



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 6, 2020 – 12:47 PM BST

PDB ID : 6BHJ  
Title : Structure of HIV-1 Reverse Transcriptase Bound to a 38-mer Hairpin  
Template-Primer RNA-DNA Aptamer  
Authors : Ruiz, F.X.; Miller, M.T.; Tuske, S.; Das, K.; Arnold, E.  
Deposited on : 2017-10-30  
Resolution : 2.81 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 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 : 2.13.1  
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) : 2.13.1

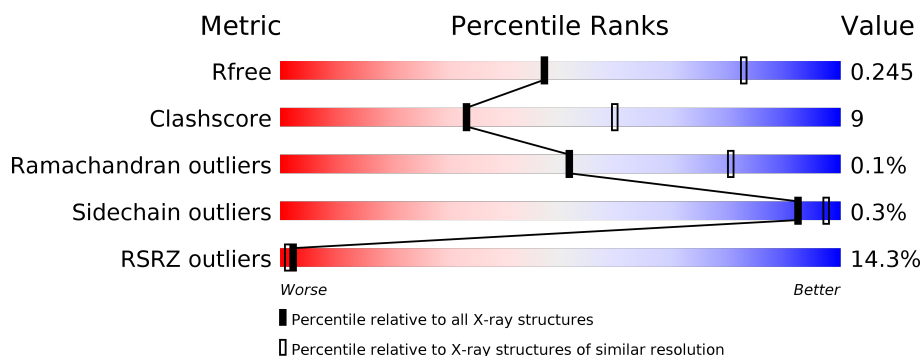
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.81 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	3617 (2.84-2.80)
Clashscore	141614	4060 (2.84-2.80)
Ramachandran outliers	138981	3978 (2.84-2.80)
Sidechain outliers	138945	3980 (2.84-2.80)
RSRZ outliers	127900	3552 (2.84-2.80)

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 on 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	557	<div> <div>13%</div> <div>92%</div> <div>7%</div> <div>..</div> </div>
1	C	557	<div> <div>25%</div> <div>90%</div> <div>9%</div> <div>.</div> </div>
2	B	444	<div> <div>5%</div> <div>90%</div> <div>6%</div> <div>.</div> </div>
2	D	444	<div> <div>8%</div> <div>88%</div> <div>8%</div> <div>.</div> </div>
3	E	38	<div> <div>18%</div> <div>47%</div> <div>26%</div> <div>8%</div> </div>
3	F	38	<div> <div>45%</div> <div>16%</div> <div>45%</div> <div>32%</div> <div>8%</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	G	2	
4	H	2	

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
6	GOL	B	2003	-	-	X	-
6	GOL	C	602	-	-	-	X
6	GOL	D	502	-	-	X	X
7	SO4	C	601	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 17498 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 HIV-1 REVERSE TRANSCRIPTASE P66 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	554	Total	C	N	O	S	0	0	0
			4509	2919	751	832	7			
1	C	553	Total	C	N	O	S	0	0	0
			4504	2916	750	831	7			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	LYS	ARG	variant	UNP A0A076Q3N8
A	280	SER	CYS	engineered mutation	UNP A0A076Q3N8
A	498	ASN	ASP	engineered mutation	UNP A0A076Q3N8
C	172	LYS	ARG	variant	UNP A0A076Q3N8
C	280	SER	CYS	engineered mutation	UNP A0A076Q3N8
C	498	ASN	ASP	engineered mutation	UNP A0A076Q3N8

- Molecule 2 is a protein called HIV-1 REVERSE TRANSCRIPTASE P51 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	416	Total	C	N	O	S	0	0	0
			3439	2242	568	622	7			
2	D	410	Total	C	N	O	S	0	0	0
			3394	2213	561	613	7			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-15	MET	-	initiating methionine	UNP A0A076Q3N8
B	-14	ALA	-	expression tag	UNP A0A076Q3N8
B	-13	HIS	-	expression tag	UNP A0A076Q3N8
B	-12	HIS	-	expression tag	UNP A0A076Q3N8
B	-11	HIS	-	expression tag	UNP A0A076Q3N8
B	-10	HIS	-	expression tag	UNP A0A076Q3N8

*Continued on next page...*

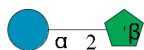
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	-9	HIS	-	expression tag	UNP A0A076Q3N8
B	-8	HIS	-	expression tag	UNP A0A076Q3N8
B	-7	ALA	-	expression tag	UNP A0A076Q3N8
B	-6	LEU	-	expression tag	UNP A0A076Q3N8
B	-5	GLU	-	expression tag	UNP A0A076Q3N8
B	-4	VAL	-	expression tag	UNP A0A076Q3N8
B	-3	LEU	-	expression tag	UNP A0A076Q3N8
B	-2	PHE	-	expression tag	UNP A0A076Q3N8
B	-1	GLN	-	expression tag	UNP A0A076Q3N8
B	0	GLY	-	expression tag	UNP A0A076Q3N8
B	172	LYS	ARG	variant	UNP A0A076Q3N8
B	280	SER	CYS	engineered mutation	UNP A0A076Q3N8
D	-15	MET	-	initiating methionine	UNP A0A076Q3N8
D	-14	ALA	-	expression tag	UNP A0A076Q3N8
D	-13	HIS	-	expression tag	UNP A0A076Q3N8
D	-12	HIS	-	expression tag	UNP A0A076Q3N8
D	-11	HIS	-	expression tag	UNP A0A076Q3N8
D	-10	HIS	-	expression tag	UNP A0A076Q3N8
D	-9	HIS	-	expression tag	UNP A0A076Q3N8
D	-8	HIS	-	expression tag	UNP A0A076Q3N8
D	-7	ALA	-	expression tag	UNP A0A076Q3N8
D	-6	LEU	-	expression tag	UNP A0A076Q3N8
D	-5	GLU	-	expression tag	UNP A0A076Q3N8
D	-4	VAL	-	expression tag	UNP A0A076Q3N8
D	-3	LEU	-	expression tag	UNP A0A076Q3N8
D	-2	PHE	-	expression tag	UNP A0A076Q3N8
D	-1	GLN	-	expression tag	UNP A0A076Q3N8
D	0	GLY	-	expression tag	UNP A0A076Q3N8
D	172	LYS	ARG	variant	UNP A0A076Q3N8
D	280	SER	CYS	engineered mutation	UNP A0A076Q3N8

- Molecule 3 is DNA/RNA hybrid called 38-MER RNA-DNA Aptamer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	35	Total	C	N	O	P	0	0	0
			726	331	130	230	35			
3	F	35	Total	C	N	O	P	0	0	0
			726	331	130	230	35			

- Molecule 4 is an oligosaccharide called beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose.

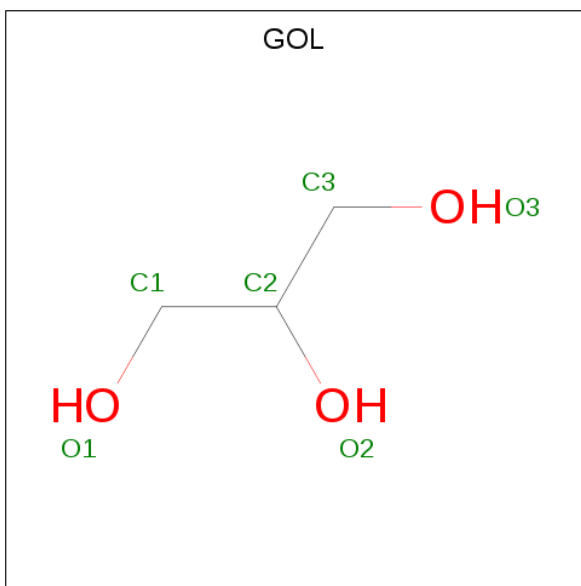


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	G	2	Total	C	O	0	0	0
			23	12	11			
4	H	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg	0	0
			1	1		
5	C	1	Total	Mg	0	0
			1	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



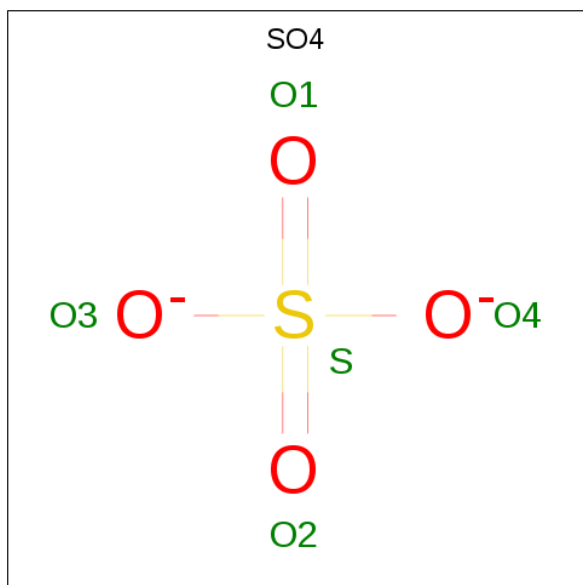
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			6	3	3		
6	B	1	Total	C	O	0	0
			6	3	3		
6	B	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			6	3	3		
6	C	1	Total	C	O	0	0
			6	3	3		
6	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

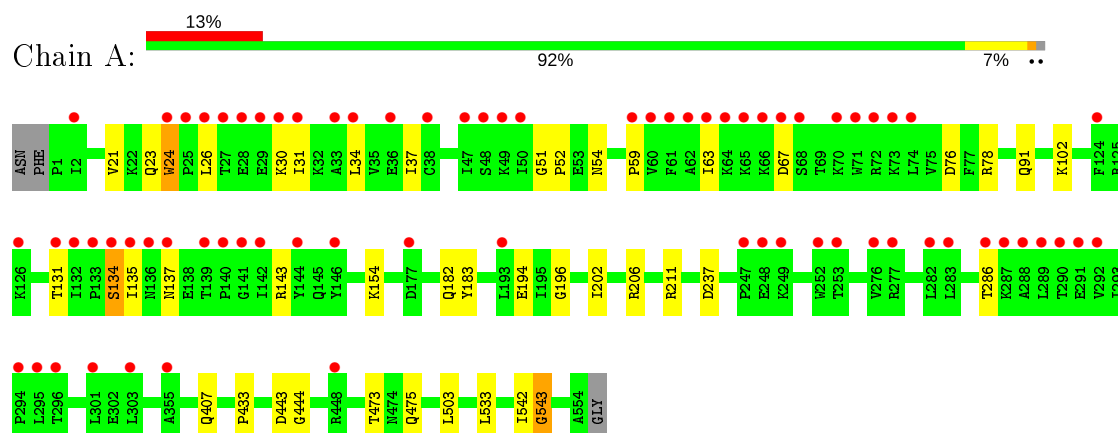
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	22	Total	O	0	0
			22	22		
8	B	43	Total	O	0	0
			43	43		
8	C	25	Total	O	0	0
			25	25		
8	D	20	Total	O	0	0
			20	20		
8	E	1	Total	O	0	0
			1	1		

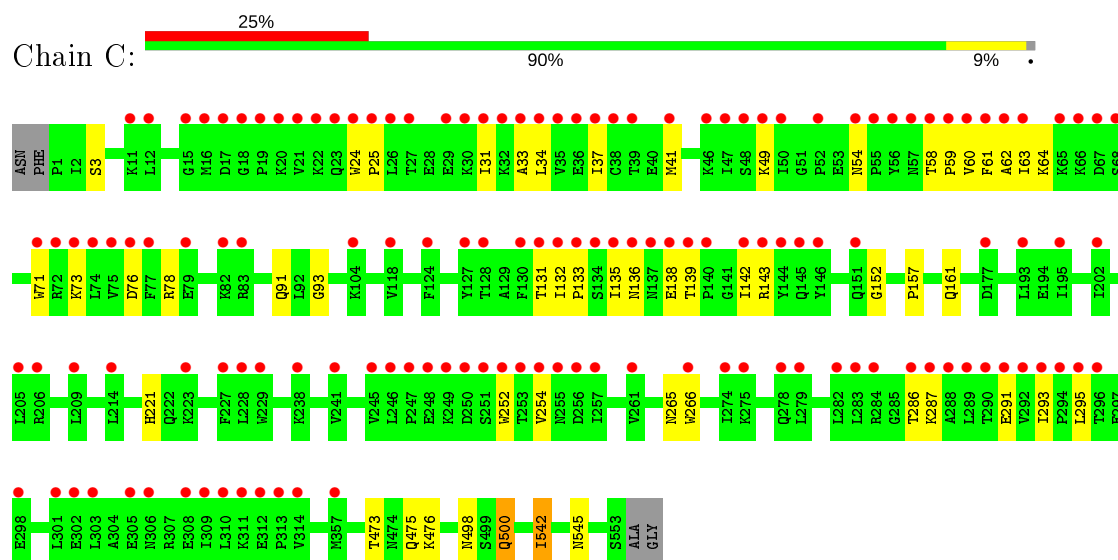
### 3 Residue-property plots [i](#)

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.

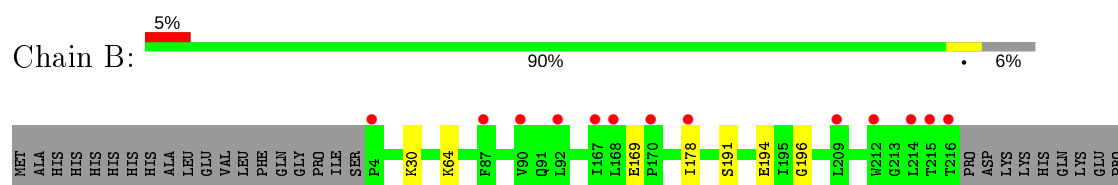
#### • Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 subunit



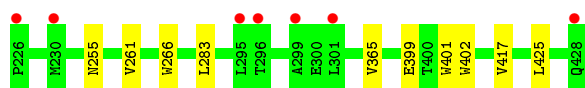
#### • Molecule 1: HIV-1 REVERSE TRANSCRIPTASE P66 subunit



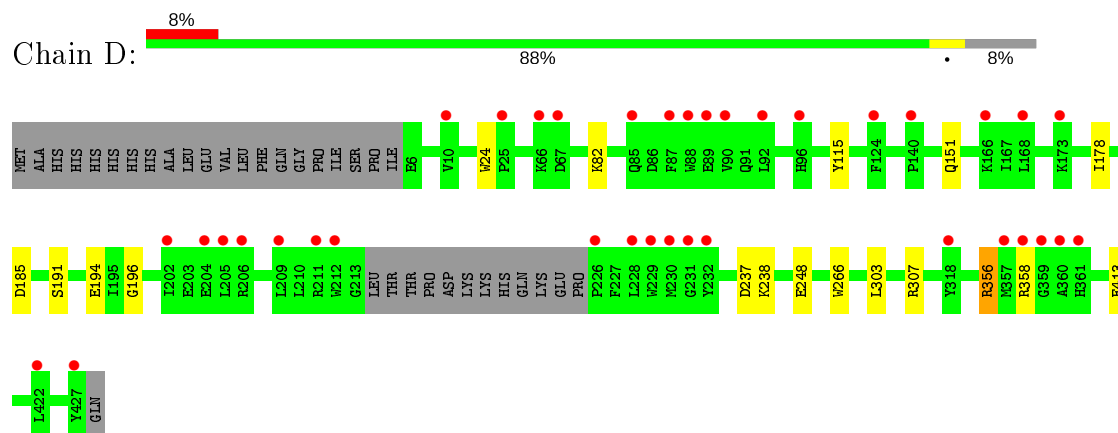
#### • Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 subunit



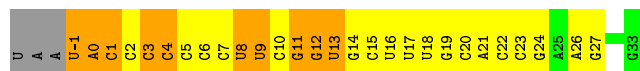
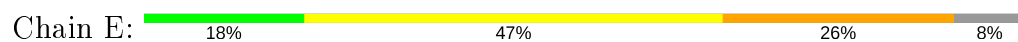




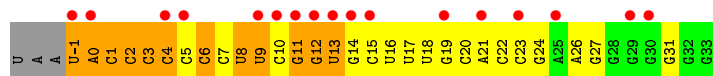
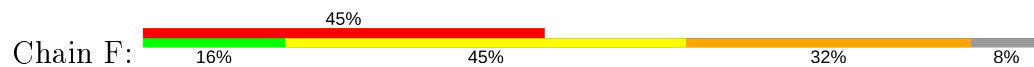
- Molecule 2: HIV-1 REVERSE TRANSCRIPTASE P51 subunit



- Molecule 3: 38-MER RNA-DNA Aptamer



- Molecule 3: 38-MER RNA-DNA Aptamer



- Molecule 4: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



- Molecule 4: beta-D-fructofuranose-(2-1)-alpha-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	90.06Å 127.46Å 132.12Å 90.00° 101.81° 90.00°	Depositor
Resolution (Å)	47.69 – 2.81 47.69 – 2.81	Depositor EDS
% Data completeness (in resolution range)	92.4 (47.69-2.81) 92.4 (47.69-2.81)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 2.81Å)	Xtriage
Refinement program	PHENIX (1.10.1 _2155: ???)	Depositor
R, $R_{free}$	0.203 , 0.244 0.203 , 0.245	Depositor DCC
$R_{free}$ test set	2679 reflections (4.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	71.9	Xtriage
Anisotropy	0.284	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 79.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	17498	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.09% 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: GOL, MG, GLC, SO4, FRU

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	0.26	0/4627	0.44	1/6286 (0.0%)
1	C	0.26	0/4622	0.45	1/6279 (0.0%)
2	B	0.27	0/3539	0.43	0/4807
2	D	0.25	0/3493	0.42	0/4745
3	E	1.41	10/811 (1.2%)	1.69	27/1254 (2.2%)
3	F	1.42	10/811 (1.2%)	1.67	26/1254 (2.1%)
All	All	0.49	20/17903 (0.1%)	0.68	55/24625 (0.2%)

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	D	0	1
3	E	0	1
3	F	0	1
All	All	0	3

The worst 5 of 20 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	13	U	O5'-C5'	-15.95	1.17	1.42
3	F	13	U	O5'-C5'	-15.93	1.17	1.42
3	F	12	G	O3'-P	-11.38	1.47	1.61
3	E	12	G	O3'-P	-11.22	1.47	1.61
3	F	13	U	C3'-O3'	10.44	1.56	1.42

The worst 5 of 55 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	13	U	C5'-C4'-C3'	-20.02	83.97	116.00
3	E	13	U	C5'-C4'-C3'	-19.99	84.02	116.00
3	E	9	U	C4'-C3'-O3'	16.47	145.94	113.00
3	F	9	U	C4'-C3'-O3'	16.45	145.91	113.00
3	F	13	U	C4'-C3'-O3'	11.15	135.31	113.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	D	356	ARG	Peptide
3	E	13	U	Sidechain
3	F	13	U	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	4509	0	4572	63	1
1	C	4504	0	4565	89	1
2	B	3439	0	3468	25	0
2	D	3394	0	3421	29	0
3	E	726	0	379	43	0
3	F	726	0	379	71	0
4	G	23	0	21	0	0
4	H	23	0	21	6	0
5	A	1	0	0	0	0
5	C	1	0	0	0	0
6	B	24	0	32	11	0
6	C	6	0	8	0	0
6	D	6	0	8	6	0
7	C	5	0	0	0	0
8	A	22	0	0	0	0
8	B	43	0	0	0	0
8	C	25	0	0	0	0
8	D	20	0	0	0	0
8	E	1	0	0	0	0
All	All	17498	0	16874	283	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 9.

The worst 5 of 283 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:F:19:DG:C2'	3:F:20:DC:H5'	1.06	1.51
1:A:63:ILE:HD11	3:E:-1:U:P	1.51	1.49
1:A:30:LYS:O	1:A:34:LEU:HD13	1.26	1.33
3:F:19:DG:H2''	3:F:20:DC:C5'	0.84	1.32
1:A:30:LYS:O	1:A:34:LEU:CD1	1.85	1.25

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:211:ARG:NH1	1:C:3:SER:OG[2_455]	1.25	0.95

## 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	552/557 (99%)	527 (96%)	23 (4%)	2 (0%)	34	64
1	C	551/557 (99%)	533 (97%)	18 (3%)	0	100	100
2	B	412/444 (93%)	403 (98%)	9 (2%)	0	100	100
2	D	406/444 (91%)	397 (98%)	9 (2%)	0	100	100
All	All	1921/2002 (96%)	1860 (97%)	59 (3%)	2 (0%)	51	80

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	134	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	543	GLY

### 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	495/497 (100%)	493 (100%)	2 (0%)	91	97
1	C	495/497 (100%)	493 (100%)	2 (0%)	91	97
2	B	377/403 (94%)	376 (100%)	1 (0%)	92	97
2	D	372/403 (92%)	371 (100%)	1 (0%)	92	97
All	All	1739/1800 (97%)	1733 (100%)	6 (0%)	92	97

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

Mol	Chain	Res	Type
2	B	169	GLU
2	D	266	TRP
1	C	221	HIS
1	A	182	GLN
1	C	500	GLN

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

Mol	Chain	Res	Type
1	A	23	GLN
2	B	255	ASN
1	C	161	GLN
1	C	265	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 ⓘ

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	GLC	G	1	4	11,11,12	1.08	1 (9%)	15,15,17	2.16	5 (33%)
4	FRU	G	2	4	11,12,12	0.58	0	10,18,18	0.94	0
4	GLC	H	1	4	11,11,12	0.70	0	15,15,17	1.25	2 (13%)
4	FRU	H	2	4	11,12,12	0.64	0	10,18,18	1.83	3 (30%)

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	GLC	G	1	4	-	0/2/19/22	0/1/1/1
4	FRU	G	2	4	-	1/5/24/24	0/1/1/1
4	GLC	H	1	4	-	0/2/19/22	0/1/1/1
4	FRU	H	2	4	-	2/5/24/24	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	1	GLC	C2-C3	-2.19	1.49	1.52

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1	GLC	C2-C3-C4	-5.04	102.17	110.89

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	1	GLC	C1-O5-C5	4.18	117.86	112.19
4	H	2	FRU	O1-C1-C2	-3.35	104.73	111.86
4	G	1	GLC	C3-C4-C5	-3.18	104.57	110.24
4	H	2	FRU	O4-C4-C3	-2.80	103.78	112.15

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	H	2	FRU	C4-C5-C6-O6
4	H	2	FRU	O5-C5-C6-O6
4	G	2	FRU	O1-C1-C2-C3

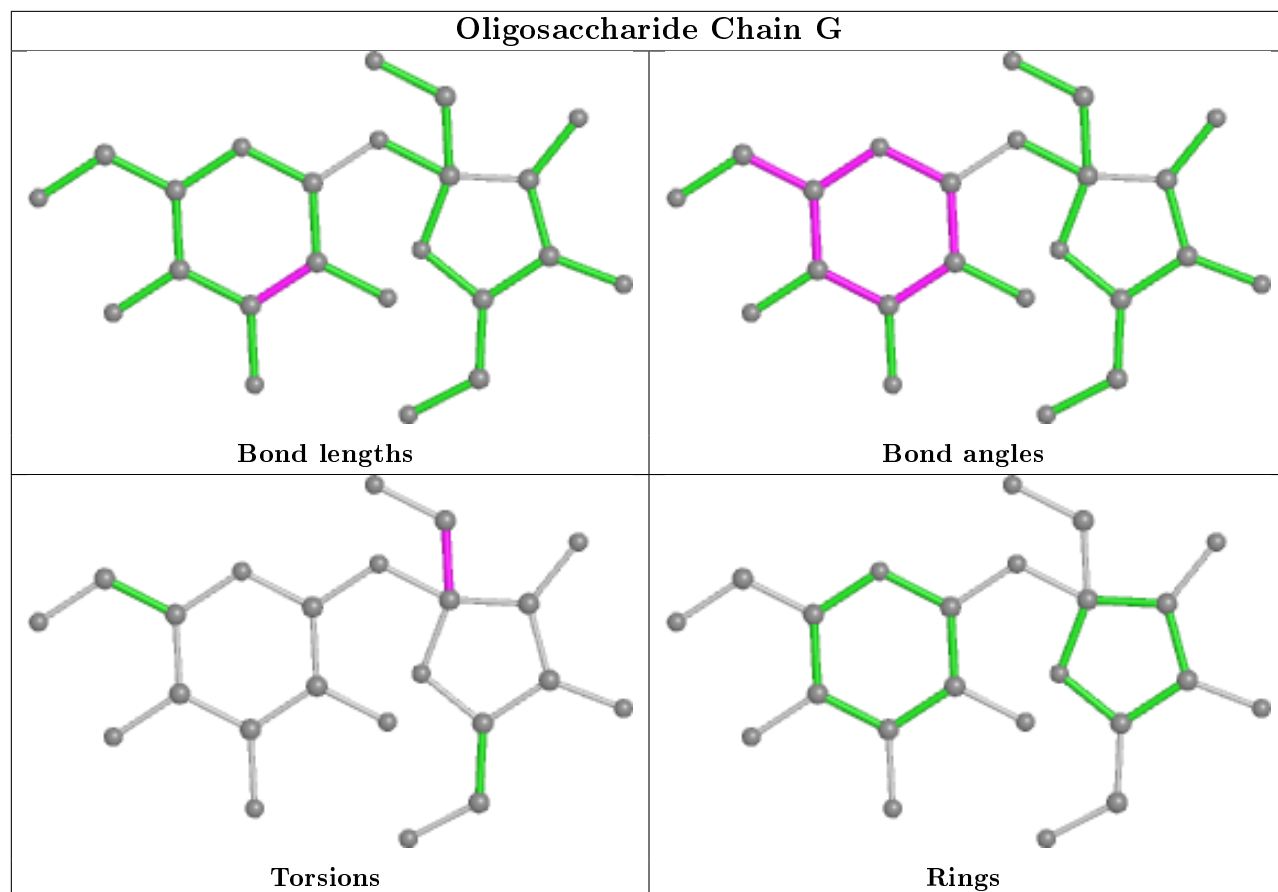
There are no ring outliers.

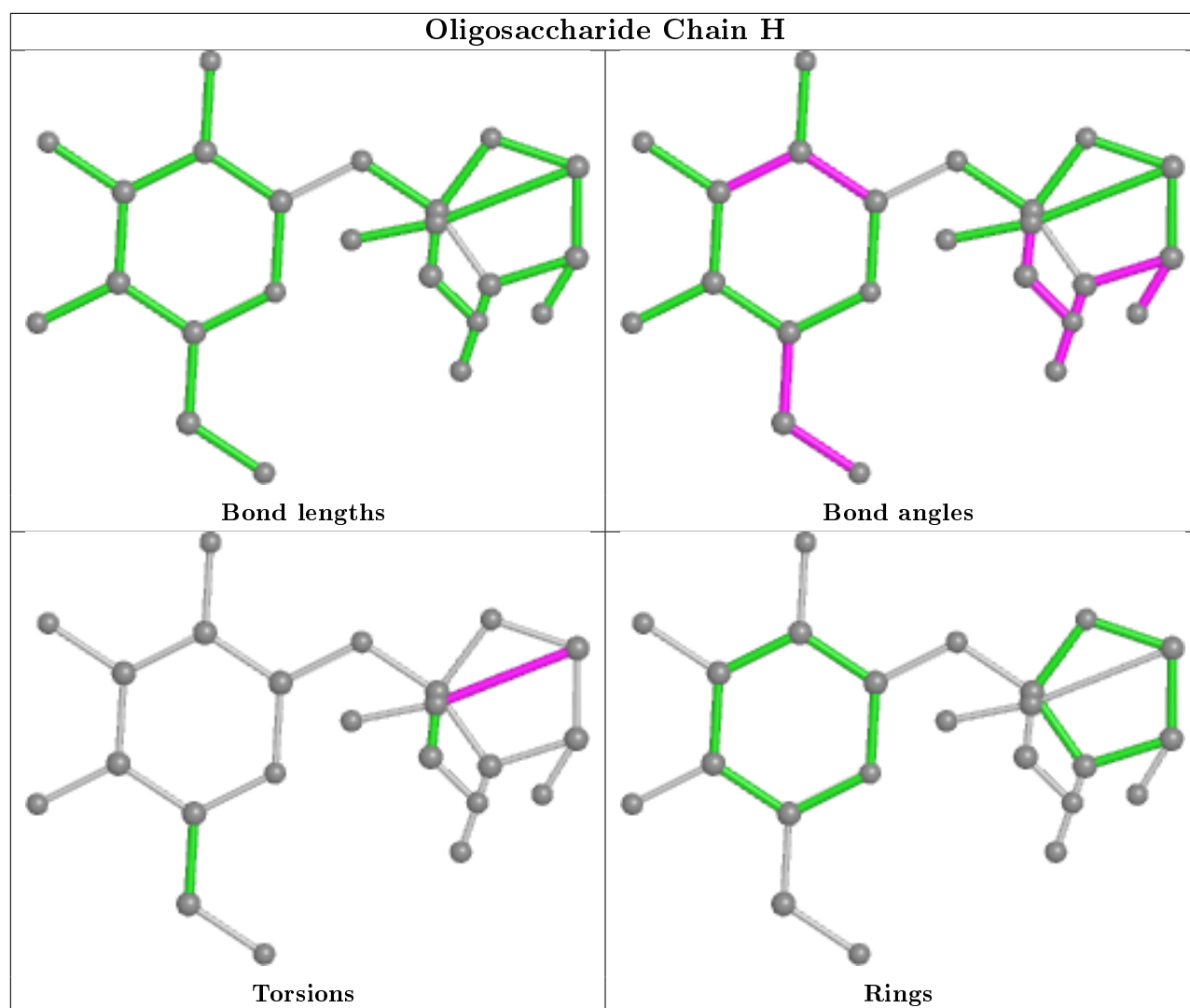
2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	H	1	GLC	5	0
4	H	2	FRU	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 9 ligands modelled in this entry, 2 are monoatomic - leaving 7 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	$\# Z  > 2$
6	GOL	B	2003	-	5,5,5	0.27	0	5,5,5	0.65	0
6	GOL	B	2004	-	5,5,5	0.38	0	5,5,5	0.24	0
6	GOL	C	602	-	5,5,5	0.17	0	5,5,5	0.47	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
6	GOL	B	2005	-	5,5,5	0.11	0	5,5,5	0.68	0
6	GOL	D	502	-	5,5,5	0.23	0	5,5,5	0.26	0
6	GOL	B	2002	-	5,5,5	0.30	0	5,5,5	0.72	0
7	SO4	C	601	-	4,4,4	0.26	0	6,6,6	0.18	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
6	GOL	B	2003	-	-	2/4/4/4	-
6	GOL	B	2004	-	-	0/4/4/4	-
6	GOL	C	602	-	-	0/4/4/4	-
6	GOL	B	2005	-	-	0/4/4/4	-
6	GOL	D	502	-	-	2/4/4/4	-
6	GOL	B	2002	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	2003	GOL	C1-C2-C3-O3
6	D	502	GOL	O1-C1-C2-C3
6	D	502	GOL	O1-C1-C2-O2
6	B	2002	GOL	O1-C1-C2-C3
6	B	2003	GOL	O2-C2-C3-O3

There are no ring outliers.

3 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	2003	GOL	9	0
6	B	2005	GOL	2	0
6	D	502	GOL	6	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	554/557 (99%)	0.65	72 (12%) <b>3</b> <b>2</b>	46, 96, 199, 268	0
1	C	553/557 (99%)	1.26	140 (25%) <b>0</b> <b>0</b>	41, 133, 228, 281	0
2	B	416/444 (93%)	0.21	20 (4%) <b>30</b> <b>21</b>	41, 78, 149, 224	0
2	D	410/444 (92%)	0.39	37 (9%) <b>9</b> <b>5</b>	43, 101, 184, 262	0
3	E	32/38 (84%)	0.28	0 <b>100</b> <b>100</b>	110, 159, 204, 220	0
3	F	32/38 (84%)	2.07	17 (53%) <b>0</b> <b>0</b>	205, 240, 283, 298	0
All	All	1997/2078 (96%)	0.69	286 (14%) <b>2</b> <b>1</b>	41, 99, 217, 298	0

The worst 5 of 286 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	62	ALA	15.8
1	A	67	ASP	14.0
1	C	26	LEU	11.9
1	C	290	THR	10.3
1	C	61	PHE	10.3

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

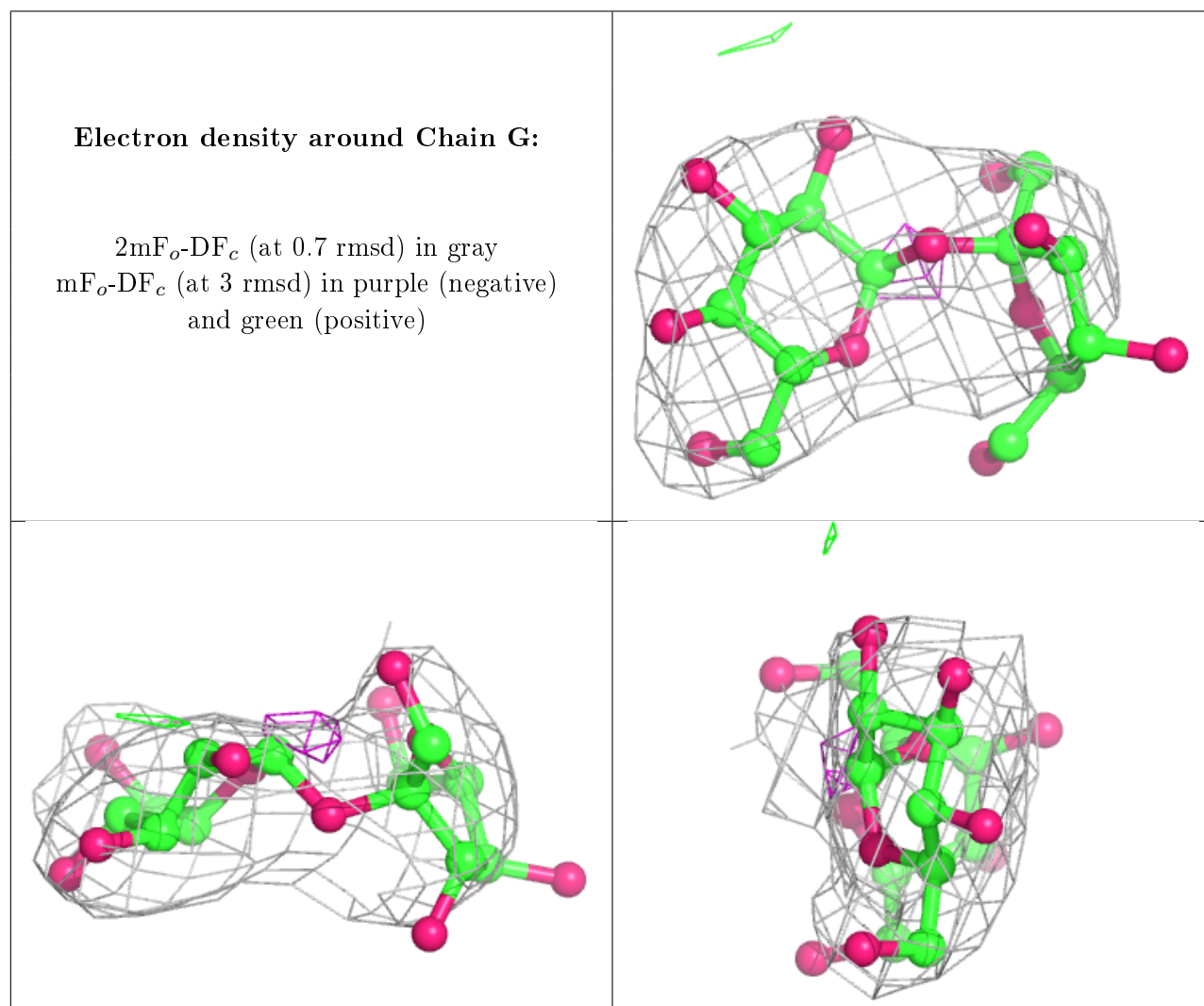
There are no non-standard protein/DNA/RNA residues in this entry.

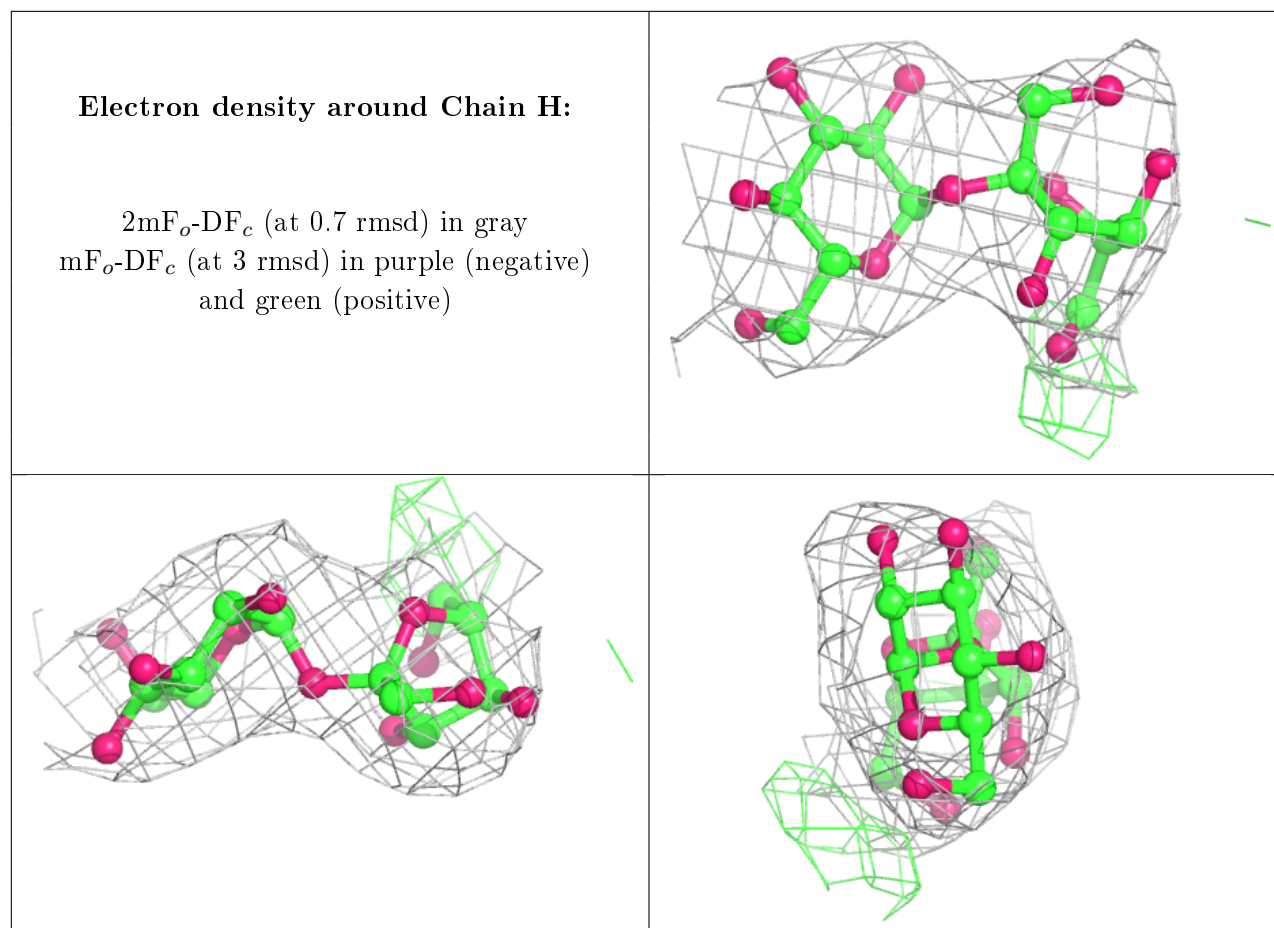
### 6.3 Carbohydrates [i](#)

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( $\text{\AA}^2$ )	Q<0.9
4	FRU	G	2	12/12	0.84	0.26	151,163,168,169	0
4	GLC	H	1	11/12	0.88	0.19	113,116,124,125	0
4	FRU	H	2	12/12	0.89	0.22	136,147,152,152	0
4	GLC	G	1	11/12	0.94	0.25	87,116,134,142	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 ⓘ

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
7	SO4	C	601	5/5	0.10	0.43	207,208,210,212	0
6	GOL	D	502	6/6	0.66	0.47	92,102,107,109	0
6	GOL	C	602	6/6	0.77	0.57	98,108,110,113	0
6	GOL	B	2005	6/6	0.81	0.33	88,92,94,96	0
6	GOL	B	2004	6/6	0.83	0.31	77,81,86,87	0
6	GOL	B	2003	6/6	0.88	0.28	69,72,75,80	0
6	GOL	B	2002	6/6	0.89	0.20	68,73,79,80	0
5	MG	A	601	1/1	0.95	0.27	81,81,81,81	0
5	MG	C	603	1/1	0.95	0.38	77,77,77,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.