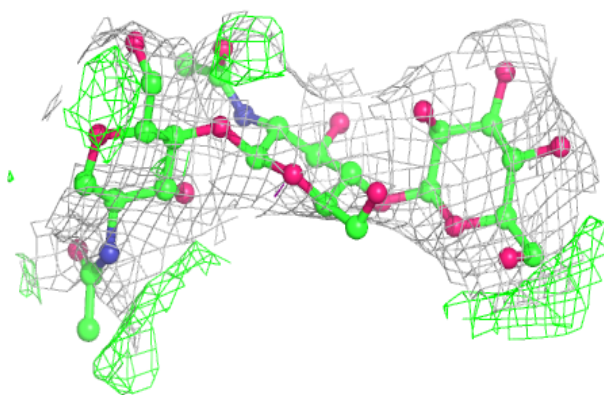
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	BMA	G	3	11/12	0.59	0.10	97,106,110,111	0
3	BMA	E	3	11/12	0.61	0.11	89,94,103,105	0
4	NAG	H	2	14/15	0.70	0.12	88,94,105,106	0
3	NAG	G	1	14/15	0.74	0.14	88,97,106,107	0
3	NAG	E	2	14/15	0.74	0.12	86,98,107,108	0
3	NAG	F	2	14/15	0.79	0.17	103,107,116,117	0
3	NAG	F	1	14/15	0.81	0.17	92,103,124,128	0
4	NAG	H	1	14/15	0.84	0.08	77,83,89,89	0
3	NAG	E	1	14/15	0.90	0.09	73,82,90,94	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.

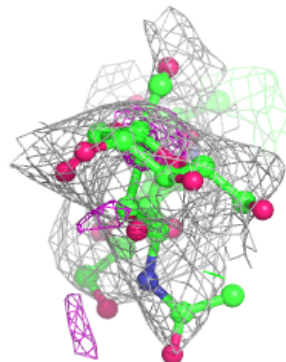
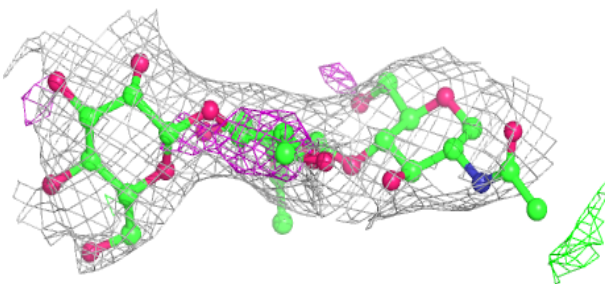
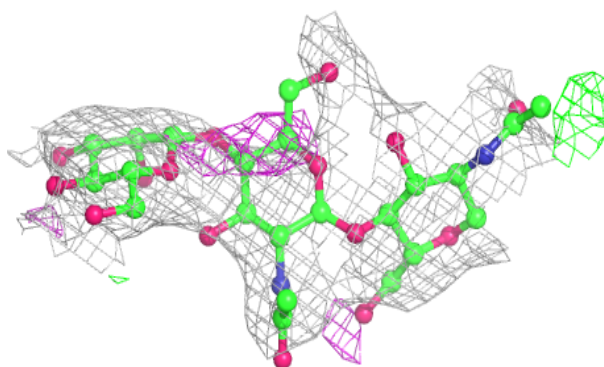


**Electron density around Chain F:**

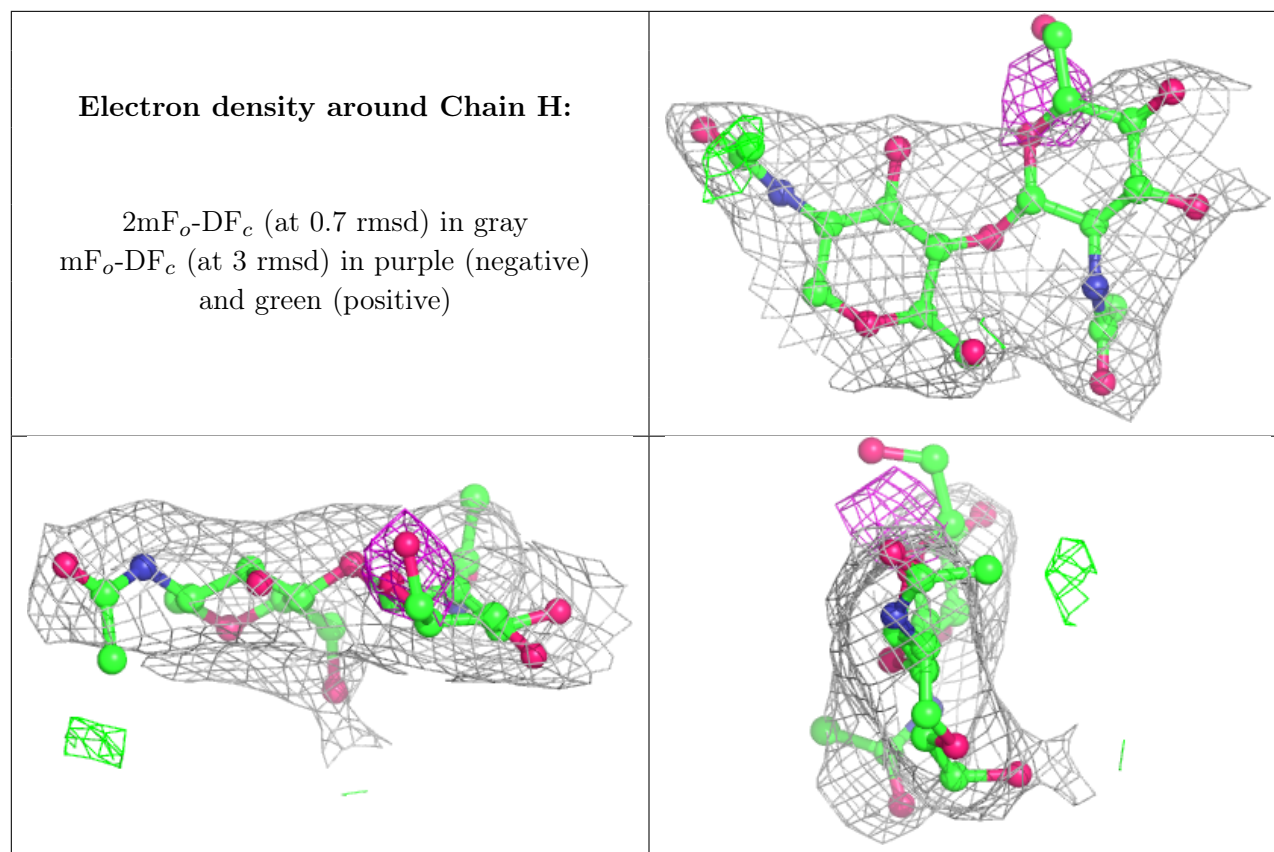
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

**Electron density around Chain G:**

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







## 6.4 Ligands [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
5	NAG	B	702	14/15	0.61	0.15	91,116,126,127	0
5	NAG	B	701	14/15	0.64	0.15	137,146,156,158	0
5	NAG	A	701	14/15	0.64	0.14	102,107,115,117	0
5	NAG	B	703	14/15	0.72	0.12	116,132,144,146	0
5	NAG	A	703	14/15	0.75	0.14	89,109,115,116	0
5	NAG	A	702	14/15	0.77	0.13	109,115,120,122	0
5	NAG	B	704	14/15	0.81	0.08	96,106,110,110	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.