



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 15, 2023 – 03:59 AM EDT

PDB ID : 7T2D  
Title : Crystal structure of the B1 TCR in complex with HLA-DP4-Ply  
Authors : Ciacchi, L.; Farenc, C.; Petersen, J.; Reid, H.H.; Rossjohn, J.  
Deposited on : 2021-12-03  
Resolution : 3.40 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

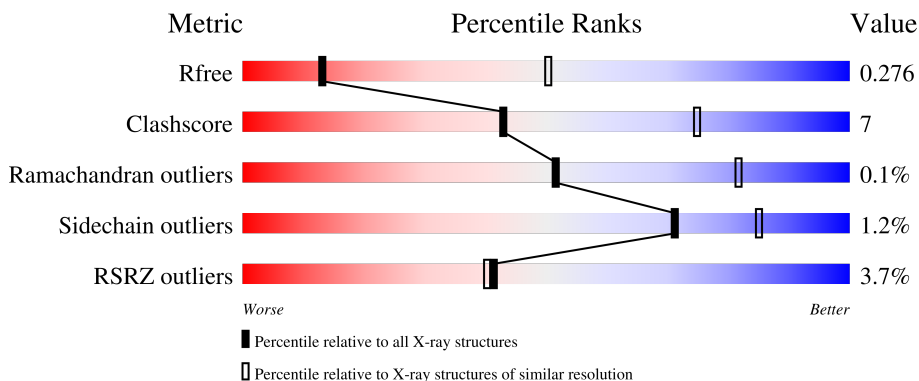
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.40 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	1026 (3.48-3.32)
Clashscore	141614	1055 (3.48-3.32)
Ramachandran outliers	138981	1038 (3.48-3.32)
Sidechain outliers	138945	1038 (3.48-3.32)
RSRZ outliers	127900	2173 (3.50-3.30)

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	181	 84% 15% ..
1	F	181	 % 88% 12% .
1	K	181	 2% 86% 13% .
1	P	181	 2% 90% 10% .
2	B	188	 13% 72% 19% 9%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	188	
2	L	188	
2	Q	188	
3	C	15	
3	H	15	
3	M	15	
3	R	15	
4	D	207	
4	I	207	
4	N	207	
4	S	207	
5	E	243	
5	J	243	
5	O	243	
5	T	243	
6	U	4	
7	V	4	
8	W	3	
8	X	3	

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
7	MAN	V	4	-	-	-	X
9	NAG	B	201	-	-	-	X
9	NAG	F	201	-	-	-	X
9	NAG	G	201	-	-	-	X
9	NAG	L	201	-	-	-	X

Continued on next page...

*Continued from previous page...*

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
9	NAG	Q	201	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 26248 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 HLA class II histocompatibility antigen, DP alpha 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	180	1473	953	238	277	5	0	0	0
1	F	180	1473	953	238	277	5	0	0	0
1	K	180	1473	953	238	277	5	0	0	0
1	P	180	1473	953	238	277	5	0	0	0

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DP beta 1 chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	171	1397	876	248	265	8	0	0	0
2	G	168	1367	855	243	261	8	0	0	0
2	L	178	1460	918	258	276	8	0	0	0
2	Q	177	1451	912	256	275	8	0	0	0

- Molecule 3 is a protein called Pneumolysin-derived peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	14	126	85	20	21	0	0	0
3	H	14	126	85	20	21	0	0	0
3	M	14	126	85	20	21	0	0	0
3	R	14	126	85	20	21	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-3	GLY	-	cloning artifact	UNP Q04IN8
C	-2	ALA	-	cloning artifact	UNP Q04IN8
H	-3	GLY	-	cloning artifact	UNP Q04IN8
H	-2	ALA	-	cloning artifact	UNP Q04IN8
M	-3	GLY	-	cloning artifact	UNP Q04IN8
M	-2	ALA	-	cloning artifact	UNP Q04IN8
R	-3	GLY	-	cloning artifact	UNP Q04IN8
R	-2	ALA	-	cloning artifact	UNP Q04IN8

- Molecule 4 is a protein called T cell receptor, B1, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	201	Total	C	N	O	S	0	0	0
			1597	1003	259	328	7			
4	I	201	Total	C	N	O	S	0	0	0
			1597	1003	259	328	7			
4	N	199	Total	C	N	O	S	0	0	0
			1575	985	257	326	7			
4	S	199	Total	C	N	O	S	0	0	0
			1575	985	257	326	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	176	CYS	THR	engineered mutation	UNP P01848
I	176	CYS	THR	engineered mutation	UNP P01848
N	176	CYS	THR	engineered mutation	UNP P01848
S	176	CYS	THR	engineered mutation	UNP P01848

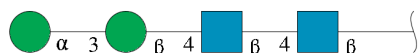
- Molecule 5 is a protein called T cell receptor, B1, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	235	Total	C	N	O	S	0	0	0
			1869	1178	328	358	5			
5	J	237	Total	C	N	O	S	0	0	0
			1883	1187	330	361	5			
5	O	238	Total	C	N	O	S	0	0	0
			1892	1192	332	363	5			
5	T	238	Total	C	N	O	S	0	0	0
			1892	1192	332	363	5			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	184	CYS	SER	engineered mutation	UNP P01850
E	202	ALA	CYS	engineered mutation	UNP P01850
J	184	CYS	SER	engineered mutation	UNP P01850
J	202	ALA	CYS	engineered mutation	UNP P01850
O	184	CYS	SER	engineered mutation	UNP P01850
O	202	ALA	CYS	engineered mutation	UNP P01850
T	184	CYS	SER	engineered mutation	UNP P01850
T	202	ALA	CYS	engineered mutation	UNP P01850

- Molecule 6 is an oligosaccharide called alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



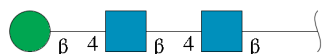
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	U	4	50	28	2	20	0	0	0

- Molecule 7 is an oligosaccharide called alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose.



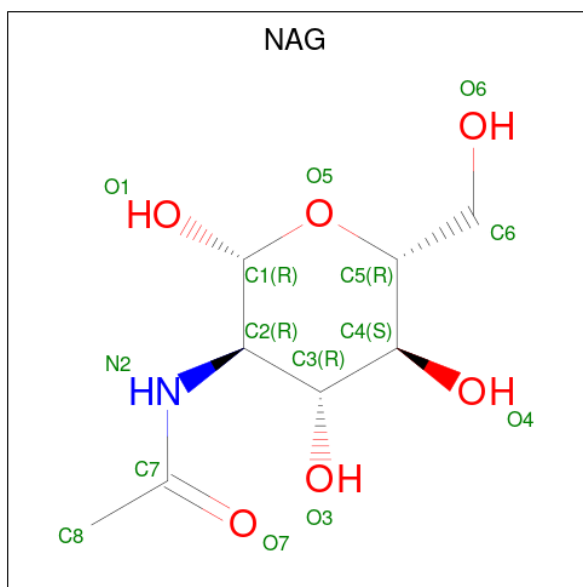
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
7	V	4	50	28	2	20	0	0	0

- Molecule 8 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
8	W	3	Total	C	N	O	0	0	0
			39	22	2	15			
8	X	3	Total	C	N	O	0	0	0
			39	22	2	15			

- Molecule 9 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
9	A	1	Total	C	N	O	0	0
			14	8	1	5		
9	B	1	Total	C	N	O	0	0
			14	8	1	5		
9	F	1	Total	C	N	O	0	0
			14	8	1	5		
9	G	1	Total	C	N	O	0	0
			14	8	1	5		
9	K	1	Total	C	N	O	0	0
			14	8	1	5		
9	L	1	Total	C	N	O	0	0
			14	8	1	5		
9	P	1	Total	C	N	O	0	0
			14	8	1	5		
9	Q	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 10 is water.




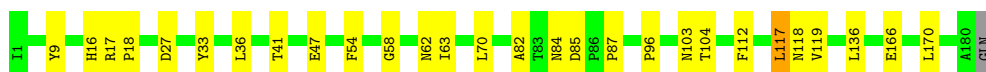
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	A	1	Total O 1 1	0	0
10	C	1	Total O 1 1	0	0
10	F	1	Total O 1 1	0	0
10	E	3	Total O 3 3	0	0
10	I	1	Total O 1 1	0	0

### 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: HLA class II histocompatibility antigen, DP alpha 1 chain

Chain A:  84% 15% ..




- Molecule 1: HLA class II histocompatibility antigen, DP alpha 1 chain

Chain F:  88% 12% .




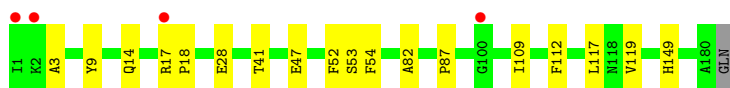
- Molecule 1: HLA class II histocompatibility antigen, DP alpha 1 chain

Chain K:  86% 13% .



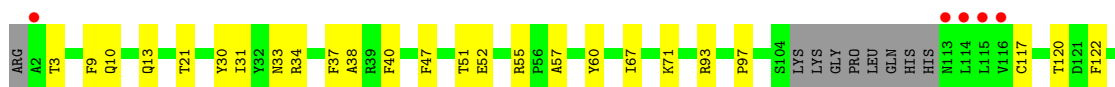
- Molecule 1: HLA class II histocompatibility antigen, DP alpha 1 chain

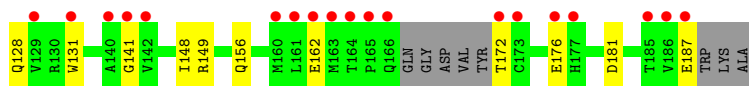
Chain P:  90% 10% .



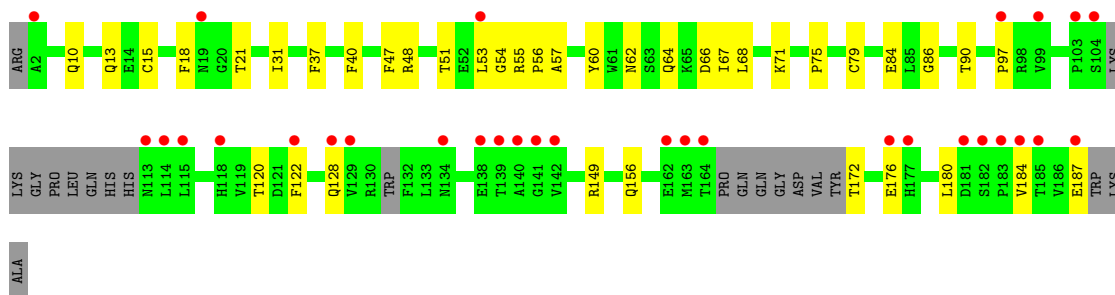
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain

Chain B:  72% 19% 9%

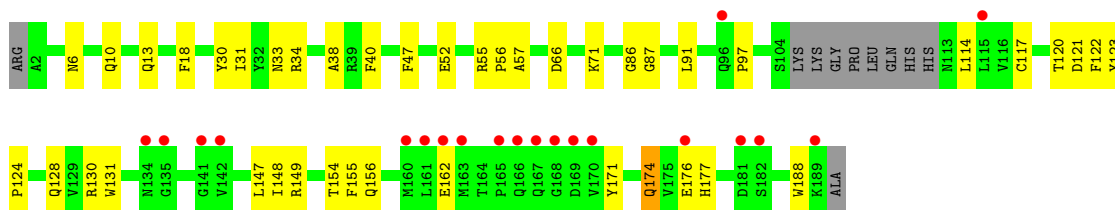




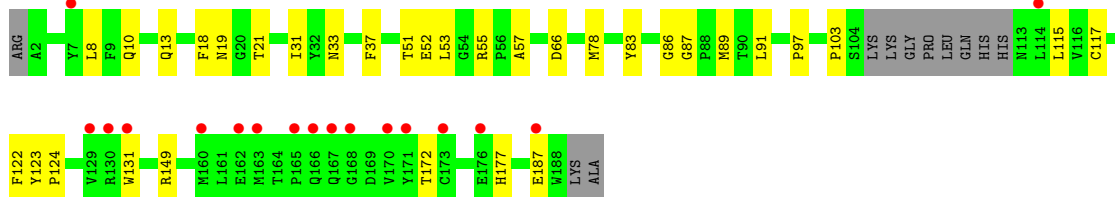
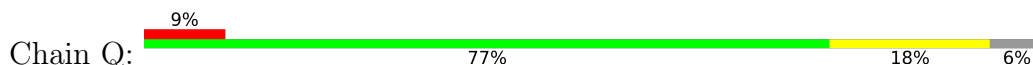
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain



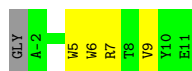
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain



- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain

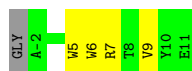


- Molecule 3: Pneumolysin-derived peptide

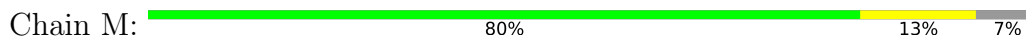


- Molecule 3: Pneumolysin-derived peptide

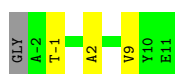




- Molecule 3: Pneumolysin-derived peptide



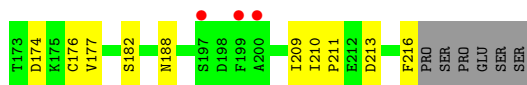
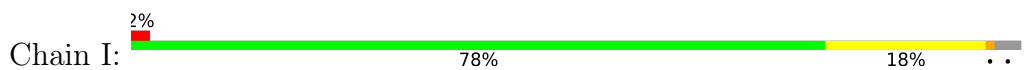
- Molecule 3: Pneumolysin-derived peptide



- Molecule 4: T cell receptor, B1, alpha chain



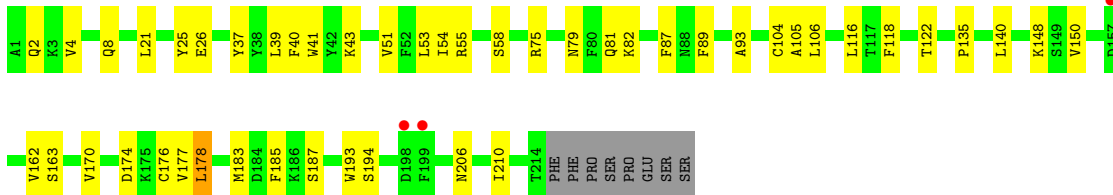
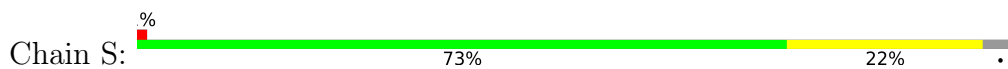
- Molecule 4: T cell receptor, B1, alpha chain



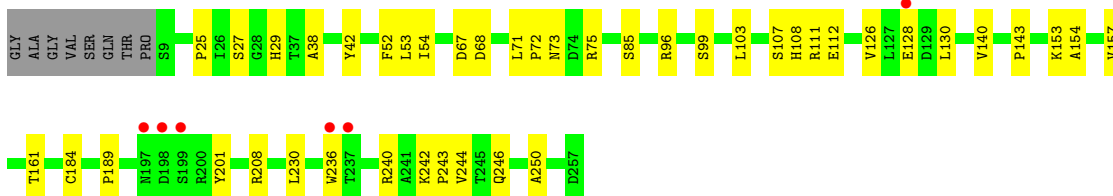
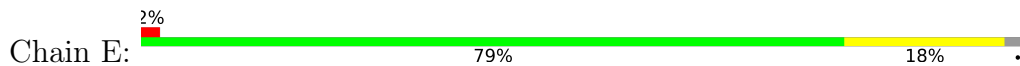
- Molecule 4: T cell receptor, B1, alpha chain



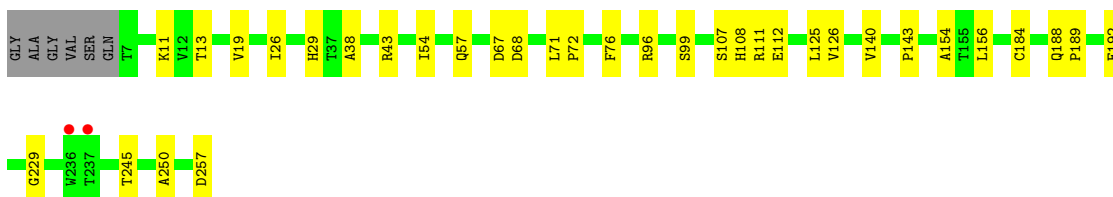
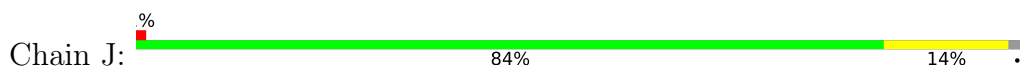
- Molecule 4: T cell receptor, B1, alpha chain



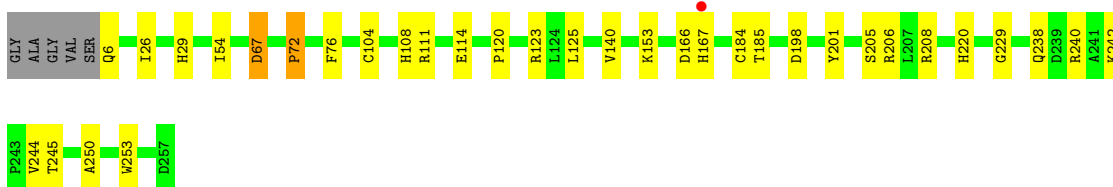
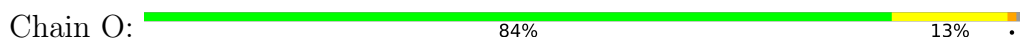
- Molecule 5: T cell receptor, B1, beta chain



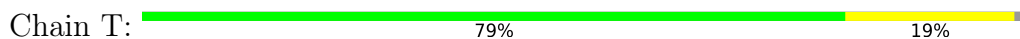
- Molecule 5: T cell receptor, B1, beta chain



- Molecule 5: T cell receptor, B1, beta chain



- Molecule 5: T cell receptor, B1, beta chain





- Molecule 6: alpha-D-mannopyranose-(1-3)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 7: alpha-D-mannopyranose-(1-6)-beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 8: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



- Molecule 8: beta-D-mannopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose-(1-4)-2-acetamido-2-deoxy-beta-D-glucopyranose



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	159.71Å 193.57Å 236.80Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.40 – 3.40 33.43 – 3.40	Depositor EDS
% Data completeness (in resolution range)	99.6 (33.40-3.40) 99.6 (33.43-3.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.67 (at 3.39Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.236 , 0.271 0.241 , 0.276	Depositor DCC
$R_{free}$ test set	5020 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	91.9	Xtrriage
Anisotropy	0.571	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 75.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	26248	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	103.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 47.56 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 9.7578e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: NAG, MAN, 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	0.24	0/1521	0.40	0/2076
1	F	0.26	0/1521	0.44	0/2076
1	K	0.24	0/1521	0.41	0/2076
1	P	0.24	0/1521	0.41	0/2076
2	B	0.24	0/1428	0.41	0/1936
2	G	0.24	0/1394	0.41	0/1886
2	L	0.24	0/1495	0.41	0/2029
2	Q	0.25	0/1486	0.42	0/2018
3	C	0.20	0/132	0.39	0/182
3	H	0.21	0/132	0.38	0/182
3	M	0.21	0/132	0.39	0/182
3	R	0.21	0/132	0.38	0/182
4	D	0.27	0/1631	0.46	0/2215
4	I	0.26	0/1631	0.45	0/2215
4	N	0.27	0/1607	0.47	0/2183
4	S	0.25	0/1607	0.44	0/2183
5	E	0.25	0/1920	0.43	0/2610
5	J	0.25	0/1935	0.43	0/2632
5	O	0.25	0/1944	0.43	0/2644
5	T	0.25	0/1944	0.44	0/2644
All	All	0.25	0/26634	0.43	0/36227

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.



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

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1473	0	1375	18	0
1	F	1473	0	1375	15	0
1	K	1473	0	1375	16	0
1	P	1473	0	1375	14	0
2	B	1397	0	1326	23	0
2	G	1367	0	1300	27	0
2	L	1460	0	1383	29	0
2	Q	1451	0	1370	25	0
3	C	126	0	113	3	0
3	H	126	0	113	3	0
3	M	126	0	113	2	0
3	R	126	0	113	3	0
4	D	1597	0	1505	33	0
4	I	1597	0	1506	32	0
4	N	1575	0	1488	31	0
4	S	1575	0	1488	35	0
5	E	1869	0	1774	26	0
5	J	1883	0	1789	27	0
5	O	1892	0	1797	25	0
5	T	1892	0	1797	30	0
6	U	50	0	43	0	0
7	V	50	0	43	0	0
8	W	39	0	34	1	0
8	X	39	0	34	0	0
9	A	14	0	13	0	0
9	B	14	0	13	0	0
9	F	14	0	13	0	0
9	G	14	0	13	0	0
9	K	14	0	13	0	0
9	L	14	0	13	0	0
9	P	14	0	13	0	0
9	Q	14	0	13	1	0
10	A	1	0	0	0	0
10	C	1	0	0	0	0
10	E	3	0	0	0	0
10	F	1	0	0	0	0
10	I	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	26248	0	24733	343	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 7.

The worst 5 of 343 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)
4:I:176:CYS:CB	5:J:184:CYS:SG	2.27	1.22
4:S:176:CYS:CB	5:T:184:CYS:SG	2.34	1.14
4:N:176:CYS:SG	5:O:184:CYS:CB	2.37	1.12
4:N:176:CYS:CB	5:O:184:CYS:SG	2.40	1.08
4:S:176:CYS:SG	5:T:184:CYS:CB	2.64	0.84

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	178/181 (98%)	175 (98%)	3 (2%)	0	100	100
1	F	178/181 (98%)	171 (96%)	7 (4%)	0	100	100
1	K	178/181 (98%)	173 (97%)	5 (3%)	0	100	100
1	P	178/181 (98%)	172 (97%)	6 (3%)	0	100	100
2	B	165/188 (88%)	157 (95%)	8 (5%)	0	100	100
2	G	160/188 (85%)	156 (98%)	4 (2%)	0	100	100
2	L	174/188 (93%)	169 (97%)	5 (3%)	0	100	100
2	Q	173/188 (92%)	163 (94%)	10 (6%)	0	100	100
3	C	12/15 (80%)	12 (100%)	0	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	H	12/15 (80%)	12 (100%)	0	0	100	100
3	M	12/15 (80%)	12 (100%)	0	0	100	100
3	R	12/15 (80%)	12 (100%)	0	0	100	100
4	D	199/207 (96%)	187 (94%)	11 (6%)	1 (0%)	29	61
4	I	199/207 (96%)	180 (90%)	18 (9%)	1 (0%)	29	61
4	N	197/207 (95%)	186 (94%)	11 (6%)	0	100	100
4	S	197/207 (95%)	188 (95%)	9 (5%)	0	100	100
5	E	233/243 (96%)	223 (96%)	9 (4%)	1 (0%)	34	67
5	J	235/243 (97%)	229 (97%)	6 (3%)	0	100	100
5	O	236/243 (97%)	224 (95%)	11 (5%)	1 (0%)	34	67
5	T	236/243 (97%)	224 (95%)	12 (5%)	0	100	100
All	All	3164/3336 (95%)	3025 (96%)	135 (4%)	4 (0%)	51	82

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	E	27	SER
4	D	120	SER
4	I	182	SER
5	O	72	PRO

### 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	161/162 (99%)	159 (99%)	2 (1%)	71	85
1	F	161/162 (99%)	161 (100%)	0	100	100
1	K	161/162 (99%)	160 (99%)	1 (1%)	86	94
1	P	161/162 (99%)	161 (100%)	0	100	100
2	B	153/167 (92%)	151 (99%)	2 (1%)	69	84

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	G	150/167 (90%)	148 (99%)	2 (1%)	69	84
2	L	159/167 (95%)	158 (99%)	1 (1%)	86	94
2	Q	158/167 (95%)	158 (100%)	0	100	100
3	C	11/11 (100%)	10 (91%)	1 (9%)	9	32
3	H	11/11 (100%)	10 (91%)	1 (9%)	9	32
3	M	11/11 (100%)	11 (100%)	0	100	100
3	R	11/11 (100%)	11 (100%)	0	100	100
4	D	184/190 (97%)	178 (97%)	6 (3%)	38	66
4	I	184/190 (97%)	179 (97%)	5 (3%)	44	70
4	N	182/190 (96%)	177 (97%)	5 (3%)	44	70
4	S	182/190 (96%)	181 (100%)	1 (0%)	88	94
5	E	201/206 (98%)	200 (100%)	1 (0%)	88	94
5	J	203/206 (98%)	202 (100%)	1 (0%)	88	94
5	O	204/206 (99%)	201 (98%)	3 (2%)	65	82
5	T	204/206 (99%)	201 (98%)	3 (2%)	65	82
All	All	2852/2944 (97%)	2817 (99%)	35 (1%)	71	85

5 of 35 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	O	67	ASP
5	O	104	CYS
5	T	83	GLU
4	D	203	ASN
4	D	183	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	14	GLN
1	P	68	ASN
4	S	81	GLN
4	S	65	ASN
5	J	29	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](#)

14 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
6	NAG	U	1	6,1	14,14,15	0.26	0	17,19,21	0.52	0
6	NAG	U	2	6	14,14,15	0.14	0	17,19,21	0.62	0
6	BMA	U	3	6	11,11,12	0.65	0	15,15,17	0.96	1 (6%)
6	MAN	U	4	6	11,11,12	0.80	0	15,15,17	1.64	1 (6%)
7	NAG	V	1	7,1	14,14,15	0.20	0	17,19,21	0.57	0
7	NAG	V	2	7	14,14,15	0.39	0	17,19,21	0.60	0
7	BMA	V	3	7	11,11,12	1.10	1 (9%)	15,15,17	1.30	1 (6%)
7	MAN	V	4	7	11,11,12	1.70	3 (27%)	15,15,17	1.51	3 (20%)
8	NAG	W	1	8,1	14,14,15	0.24	0	17,19,21	0.47	0
8	NAG	W	2	8	14,14,15	0.34	0	17,19,21	1.28	2 (11%)
8	BMA	W	3	8	11,11,12	0.58	0	15,15,17	0.84	0
8	NAG	X	1	8,1	14,14,15	0.21	0	17,19,21	0.47	0
8	NAG	X	2	8	14,14,15	0.34	0	17,19,21	0.55	0
8	BMA	X	3	8	11,11,12	0.95	0	15,15,17	0.87	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	NAG	U	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	U	2	6	-	1/6/23/26	0/1/1/1
6	BMA	U	3	6	-	0/2/19/22	0/1/1/1
6	MAN	U	4	6	-	0/2/19/22	0/1/1/1
7	NAG	V	1	7,1	-	0/6/23/26	0/1/1/1
7	NAG	V	2	7	-	2/6/23/26	0/1/1/1
7	BMA	V	3	7	-	2/2/19/22	0/1/1/1
7	MAN	V	4	7	-	0/2/19/22	1/1/1/1
8	NAG	W	1	8,1	-	0/6/23/26	0/1/1/1
8	NAG	W	2	8	-	3/6/23/26	0/1/1/1
8	BMA	W	3	8	-	2/2/19/22	0/1/1/1
8	NAG	X	1	8,1	-	0/6/23/26	0/1/1/1
8	NAG	X	2	8	-	0/6/23/26	0/1/1/1
8	BMA	X	3	8	-	1/2/19/22	0/1/1/1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	V	4	MAN	O5-C5	3.63	1.50	1.43
7	V	4	MAN	O5-C1	3.33	1.49	1.43
7	V	4	MAN	C1-C2	2.42	1.57	1.52
7	V	3	BMA	O5-C5	2.13	1.47	1.43

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	U	4	MAN	C1-O5-C5	5.10	119.11	112.19
8	W	2	NAG	C2-N2-C7	4.37	129.12	122.90
7	V	3	BMA	O5-C5-C6	4.30	113.94	107.20
7	V	4	MAN	C1-O5-C5	3.74	117.26	112.19
7	V	4	MAN	O2-C2-C3	-2.38	105.38	110.14

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	V	3	BMA	O5-C5-C6-O6
7	V	3	BMA	C4-C5-C6-O6
8	W	2	NAG	C8-C7-N2-C2
8	W	2	NAG	O7-C7-N2-C2
8	W	3	BMA	O5-C5-C6-O6

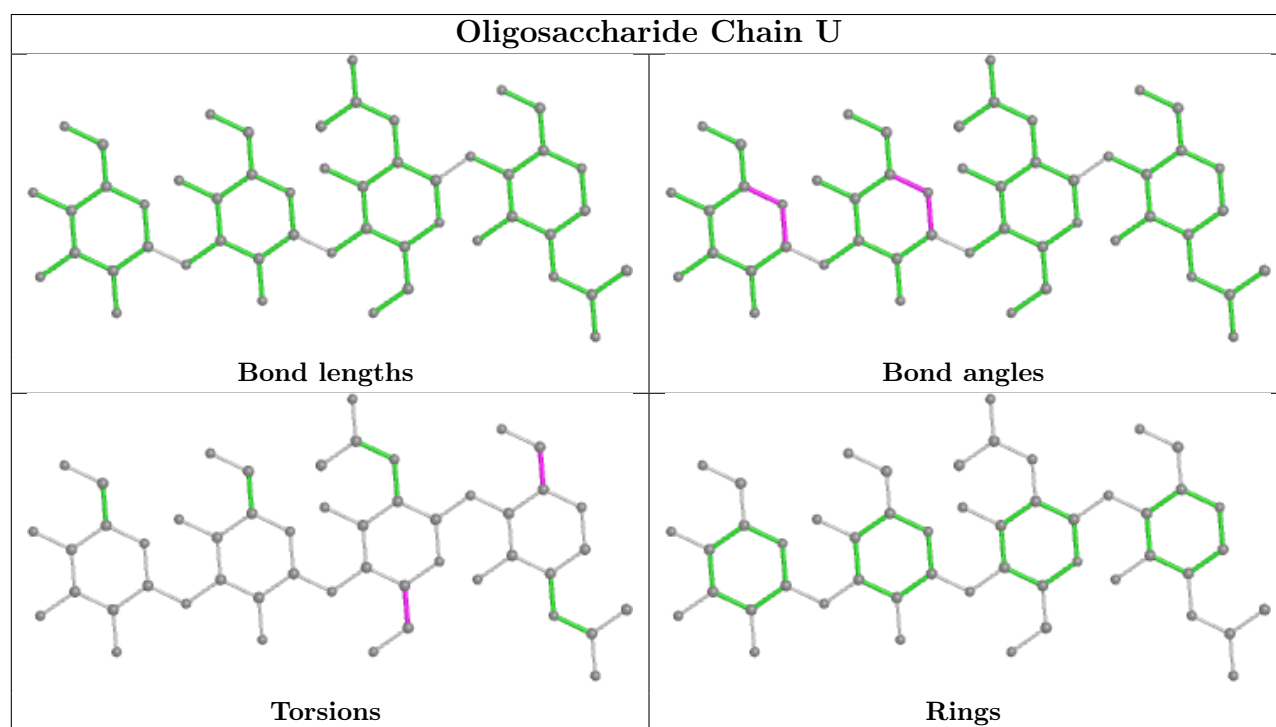
All (1) ring outliers are listed below:

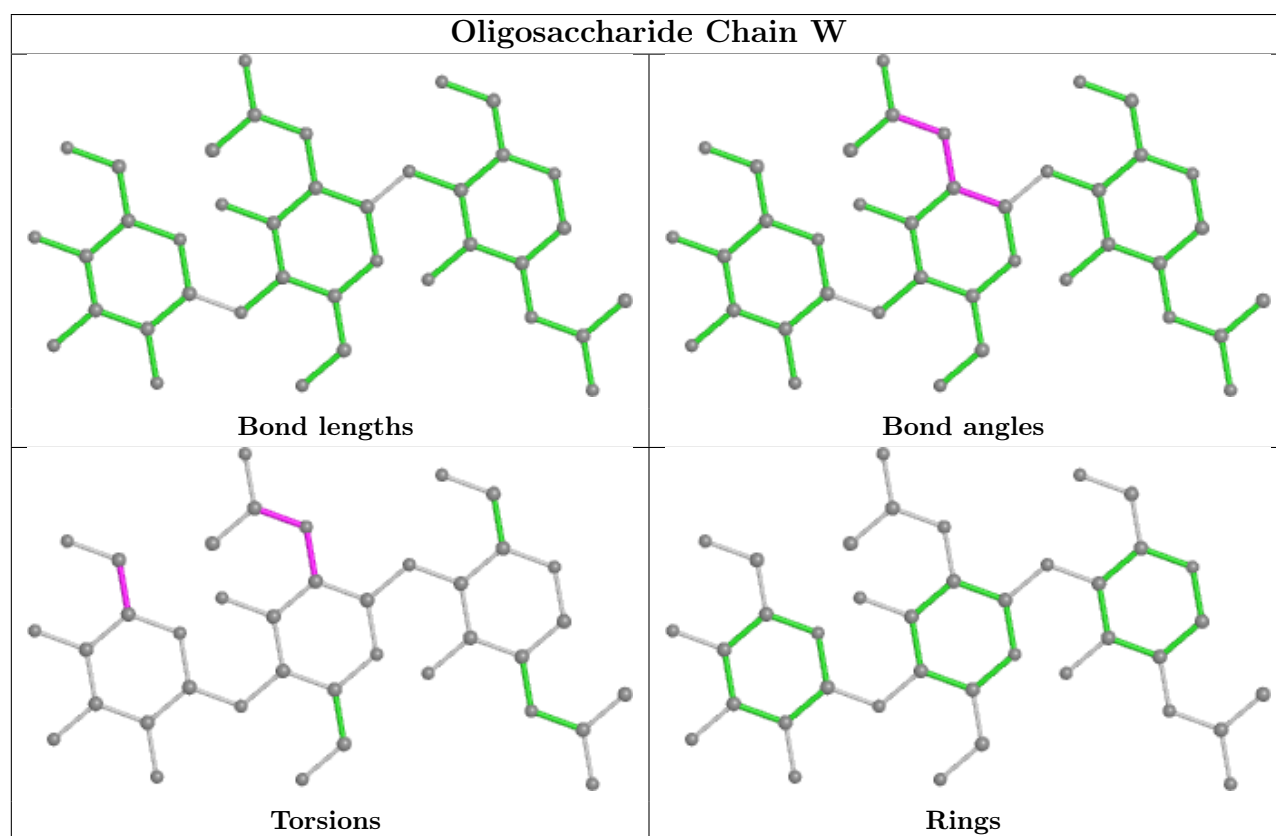
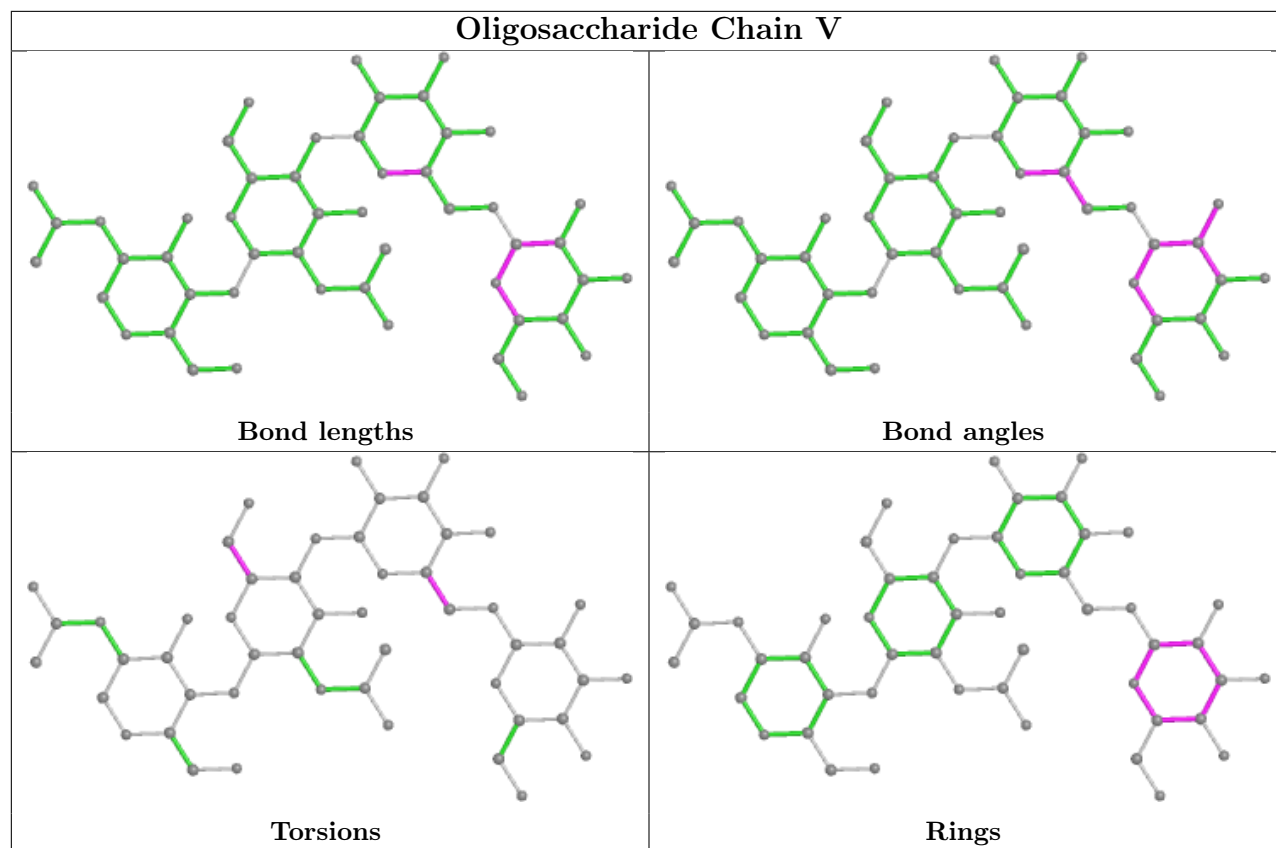
Mol	Chain	Res	Type	Atoms
7	V	4	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 1 short contact:

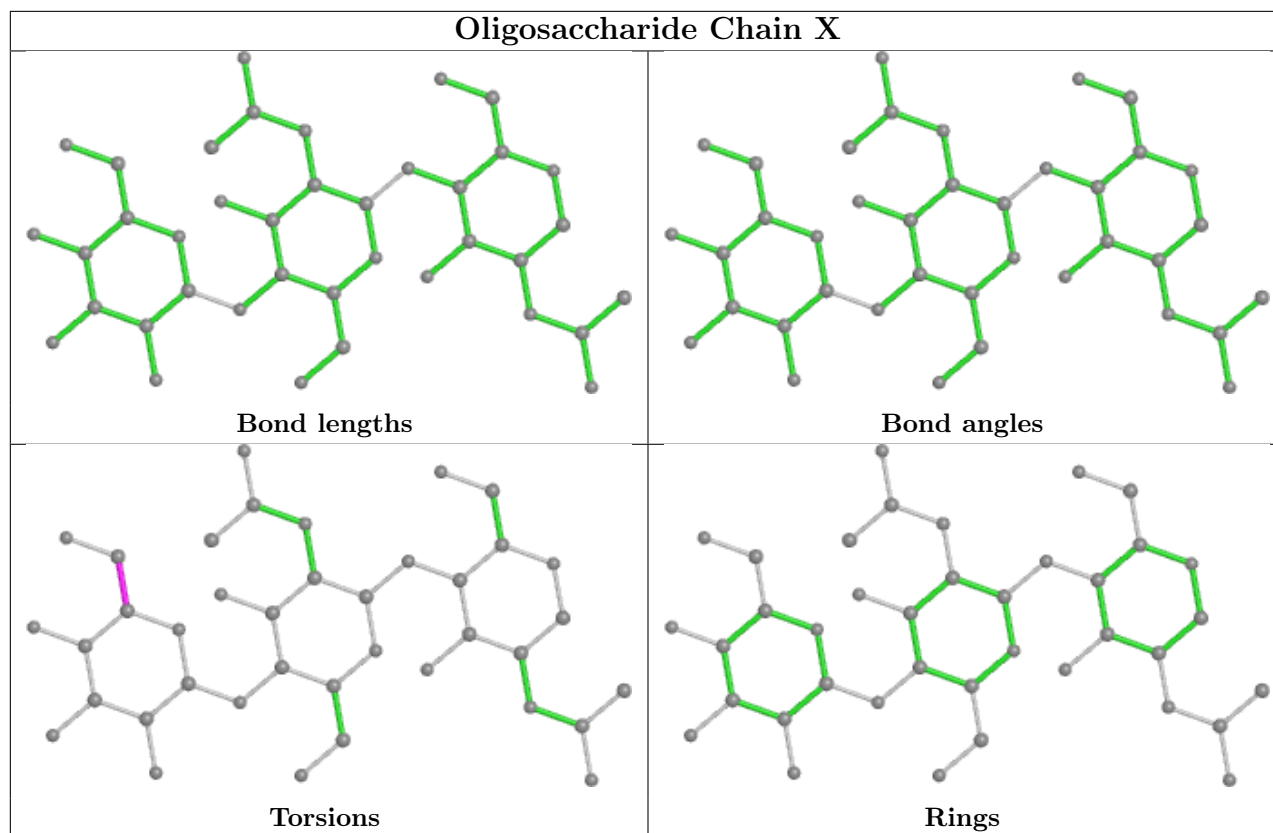
Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	W	2	NAG	1	0

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](#)

8 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	$\# Z  > 2$
9	NAG	A	201	1	14,14,15	0.24	0	17,19,21	0.43	0
9	NAG	Q	201	2	14,14,15	1.10	1 (7%)	17,19,21	1.08	1 (5%)
9	NAG	L	201	2	14,14,15	0.27	0	17,19,21	0.45	0
9	NAG	B	201	2	14,14,15	0.26	0	17,19,21	0.44	0
9	NAG	P	201	1	14,14,15	0.36	0	17,19,21	0.59	0
9	NAG	F	201	1	14,14,15	0.24	0	17,19,21	0.56	0
9	NAG	K	201	1	14,14,15	0.24	0	17,19,21	0.44	0
9	NAG	G	201	2	14,14,15	0.20	0	17,19,21	0.57	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
9	NAG	A	201	1	-	0/6/23/26	0/1/1/1
9	NAG	Q	201	2	-	1/6/23/26	0/1/1/1
9	NAG	L	201	2	-	1/6/23/26	0/1/1/1
9	NAG	B	201	2	-	0/6/23/26	0/1/1/1
9	NAG	P	201	1	-	0/6/23/26	0/1/1/1
9	NAG	F	201	1	-	1/6/23/26	0/1/1/1
9	NAG	K	201	1	-	0/6/23/26	0/1/1/1
9	NAG	G	201	2	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	Q	201	NAG	C1-C2	3.17	1.57	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	Q	201	NAG	C2-N2-C7	3.10	127.32	122.90

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	Q	201	NAG	C1-C2-N2-C7
9	L	201	NAG	O5-C5-C6-O6
9	G	201	NAG	C3-C2-N2-C7
9	F	201	NAG	C3-C2-N2-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	Q	201	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	180/181 (99%)	-0.06	0 <b>100</b> <b>100</b>	61, 85, 114, 143	0
1	F	180/181 (99%)	0.06	1 (0%) <b>89</b> <b>89</b>	70, 98, 138, 176	0
1	K	180/181 (99%)	-0.02	3 (1%) <b>70</b> <b>68</b>	65, 88, 128, 186	0
1	P	180/181 (99%)	0.05	4 (2%) <b>62</b> <b>60</b>	63, 88, 133, 189	0
2	B	171/188 (90%)	0.50	24 (14%) <b>2</b> <b>3</b>	67, 106, 176, 228	0
2	G	168/188 (89%)	0.80	31 (18%) <b>1</b> <b>1</b>	78, 123, 202, 234	8 (4%)
2	L	178/188 (94%)	0.44	20 (11%) <b>5</b> <b>6</b>	74, 112, 188, 222	0
2	Q	177/188 (94%)	0.42	17 (9%) <b>8</b> <b>9</b>	71, 110, 185, 224	0
3	C	14/15 (93%)	0.19	0 <b>100</b> <b>100</b>	64, 75, 108, 113	0
3	H	14/15 (93%)	0.22	0 <b>100</b> <b>100</b>	73, 89, 121, 140	0
3	M	14/15 (93%)	0.25	0 <b>100</b> <b>100</b>	70, 86, 122, 131	0
3	R	14/15 (93%)	0.28	0 <b>100</b> <b>100</b>	65, 79, 108, 132	0
4	D	201/207 (97%)	-0.03	1 (0%) <b>91</b> <b>90</b>	59, 93, 155, 196	0
4	I	201/207 (97%)	0.03	5 (2%) <b>57</b> <b>55</b>	64, 100, 159, 202	0
4	N	199/207 (96%)	0.07	0 <b>100</b> <b>100</b>	60, 101, 161, 193	0
4	S	199/207 (96%)	-0.01	3 (1%) <b>73</b> <b>72</b>	60, 96, 162, 200	0
5	E	235/243 (96%)	-0.08	6 (2%) <b>56</b> <b>54</b>	59, 83, 130, 173	0
5	J	237/243 (97%)	-0.12	2 (0%) <b>86</b> <b>85</b>	63, 87, 125, 171	0
5	O	238/243 (97%)	-0.07	1 (0%) <b>92</b> <b>92</b>	62, 98, 153, 185	0
5	T	238/243 (97%)	-0.05	1 (0%) <b>92</b> <b>92</b>	61, 99, 151, 194	0
All	All	3218/3336 (96%)	0.10	119 (3%) <b>41</b> <b>40</b>	59, 95, 167, 234	8 (0%)

The worst 5 of 119 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	G	113	ASN	4.9
2	G	163	MET	4.8
2	G	140	ALA	4.5
2	G	134	ASN	4.5
2	B	166	GLN	4.4

## 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](#)

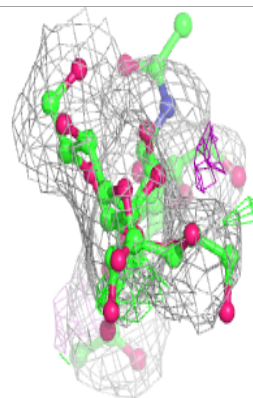
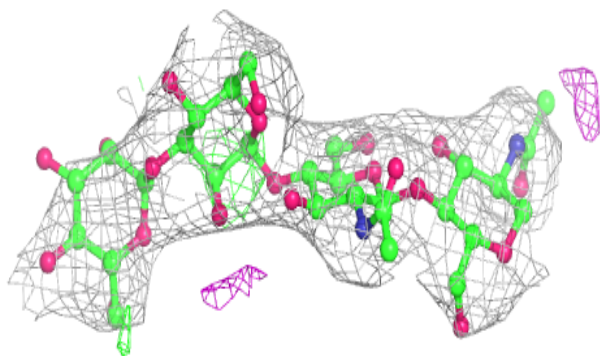
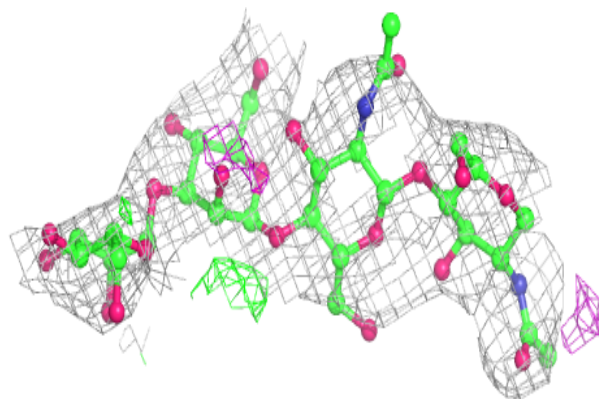
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 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
7	MAN	V	4	11/12	0.56	0.45	203,210,217,219	0
8	BMA	X	3	11/12	0.76	0.24	148,157,178,195	0
7	BMA	V	3	11/12	0.77	0.28	162,164,177,182	0
7	NAG	V	2	14/15	0.80	0.28	122,138,154,160	0
6	BMA	U	3	11/12	0.81	0.23	120,148,156,158	0
6	MAN	U	4	11/12	0.81	0.33	151,193,202,219	0
8	NAG	W	2	14/15	0.87	0.35	136,159,176,177	0
8	NAG	X	2	14/15	0.88	0.33	136,142,160,168	0
8	BMA	W	3	11/12	0.89	0.17	121,151,163,164	0
6	NAG	U	2	14/15	0.91	0.24	107,124,139,144	0
7	NAG	V	1	14/15	0.92	0.21	89,101,130,133	0
8	NAG	X	1	14/15	0.94	0.24	100,107,133,138	0
8	NAG	W	1	14/15	0.95	0.20	91,115,136,141	0
6	NAG	U	1	14/15	0.95	0.17	75,99,111,116	0

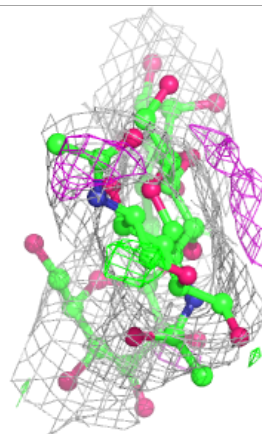
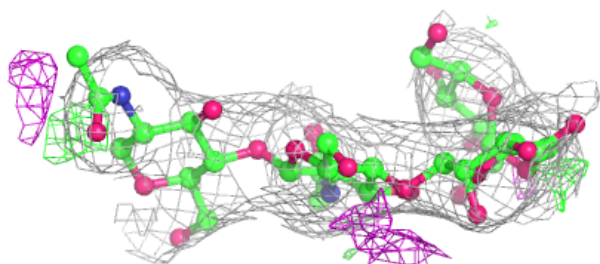
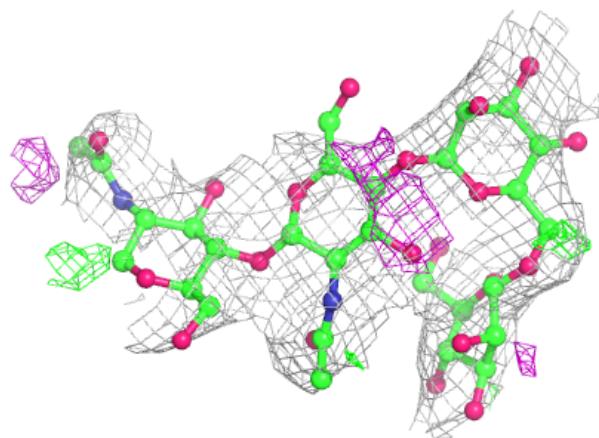
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain U:**

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

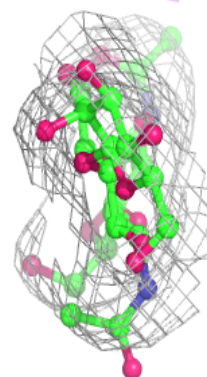
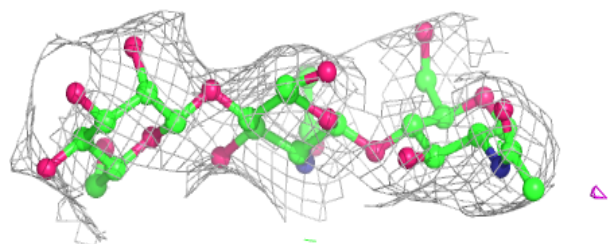
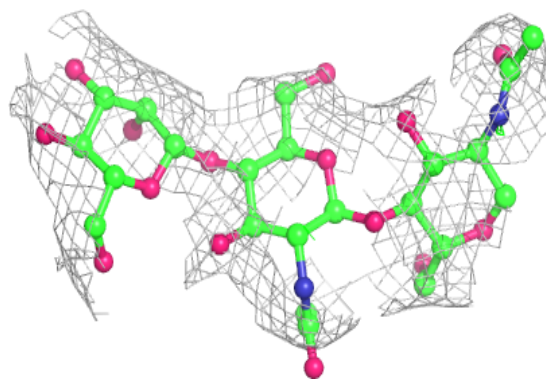
**Electron density around Chain V:**

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

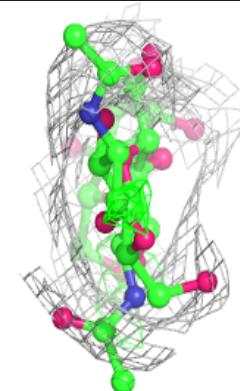
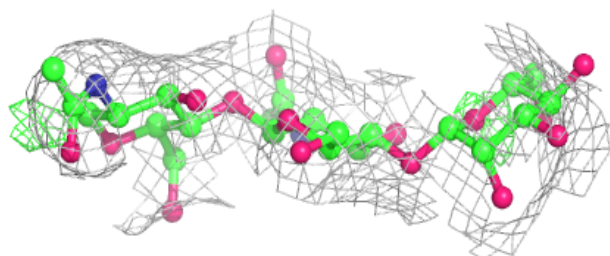
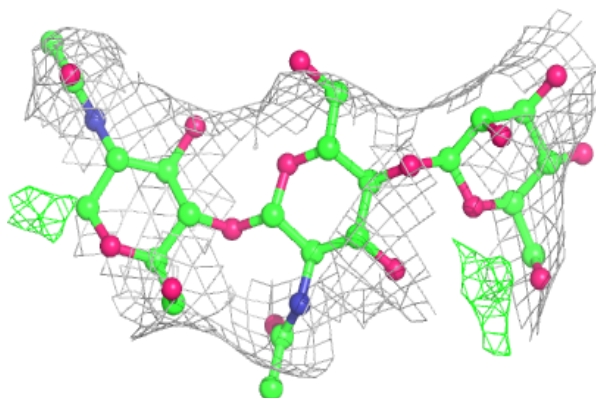


**Electron density around Chain W:**

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

**Electron density around Chain X:**

$2mF_o-DF_c$  (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
9	NAG	G	201	14/15	0.70	0.51	132,183,192,195	0
9	NAG	L	201	14/15	0.72	0.44	158,178,195,196	0
9	NAG	K	201	14/15	0.74	0.39	146,156,172,173	0
9	NAG	B	201	14/15	0.75	0.45	126,170,183,188	0
9	NAG	Q	201	14/15	0.75	0.55	133,187,193,196	0
9	NAG	P	201	14/15	0.79	0.39	144,166,180,181	0
9	NAG	F	201	14/15	0.79	0.40	147,167,180,181	0
9	NAG	A	201	14/15	0.80	0.31	119,138,155,158	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.