



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 7, 2020 – 09:06 PM BST

PDB ID : 5KDX  
Title : IMPa metallopeptidase in complex with T-antigen  
Authors : Noach, I.; Boraston, A.B.  
Deposited on : 2016-06-08  
Resolution : 2.40 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

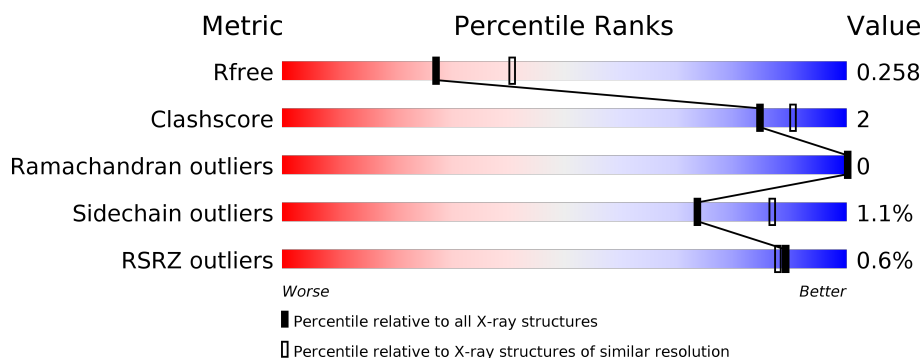
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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	885	<div> <div>%</div> <div> <div></div> <div>94%</div> <div>6%</div> </div> </div>
1	B	885	<div> <div>%</div> <div> <div></div> <div>95%</div> <div>.</div> </div> </div>
2	C	2	<div> <div></div> <div>100%</div> </div>
2	D	2	<div> <div></div> <div>100%</div> </div>

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	EDO	B	1008	-	-	X	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 15077 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.

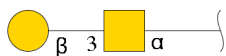
- Molecule 1 is a protein called Metallopeptidase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	884	Total	C	N	O	S	0	0	0
			6789	4250	1214	1311	14			
1	B	883	Total	C	N	O	S	0	0	0
			6780	4245	1213	1308	14			

There are 6 discrepancies between the modelled and reference sequences:

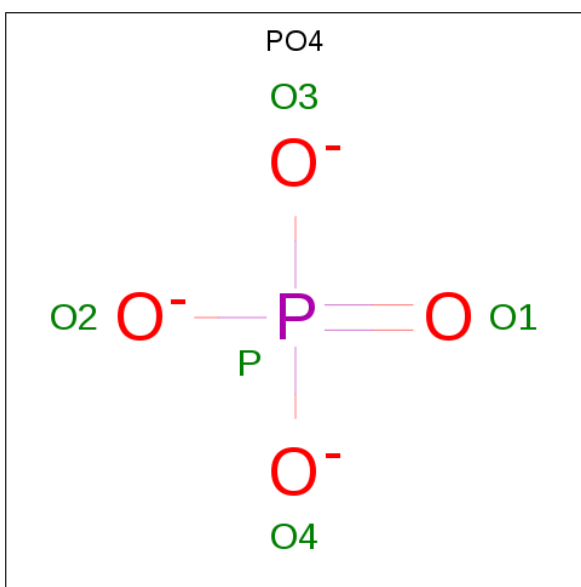
Chain	Residue	Modelled	Actual	Comment	Reference
A	41	MET	-	initiating methionine	UNP Q9I5W4
A	924	LEU	-	expression tag	UNP Q9I5W4
A	925	GLU	-	expression tag	UNP Q9I5W4
B	41	MET	-	initiating methionine	UNP Q9I5W4
B	924	LEU	-	expression tag	UNP Q9I5W4
B	925	GLU	-	expression tag	UNP Q9I5W4

- Molecule 2 is an oligosaccharide called beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-alpha-D-galactopyranose.



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	2	Total	C	N	O	0	0	0
			25	14	1	10			
2	D	2	Total	C	N	O	0	0	0
			25	14	1	10			

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	A	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		
3	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Zn	0	0
			1	1		
4	A	1	Total	Zn	0	0
			1	1		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



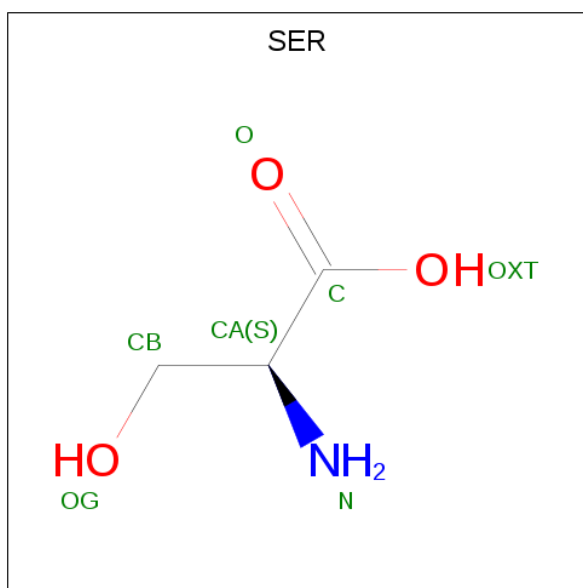
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		
5	A	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		
5	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is SERINE (three-letter code: SER) (formula:  $C_3H_7NO_3$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	N	O	0	0
			7	3	1	3		
6	B	1	Total	C	N	O	0	0
			7	3	1	3		

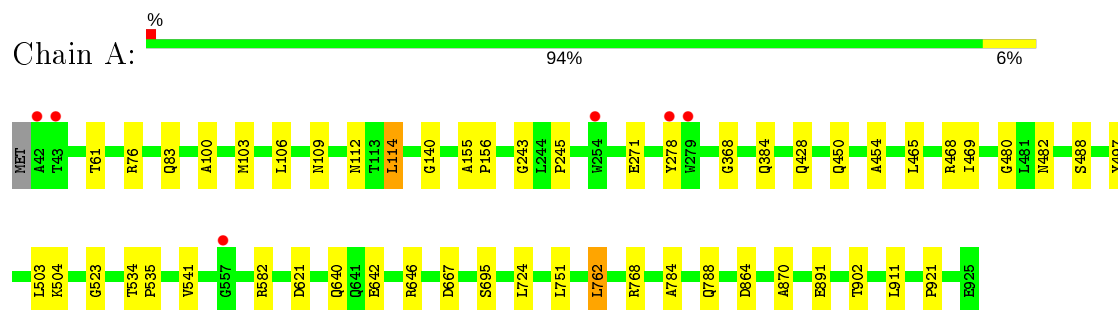
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	657	Total	O	0	0
			657	657		
7	B	656	Total	O	0	0
			656	656		

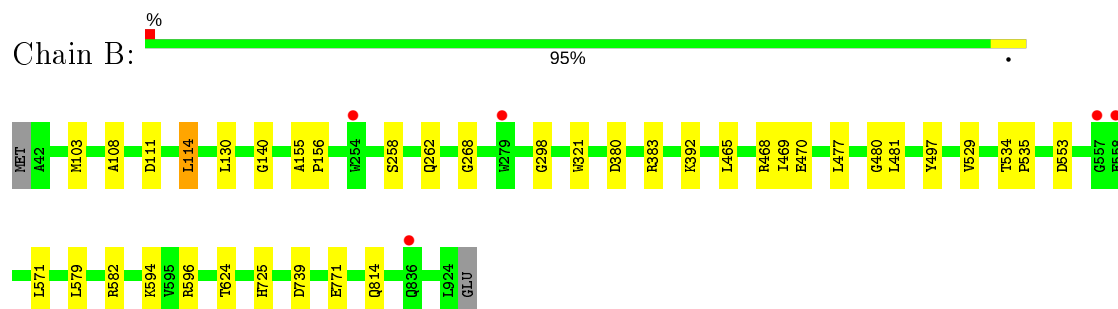
### 3 Residue-property plots [i](#)

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.

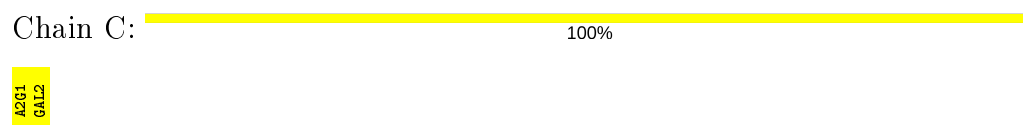
- Molecule 1: Metallopeptidase



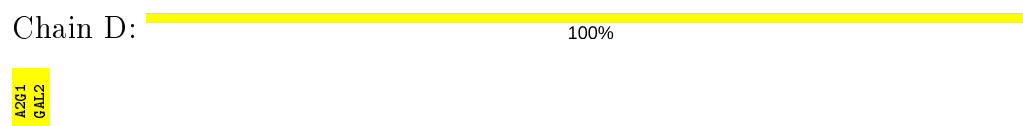
- Molecule 1: Metallopeptidase



- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-alpha-D-galactopyranose



- Molecule 2: beta-D-galactopyranose-(1-3)-2-acetamido-2-deoxy-alpha-D-galactopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.41Å 133.10Å 103.78Å 90.00° 103.12° 90.00°	Depositor
Resolution (Å)	48.12 – 2.40 48.12 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.12-2.40) 99.6 (48.12-2.40)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.54 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, $R_{free}$	0.201 , 0.254 0.208 , 0.258	Depositor DCC
$R_{free}$ test set	5027 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.2	Xtriage
Anisotropy	0.179	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 40.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.090 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	15077	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.20% 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: A2G, ZN, GAL, PO4, EDO

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	0.48	0/6937	0.66	2/9442 (0.0%)
1	B	0.48	0/6928	0.67	1/9430 (0.0%)
All	All	0.48	0/13865	0.66	3/18872 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	114	LEU	CA-CB-CG	6.32	129.84	115.30
1	A	667	ASP	CB-CG-OD1	5.20	122.98	118.30
1	B	114	LEU	CA-CB-CG	5.06	126.94	115.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6789	0	6592	36	0
1	B	6780	0	6586	26	0
2	C	25	0	21	0	0
2	D	25	0	21	0	0
3	A	10	0	0	0	0
3	B	15	0	0	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	60	0	90	10	0
5	B	44	0	66	9	0
6	A	7	0	3	0	0
6	B	7	0	3	0	0
7	A	657	0	0	4	0
7	B	656	0	0	4	0
All	All	15077	0	13382	63	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 2.

All (63) 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:B:468:ARG:NH2	1:B:470:GLU:OE2	2.11	0.83
1:A:61:THR:HG22	1:A:384:GLN:HG2	1.70	0.74
1:B:298:GLY:O	7:B:1101:HOH:O	2.07	0.72
1:B:469:ILE:CD1	1:B:477:LEU:HD13	2.22	0.70
1:A:454:ALA:O	1:A:582:ARG:NH1	2.25	0.69
1:A:480:GLY:HA3	5:A:1011:EDO:C1	2.31	0.61
1:B:469:ILE:HD12	1:B:477:LEU:HD13	1.83	0.60
1:A:468:ARG:HD3	7:A:1177:HOH:O	2.00	0.59
1:A:902:THR:HG22	5:A:1006:EDO:H22	1.84	0.57
1:B:480:GLY:HA3	5:B:1009:EDO:H12	1.86	0.57
1:B:380:ASP:OD1	1:B:383:ARG:NH1	2.32	0.57
1:A:428:GLN:H	1:A:428:GLN:CD	2.08	0.56
1:A:469:ILE:HG22	1:A:541:VAL:HG22	1.88	0.56
1:A:621:ASP:OD2	5:A:1008:EDO:H12	2.07	0.55
1:A:724:LEU:HB3	5:A:1006:EDO:H11	1.88	0.54
1:B:571:LEU:O	1:B:596:ARG:NH2	2.43	0.52
1:A:155:ALA:HB3	1:A:156:PRO:HD3	1.92	0.52
1:B:321:TRP:CE2	1:B:392:LYS:HG2	2.44	0.52
1:A:100:ALA:HB1	1:A:278:TYR:HA	1.91	0.52
1:B:579:LEU:HD12	1:B:579:LEU:C	2.31	0.52
5:B:1009:EDO:H11	7:B:1212:HOH:O	2.10	0.51
1:A:469:ILE:HD12	1:A:469:ILE:O	2.11	0.51
1:A:106:LEU:HD12	1:A:106:LEU:N	2.25	0.50
1:A:465:LEU:C	1:A:465:LEU:HD12	2.32	0.50
1:A:480:GLY:HA3	5:A:1011:EDO:H12	1.95	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:112:ASN:ND2	7:A:1121:HOH:O	2.46	0.48
1:A:762:LEU:CD2	1:A:870:ALA:HA	2.43	0.48
1:A:523:GLY:O	5:A:1009:EDO:O1	2.28	0.47
1:B:481:LEU:CB	5:B:1008:EDO:H12	2.45	0.47
1:A:103:MET:HE2	1:A:271:GLU:HG2	1.96	0.47
1:A:482:ASN:O	1:A:504:LYS:HG3	2.14	0.47
1:B:624:THR:HG21	5:B:1010:EDO:H11	1.97	0.47
1:B:108:ALA:HB3	1:B:268:GLY:O	2.15	0.46
1:A:243:GLY:O	1:A:245:PRO:HD3	2.16	0.46
1:A:695:SER:HG	5:A:1008:EDO:C2	2.29	0.45
1:A:109:ASN:ND2	1:A:368:GLY:O	2.47	0.45
1:A:488:SER:HB2	5:A:1007:EDO:H11	1.99	0.45
1:A:642:GLU:HG2	7:A:1297:HOH:O	2.17	0.44
1:B:624:THR:HG21	5:B:1010:EDO:C1	2.48	0.44
1:B:155:ALA:HB3	1:B:156:PRO:HD3	1.99	0.43
1:A:864:ASP:HB3	1:A:891:GLU:HA	2.01	0.43
1:A:534:THR:O	1:A:535:PRO:C	2.57	0.43
1:A:103:MET:O	1:A:140:GLY:HA3	2.19	0.43
1:B:258:SER:O	1:B:262:GLN:HG3	2.19	0.42
1:A:428:GLN:CD	1:A:428:GLN:N	2.73	0.42
1:B:321:TRP:CD2	1:B:392:LYS:HG2	2.55	0.42
1:B:111:ASP:OD1	3:B:1003:PO4:O4	2.37	0.41
1:B:481:LEU:HB3	5:B:1008:EDO:H11	2.02	0.41
1:B:534:THR:O	1:B:535:PRO:C	2.57	0.41
1:A:762:LEU:HD22	1:A:870:ALA:HA	2.02	0.41
1:B:725:HIS:NE2	7:B:1106:HOH:O	2.36	0.41
1:B:553:ASP:OD1	1:B:582:ARG:HD2	2.21	0.41
1:A:621:ASP:OD2	5:A:1008:EDO:C1	2.69	0.41
1:A:784:ALA:O	1:A:788:GLN:HG2	2.20	0.41
1:B:481:LEU:HB2	5:B:1008:EDO:H12	2.02	0.41
1:A:503:LEU:O	5:A:1009:EDO:C2	2.69	0.41
1:A:911:LEU:HD21	1:A:921:PRO:HG3	2.04	0.40
1:B:481:LEU:HB3	5:B:1008:EDO:C1	2.50	0.40
1:B:529:VAL:HG12	5:B:1014:EDO:H11	2.03	0.40
1:A:751:LEU:HD22	1:A:768:ARG:HB3	2.04	0.40
1:B:103:MET:O	1:B:140:GLY:HA3	2.21	0.40
1:A:642:GLU:CG	7:A:1297:HOH:O	2.69	0.40
1:B:594:LYS:NZ	7:B:1149:HOH:O	2.54	0.40

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	882/885 (100%)	864 (98%)	18 (2%)	0	100	100
1	B	881/885 (100%)	860 (98%)	21 (2%)	0	100	100
All	All	1763/1770 (100%)	1724 (98%)	39 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	700/701 (100%)	692 (99%)	8 (1%)	73	87
1	B	699/701 (100%)	692 (99%)	7 (1%)	76	88
All	All	1399/1402 (100%)	1384 (99%)	15 (1%)	73	87

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	76	ARG
1	A	83	GLN
1	A	114	LEU
1	A	450	GLN
1	A	497	TYR
1	A	640	GLN
1	A	646	ARG
1	A	762	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	114	LEU
1	B	130	LEU
1	B	465	LEU
1	B	497	TYR
1	B	739	ASP
1	B	771	GLU
1	B	814	GLN

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

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

2 non-standard protein/DNA/RNA residues 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	GAL	C	2	2	11,11,12	0.73	0	15,15,17	1.06	1 (6%)
2	GAL	D	2	2	11,11,12	0.69	0	15,15,17	1.18	2 (13%)

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
2	GAL	C	2	2	-	1/2/19/22	0/1/1/1
2	GAL	D	2	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	2	GAL	C1-C2-C3	-2.50	106.59	109.67
2	D	2	GAL	C1-C2-C3	-2.28	106.87	109.67
2	D	2	GAL	O3-C3-C2	-2.23	105.72	109.99

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	2	GAL	C4-C5-C6-O6
2	D	2	GAL	O5-C5-C6-O6
2	C	2	GAL	C4-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

## 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	$\# Z  > 2$
2	A2G	C	1	2,6	14,14,15	0.71	0	17,19,21	1.26	1 (5%)
2	GAL	C	2	2	11,11,12	0.73	0	15,15,17	1.06	1 (6%)
2	A2G	D	1	2,6	14,14,15	0.45	0	17,19,21	0.99	1 (5%)
2	GAL	D	2	2	11,11,12	0.69	0	15,15,17	1.18	2 (13%)

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
2	A2G	C	1	2,6	-	2/6/23/26	0/1/1/1
2	GAL	C	2	2	-	1/2/19/22	0/1/1/1
2	A2G	D	1	2,6	-	1/6/23/26	0/1/1/1
2	GAL	D	2	2	-	2/2/19/22	0/1/1/1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	1	A2G	O5-C5-C6	2.65	111.35	107.20
2	C	2	GAL	C1-C2-C3	-2.50	106.59	109.67
2	D	2	GAL	C1-C2-C3	-2.28	106.87	109.67
2	D	2	GAL	O3-C3-C2	-2.23	105.72	109.99
2	D	1	A2G	C1-O5-C5	2.16	115.12	112.19

There are no chirality outliers.

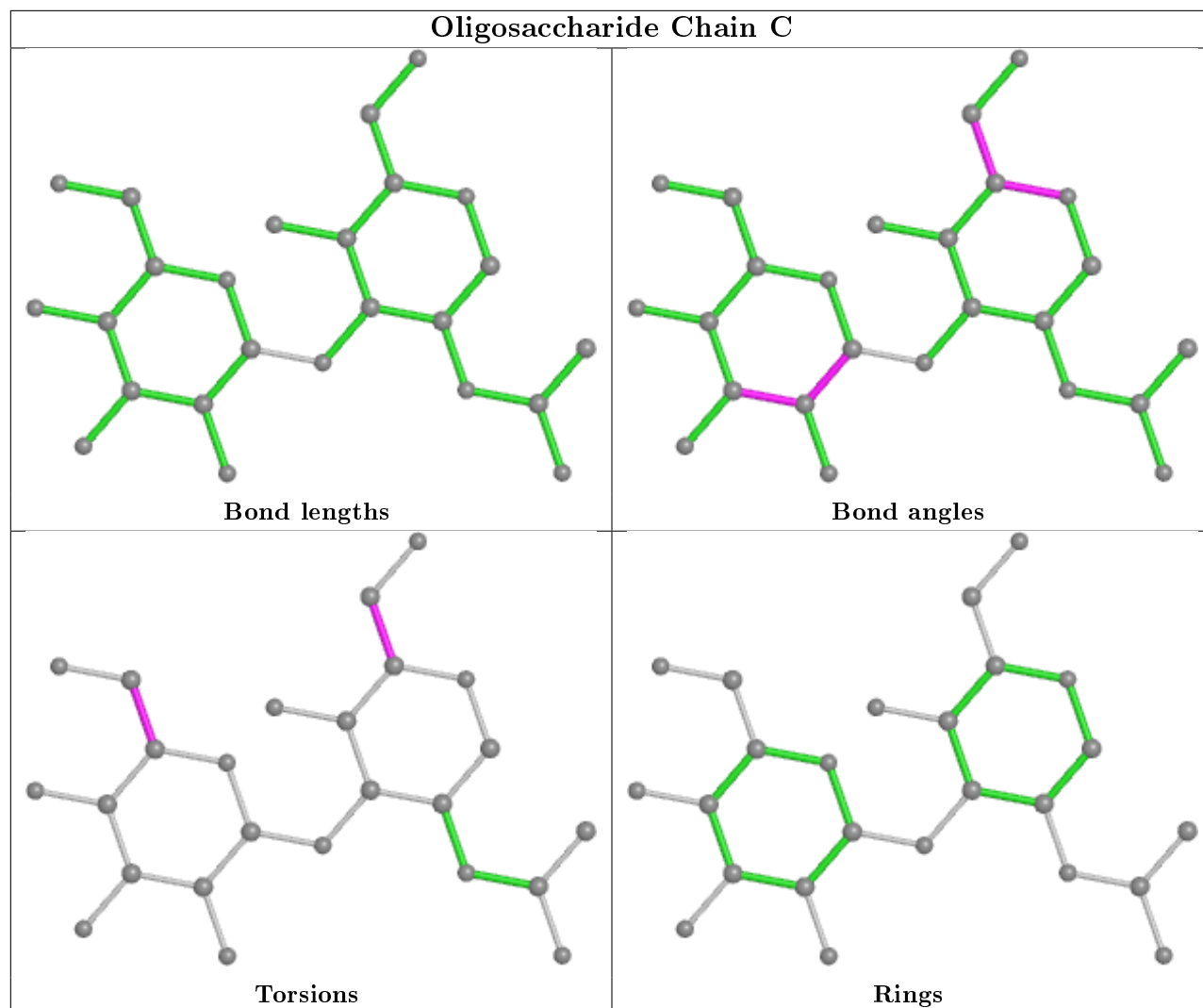
All (6) torsion outliers are listed below:

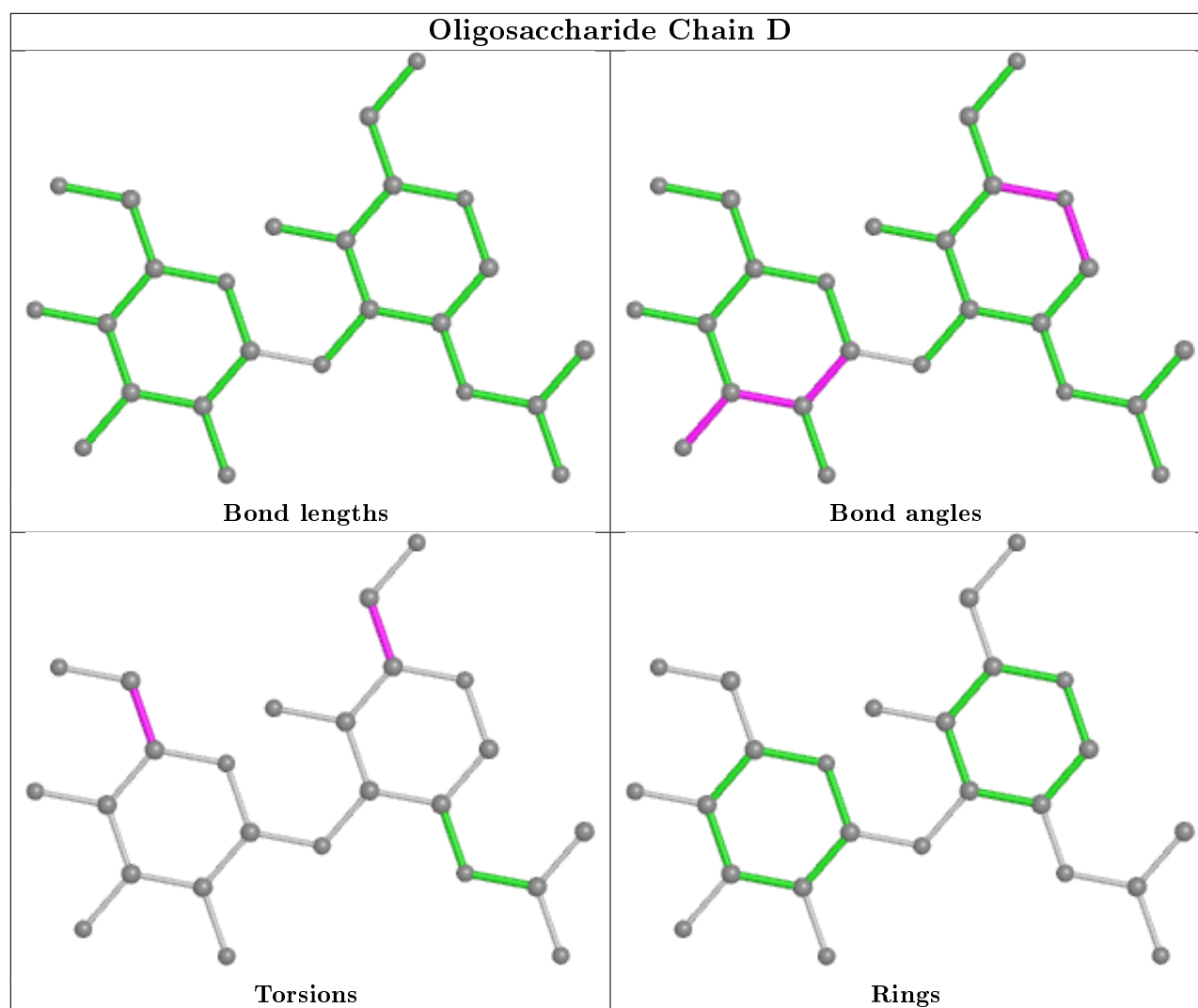
Mol	Chain	Res	Type	Atoms
2	D	2	GAL	C4-C5-C6-O6
2	D	2	GAL	O5-C5-C6-O6
2	C	2	GAL	C4-C5-C6-O6
2	C	1	A2G	C4-C5-C6-O6
2	C	1	A2G	O5-C5-C6-O6
2	D	1	A2G	C4-C5-C6-O6

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 35 ligands modelled in this entry, 2 are monoatomic - leaving 33 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	$\# Z  > 2$
5	EDO	A	1006	-	3,3,3	0.42	0	2,2,2	0.12	0
3	PO4	B	1003	-	4,4,4	0.79	0	6,6,6	0.65	0
5	EDO	A	1011	-	3,3,3	0.59	0	2,2,2	0.31	0
5	EDO	B	1007	-	3,3,3	0.44	0	2,2,2	0.33	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	EDO	A	1007	-	3,3,3	0.50	0	2,2,2	0.09	0
3	PO4	A	1004	-	4,4,4	0.74	0	6,6,6	0.64	0
5	EDO	B	1011	-	3,3,3	0.49	0	2,2,2	0.30	0
5	EDO	A	1016	-	3,3,3	0.40	0	2,2,2	0.42	0
5	EDO	A	1013	-	3,3,3	0.42	0	2,2,2	0.41	0
3	PO4	B	1005	-	4,4,4	0.76	0	6,6,6	0.80	0
5	EDO	B	1017	-	3,3,3	0.46	0	2,2,2	0.48	0
5	EDO	B	1015	-	3,3,3	0.55	0	2,2,2	0.13	0
5	EDO	A	1009	-	3,3,3	0.37	0	2,2,2	0.16	0
5	EDO	B	1016	-	3,3,3	0.52	0	2,2,2	0.08	0
5	EDO	A	1019	-	3,3,3	0.38	0	2,2,2	0.40	0
5	EDO	A	1015	-	3,3,3	0.66	0	2,2,2	0.17	0
5	EDO	A	1018	-	3,3,3	0.53	0	2,2,2	0.04	0
5	EDO	B	1014	-	3,3,3	0.45	0	2,2,2	0.29	0
5	EDO	A	1014	-	3,3,3	0.47	0	2,2,2	0.66	0
5	EDO	A	1020	-	3,3,3	0.48	0	2,2,2	0.39	0
5	EDO	B	1010	-	3,3,3	0.47	0	2,2,2	0.25	0
5	EDO	B	1008	-	3,3,3	0.26	0	2,2,2	0.37	0
3	PO4	A	1003	-	4,4,4	0.78	0	6,6,6	0.47	0
5	EDO	A	1017	-	3,3,3	0.45	0	2,2,2	0.33	0
5	EDO	A	1012	-	3,3,3	0.58	0	2,2,2	0.07	0
5	EDO	B	1012	-	3,3,3	0.39	0	2,2,2	0.44	0
5	EDO	B	1009	-	3,3,3	0.47	0	2,2,2	0.25	0
5	EDO	B	1013	-	3,3,3	0.54	0	2,2,2	0.29	0
3	PO4	B	1004	-	4,4,4	1.04	0	6,6,6	0.32	0
5	EDO	A	1008	-	3,3,3	0.46	0	2,2,2	0.33	0
5	EDO	A	1010	-	3,3,3	0.70	0	2,2,2	0.32	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	EDO	A	1006	-	-	1/1/1/1	-
5	EDO	A	1011	-	-	1/1/1/1	-
5	EDO	B	1007	-	-	1/1/1/1	-
5	EDO	A	1007	-	-	1/1/1/1	-
5	EDO	B	1011	-	-	0/1/1/1	-
5	EDO	A	1016	-	-	1/1/1/1	-
5	EDO	A	1013	-	-	1/1/1/1	-
5	EDO	B	1017	-	-	0/1/1/1	-
5	EDO	B	1015	-	-	1/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	EDO	A	1009	-	-	0/1/1/1	-
5	EDO	B	1016	-	-	0/1/1/1	-
5	EDO	A	1019	-	-	0/1/1/1	-
5	EDO	A	1015	-	-	0/1/1/1	-
5	EDO	A	1018	-	-	1/1/1/1	-
5	EDO	B	1014	-	-	1/1/1/1	-
5	EDO	A	1014	-	-	1/1/1/1	-
5	EDO	A	1020	-	-	1/1/1/1	-
5	EDO	B	1010	-	-	0/1/1/1	-
5	EDO	B	1008	-	-	0/1/1/1	-
5	EDO	A	1017	-	-	0/1/1/1	-
5	EDO	A	1012	-	-	0/1/1/1	-
5	EDO	B	1012	-	-	1/1/1/1	-
5	EDO	B	1009	-	-	0/1/1/1	-
5	EDO	B	1013	-	-	1/1/1/1	-
5	EDO	A	1008	-	-	1/1/1/1	-
5	EDO	A	1010	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	1007	EDO	O1-C1-C2-O2
5	A	1016	EDO	O1-C1-C2-O2
5	B	1015	EDO	O1-C1-C2-O2
5	B	1014	EDO	O1-C1-C2-O2
5	B	1012	EDO	O1-C1-C2-O2
5	A	1008	EDO	O1-C1-C2-O2
5	A	1011	EDO	O1-C1-C2-O2
5	B	1007	EDO	O1-C1-C2-O2
5	A	1020	EDO	O1-C1-C2-O2
5	A	1018	EDO	O1-C1-C2-O2
5	B	1013	EDO	O1-C1-C2-O2
5	A	1006	EDO	O1-C1-C2-O2
5	A	1013	EDO	O1-C1-C2-O2
5	A	1014	EDO	O1-C1-C2-O2

There are no ring outliers.

10 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1006	EDO	2	0
3	B	1003	PO4	1	0
5	A	1011	EDO	2	0
5	A	1007	EDO	1	0
5	A	1009	EDO	2	0
5	B	1014	EDO	1	0
5	B	1010	EDO	2	0
5	B	1008	EDO	4	0
5	B	1009	EDO	2	0
5	A	1008	EDO	3	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

### 6.1 Protein, DNA and RNA chains

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	884/885 (99%)	-0.23	6 (0%) 87 86	16, 26, 40, 73	0
1	B	883/885 (99%)	-0.26	5 (0%) 89 88	16, 25, 40, 68	0
All	All	1767/1770 (99%)	-0.24	11 (0%) 89 88	16, 26, 40, 73	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	254	TRP	2.9
1	B	279	TRP	2.8
1	A	279	TRP	2.8
1	A	557	GLY	2.7
1	A	278	TYR	2.6
1	A	42	ALA	2.6
1	B	254	TRP	2.4
1	B	557	GLY	2.3
1	B	836	GLN	2.1
1	A	43	THR	2.0
1	B	558	GLU	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains

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	GAL	D	2	11/12	0.74	0.28	42,45,47,47	11
2	GAL	C	2	11/12	0.86	0.21	40,43,45,46	11

### 6.3 Carbohydrates

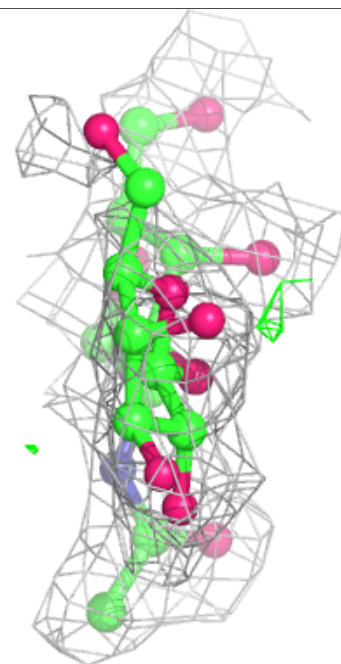
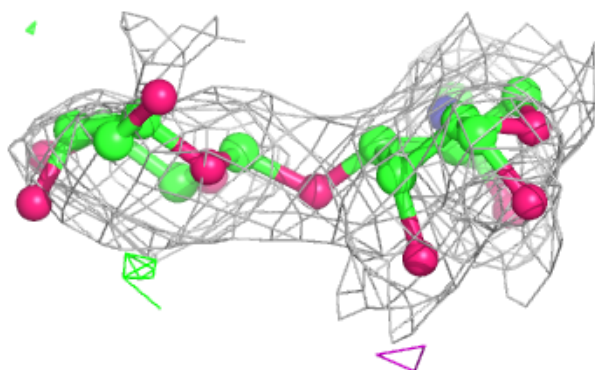
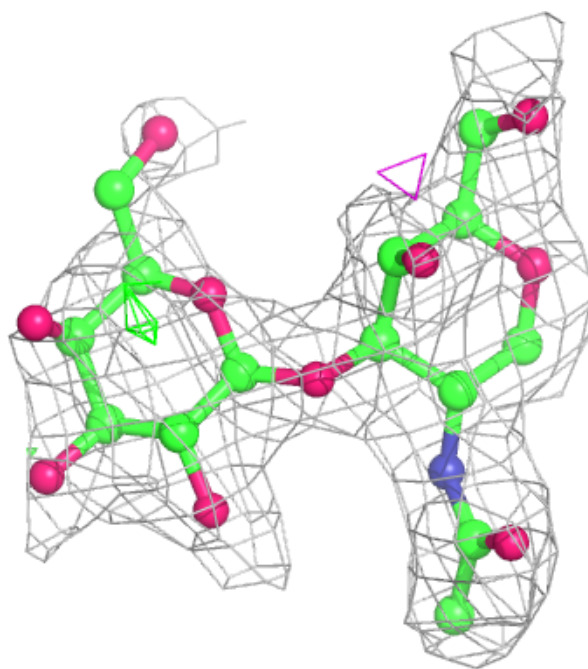
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

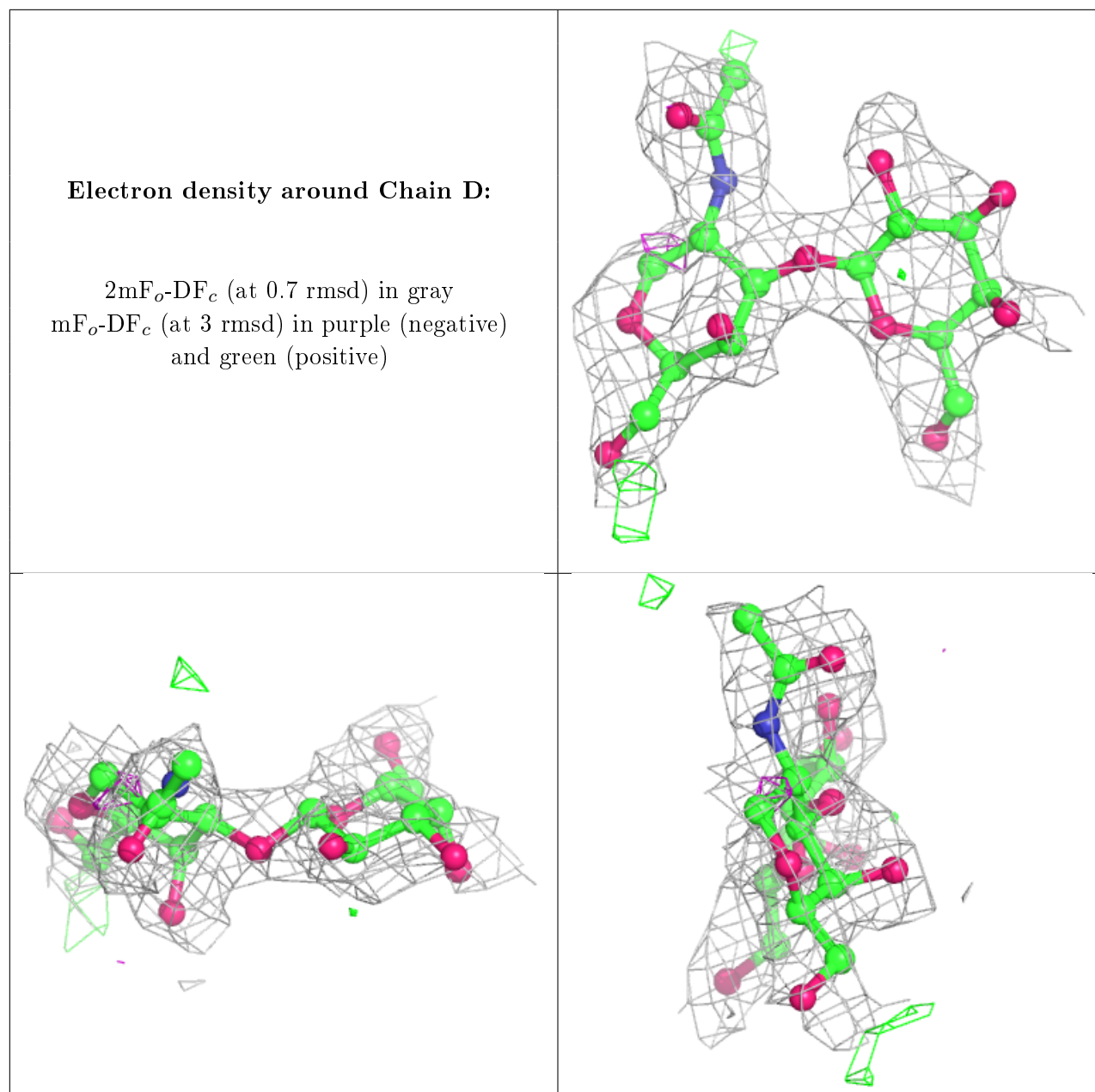
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	GAL	D	2	11/12	0.74	0.28	42,45,47,47	11
2	GAL	C	2	11/12	0.86	0.21	40,43,45,46	11
2	A2G	D	1	14/15	0.91	0.19	35,41,43,45	0
2	A2G	C	1	14/15	0.92	0.16	30,36,39,39	0

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 C:**

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





## 6.4 Ligands ⓘ

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	EDO	B	1015	4/4	0.68	0.24	45,45,45,46	0
5	EDO	B	1013	4/4	0.72	0.21	44,47,47,48	0
5	EDO	A	1015	4/4	0.74	0.23	38,38,39,41	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
5	EDO	A	1014	4/4	0.82	0.21	37,39,39,40	0
5	EDO	A	1020	4/4	0.83	0.26	53,53,54,57	0
5	EDO	A	1011	4/4	0.84	0.29	34,35,37,39	0
5	EDO	A	1013	4/4	0.84	0.15	41,42,43,45	0
5	EDO	B	1014	4/4	0.84	0.23	44,45,47,47	0
5	EDO	A	1010	4/4	0.84	0.20	31,36,37,37	0
5	EDO	B	1012	4/4	0.85	0.20	37,38,40,42	0
5	EDO	B	1017	4/4	0.88	0.21	39,41,41,45	0
5	EDO	A	1016	4/4	0.89	0.22	38,40,41,44	0
6	SER	B	1018	7/7	0.90	0.14	36,40,43,43	0
5	EDO	A	1007	4/4	0.90	0.32	38,39,39,41	0
5	EDO	B	1008	4/4	0.91	0.20	20,21,21,21	0
5	EDO	A	1019	4/4	0.91	0.17	41,43,43,44	0
5	EDO	B	1010	4/4	0.92	0.28	34,35,35,35	0
5	EDO	B	1016	4/4	0.93	0.22	42,42,43,43	0
5	EDO	A	1017	4/4	0.93	0.18	34,34,34,35	0
3	PO4	A	1004	5/5	0.93	0.16	43,49,50,52	1
6	SER	A	1021	7/7	0.93	0.13	28,29,33,36	0
3	PO4	B	1005	5/5	0.93	0.13	37,38,39,41	1
5	EDO	A	1006	4/4	0.94	0.24	27,30,30,31	0
5	EDO	A	1012	4/4	0.95	0.20	32,35,35,35	0
5	EDO	B	1007	4/4	0.95	0.18	40,40,41,42	0
3	PO4	B	1003	5/5	0.95	0.15	40,40,42,44	0
5	EDO	B	1009	4/4	0.95	0.21	35,37,37,37	0
3	PO4	A	1003	5/5	0.95	0.17	44,45,47,47	0
5	EDO	A	1008	4/4	0.95	0.18	27,27,28,28	0
5	EDO	A	1018	4/4	0.95	0.20	38,40,41,44	0
5	EDO	A	1009	4/4	0.96	0.20	19,20,21,21	0
5	EDO	B	1011	4/4	0.96	0.12	36,36,37,37	0
3	PO4	B	1004	5/5	0.97	0.12	36,38,39,40	0
4	ZN	B	1006	1/1	0.99	0.07	24,24,24,24	0
4	ZN	A	1005	1/1	1.00	0.06	22,22,22,22	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.