



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 14, 2023 – 08:25 PM EDT

PDB ID : 7T2B  
Title : Crystal structure of the 5F 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 : 2.80 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

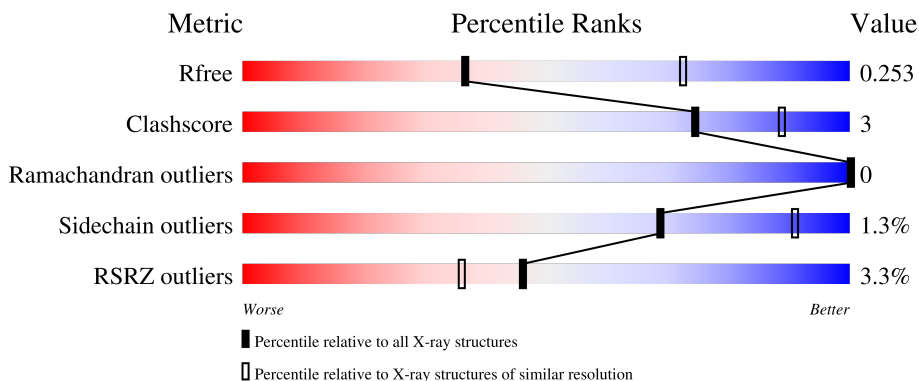
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

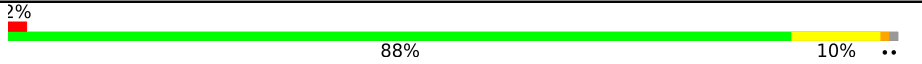


The reported resolution of this entry is 2.80 Å.

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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	 2% 88% 10% ..
1	F	181	 92% 7% ..
1	K	181	 % 90% 9% ..
1	P	181	 2% 87% 12% ..
2	B	190	 14% 78% 16% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	190	 3% 83% 15% .
2	L	190	 6% 82% 13% 6%
2	Q	190	 4% 88% 12% .
3	C	15	 87% 7% 7%
3	H	15	 87% 7% 7%
3	M	15	 67% 20% 13%
3	R	15	 87% 7% 7%
4	D	207	 3% 90% . 5%
4	I	207	 11% 89% . 8%
4	N	207	 88% 7% 5%
4	S	207	 2% 90% . 5%
5	E	241	 2% 94% 5% .
5	J	241	 88% 10% ..
5	O	241	 2% 91% 8% .
5	T	241	 90% 9% .
6	U	2	 100%
6	V	2	 50% 50%
6	W	2	 50% 50%

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 26222 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	180	1473	924	262	279	8	0	0	0
2	G	185	1517	954	271	284	8	0	0	0
2	L	179	1470	924	261	277	8	0	0	0
2	Q	189	1548	972	279	289	8	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	ALA	-	expression tag	UNP P04440
B	0	SER	-	expression tag	UNP P04440
G	-1	ALA	-	expression tag	UNP P04440
G	0	SER	-	expression tag	UNP P04440
L	-1	ALA	-	expression tag	UNP P04440
L	0	SER	-	expression tag	UNP P04440
Q	-1	ALA	-	expression tag	UNP P04440
Q	0	SER	-	expression tag	UNP P04440

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	14	Total	C	N	O	0	0	0
			126	85	20	21			
3	H	14	Total	C	N	O	0	0	0
			126	85	20	21			
3	M	13	Total	C	N	O	0	0	0
			121	82	19	20			
3	R	14	Total	C	N	O	0	0	0
			126	85	20	21			

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, 5F, alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	196	Total	C	N	O	S	0	0	0
			1515	935	266	306	8			
4	I	190	Total	C	N	O	S	0	0	0
			1466	907	258	293	8			
4	N	197	Total	C	N	O	S	0	0	0
			1522	940	267	307	8			
4	S	197	Total	C	N	O	S	0	0	0
			1522	940	267	307	8			

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, 5F, beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	239	Total 1889	C 1190	N 328	O 361	S 10	0	0	0
5	J	239	Total 1889	C 1190	N 328	O 361	S 10	0	0	0
5	O	239	Total 1889	C 1190	N 328	O 361	S 10	0	0	0
5	T	239	Total 1889	C 1190	N 328	O 361	S 10	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

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

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



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
6	U	2	Total 28	C 16	N 2	O 10	0	0	0
6	V	2	Total 28	C 16	N 2	O 10	0	0	0
6	W	2	Total 28	C 16	N 2	O 10	0	0	0

- Molecule 7 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 8 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
8	A	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0
8	F	1	Total C O 4 2 2	0	0
8	P	1	Total C O 4 2 2	0	0
8	E	1	Total C O 4 2 2	0	0
8	J	1	Total C O 4 2 2	0	0
8	N	1	Total C O 4 2 2	0	0
8	N	1	Total C O 4 2 2	0	0
8	S	1	Total C O 4 2 2	0	0

- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	2	Total O 2 2	0	0
9	F	3	Total O 3 3	0	0
9	G	1	Total O 1 1	0	0

*Continued on next page...*



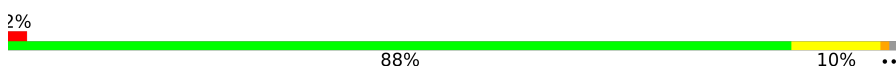
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
9	P	1	Total O 1 1	0	0
9	Q	1	Total O 1 1	0	0
9	D	1	Total O 1 1	0	0
9	O	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: 

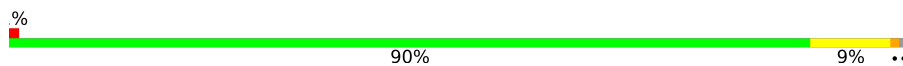


- Molecule 1: HLA class II histocompatibility antigen, DP alpha 1 chain

Chain F: 




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

Chain K: 




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

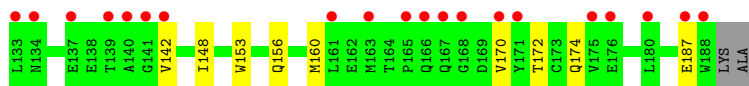
Chain P: 



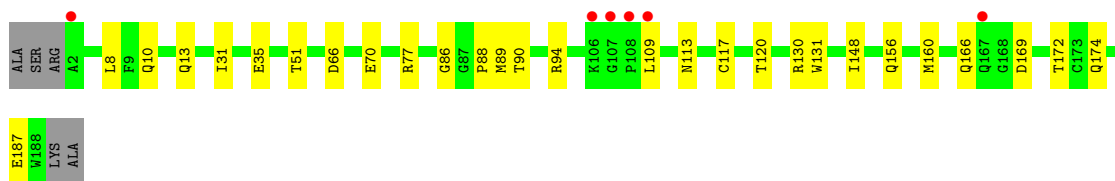
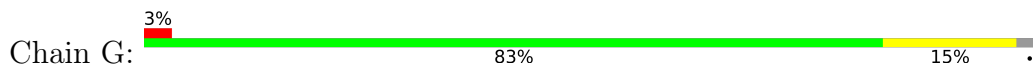
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain

Chain B: 

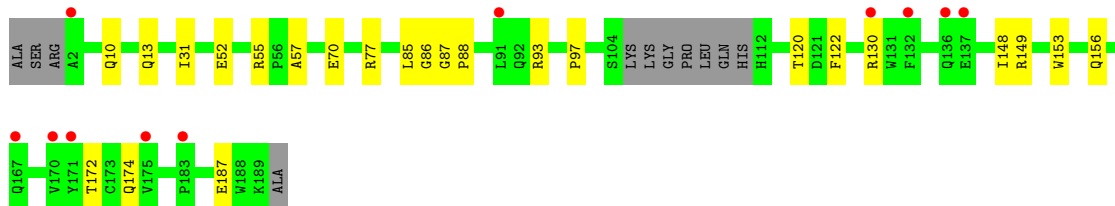
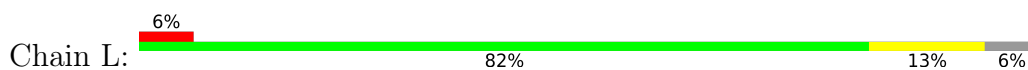




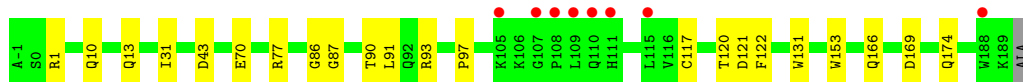
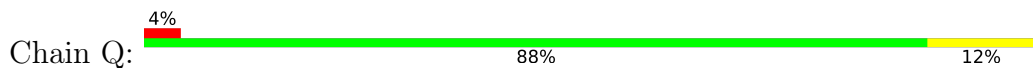
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain



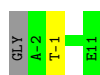
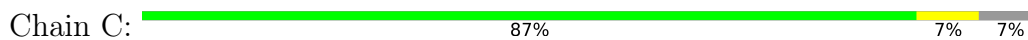
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain



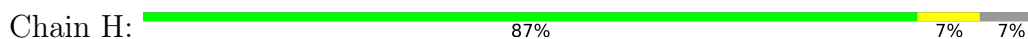
- Molecule 2: HLA class II histocompatibility antigen, DP beta 1 chain



- Molecule 3: Pneumolysin-derived peptide



- Molecule 3: Pneumolysin-derived peptide

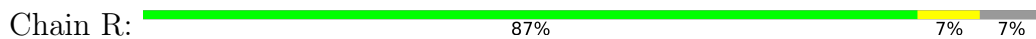


- Molecule 3: Pneumolysin-derived peptide

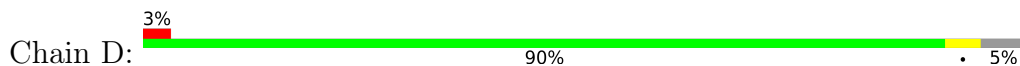




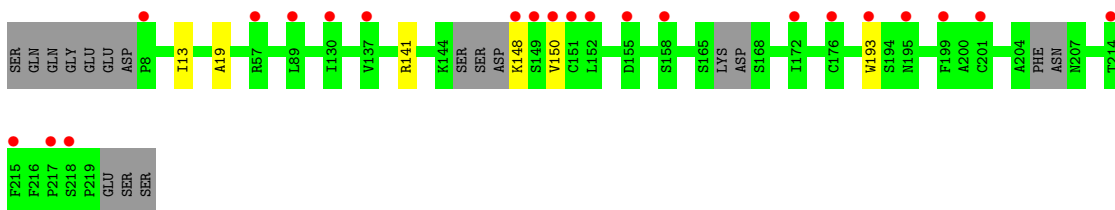
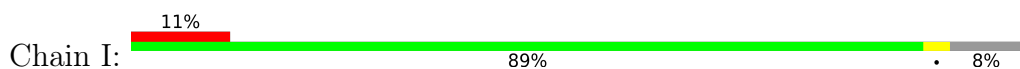
- Molecule 3: Pneumolysin-derived peptide



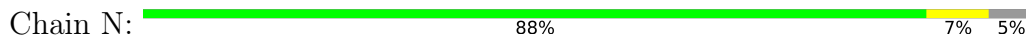
- Molecule 4: T cell receptor, 5F, alpha chain



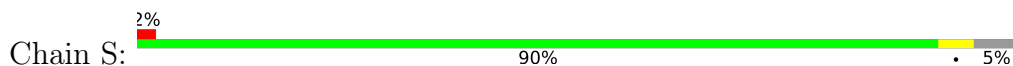
- Molecule 4: T cell receptor, 5F, alpha chain



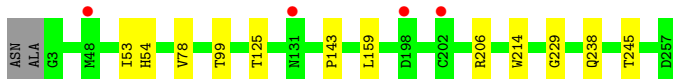
- Molecule 4: T cell receptor, 5F, alpha chain



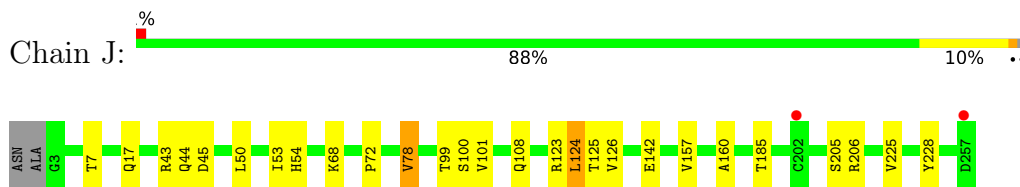
- Molecule 4: T cell receptor, 5F, alpha chain



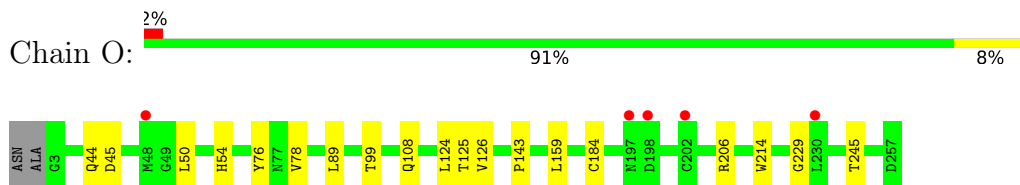
- Molecule 5: T cell receptor, 5F, beta chain



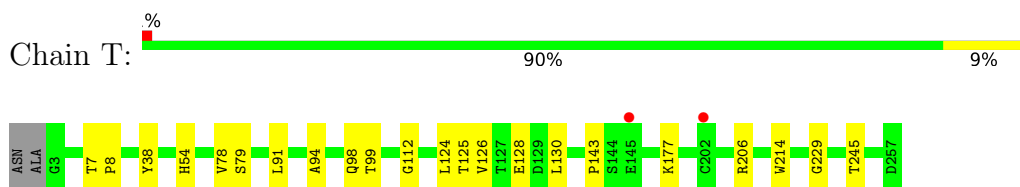
- Molecule 5: T cell receptor, 5F, beta chain



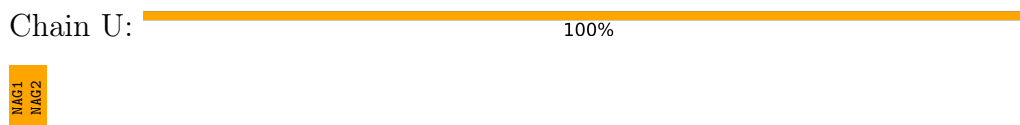
- Molecule 5: T cell receptor, 5F, beta chain



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



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



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



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



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	107.05Å 110.78Å 403.25Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.83 – 2.80 49.19 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (48.83-2.80) 99.7 (49.19-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.11 (at 2.81Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, $R_{free}$	0.216 , 0.250 0.220 , 0.253	Depositor DCC
$R_{free}$ test set	1990 reflections (1.69%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	58.8	Xtrriage
Anisotropy	0.869	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 44.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	26222	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	72.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, 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	0.25	0/1521	0.43	0/2076
1	F	0.25	0/1521	0.44	0/2076
1	K	0.24	0/1521	0.42	0/2076
1	P	0.26	0/1521	0.45	0/2076
2	B	0.25	0/1508	0.45	0/2047
2	G	0.25	0/1556	0.43	0/2113
2	L	0.24	0/1506	0.43	0/2044
2	Q	0.24	0/1587	0.43	0/2153
3	C	0.21	0/132	0.41	0/182
3	H	0.20	0/132	0.40	0/182
3	M	0.21	0/127	0.38	0/175
3	R	0.20	0/132	0.39	0/182
4	D	0.26	0/1540	0.45	0/2082
4	I	0.24	0/1488	0.44	0/2008
4	N	0.25	0/1548	0.45	0/2093
4	S	0.25	0/1548	0.44	0/2093
5	E	0.25	0/1940	0.45	0/2636
5	J	0.25	0/1940	0.45	0/2636
5	O	0.24	0/1940	0.44	0/2636
5	T	0.25	0/1940	0.45	0/2636
All	All	0.25	0/26648	0.44	0/36202

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1473	0	1375	13	0
1	F	1473	0	1375	11	0
1	K	1473	0	1375	12	0
1	P	1473	0	1375	13	0
2	B	1473	0	1396	20	0
2	G	1517	0	1440	14	0
2	L	1470	0	1391	16	0
2	Q	1548	0	1479	12	0
3	C	126	0	113	1	0
3	H	126	0	113	1	0
3	M	121	0	108	3	0
3	R	126	0	113	0	0
4	D	1515	0	1474	6	0
4	I	1466	0	1433	5	0
4	N	1522	0	1483	12	0
4	S	1522	0	1482	5	0
5	E	1889	0	1798	8	0
5	J	1889	0	1798	14	0
5	O	1889	0	1799	11	0
5	T	1889	0	1798	13	0
6	U	28	0	25	1	0
6	V	28	0	25	1	0
6	W	28	0	25	0	0
7	A	14	0	13	0	0
7	B	14	0	13	0	0
7	F	14	0	13	0	0
7	G	14	0	13	0	0
7	K	28	0	26	1	0
7	P	14	0	13	0	0
7	Q	14	0	13	0	0
8	A	4	0	6	0	0
8	B	4	0	6	0	0
8	E	4	0	6	0	0
8	F	4	0	6	1	0
8	J	4	0	6	1	0
8	N	8	0	12	2	0
8	P	4	0	6	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	S	4	0	6	0	0
9	A	2	0	0	0	0
9	D	1	0	0	0	0
9	F	3	0	0	0	0
9	G	1	0	0	0	0
9	O	1	0	0	0	0
9	P	1	0	0	1	0
9	Q	1	0	0	0	0
All	All	26222	0	24951	162	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 3.

The worst 5 of 162 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:N:176:CYS:CB	5:O:184:CYS:SG	2.36	1.13
5:T:229:GLY:H	5:T:245:THR:HG22	1.50	0.76
1:K:77:SER:HA	7:K:201:NAG:H82	1.69	0.74
5:O:229:GLY:H	5:O:245:THR:HG22	1.52	0.73
2:Q:97:PRO:HB3	2:Q:122:PHE:HB3	1.75	0.69

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%)	178 (100%)	0	0	100 100
1	F	178/181 (98%)	178 (100%)	0	0	100 100
1	K	178/181 (98%)	176 (99%)	2 (1%)	0	100 100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	P	178/181 (98%)	178 (100%)	0	0	100	100
2	B	176/190 (93%)	171 (97%)	5 (3%)	0	100	100
2	G	183/190 (96%)	180 (98%)	3 (2%)	0	100	100
2	L	175/190 (92%)	172 (98%)	3 (2%)	0	100	100
2	Q	187/190 (98%)	180 (96%)	7 (4%)	0	100	100
3	C	12/15 (80%)	12 (100%)	0	0	100	100
3	H	12/15 (80%)	12 (100%)	0	0	100	100
3	M	11/15 (73%)	11 (100%)	0	0	100	100
3	R	12/15 (80%)	12 (100%)	0	0	100	100
4	D	194/207 (94%)	189 (97%)	5 (3%)	0	100	100
4	I	182/207 (88%)	178 (98%)	4 (2%)	0	100	100
4	N	195/207 (94%)	192 (98%)	3 (2%)	0	100	100
4	S	195/207 (94%)	190 (97%)	5 (3%)	0	100	100
5	E	237/241 (98%)	231 (98%)	6 (2%)	0	100	100
5	J	237/241 (98%)	233 (98%)	4 (2%)	0	100	100
5	O	237/241 (98%)	232 (98%)	5 (2%)	0	100	100
5	T	237/241 (98%)	233 (98%)	4 (2%)	0	100	100
All	All	3194/3336 (96%)	3138 (98%)	56 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	161/162 (99%)	159 (99%)	2 (1%)	71	92
1	F	161/162 (99%)	160 (99%)	1 (1%)	86	96
1	K	161/162 (99%)	160 (99%)	1 (1%)	86	96
1	P	161/162 (99%)	159 (99%)	2 (1%)	71	92

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	160/168 (95%)	158 (99%)	2 (1%)	69	91
2	G	165/168 (98%)	160 (97%)	5 (3%)	41	75
2	L	160/168 (95%)	160 (100%)	0	100	100
2	Q	168/168 (100%)	165 (98%)	3 (2%)	59	86
3	C	11/11 (100%)	11 (100%)	0	100	100
3	H	11/11 (100%)	11 (100%)	0	100	100
3	M	11/11 (100%)	11 (100%)	0	100	100
3	R	11/11 (100%)	10 (91%)	1 (9%)	9	27
4	D	173/183 (94%)	172 (99%)	1 (1%)	86	96
4	I	167/183 (91%)	167 (100%)	0	100	100
4	N	174/183 (95%)	172 (99%)	2 (1%)	73	92
4	S	174/183 (95%)	172 (99%)	2 (1%)	73	92
5	E	205/206 (100%)	203 (99%)	2 (1%)	76	93
5	J	205/206 (100%)	199 (97%)	6 (3%)	42	76
5	O	205/206 (100%)	202 (98%)	3 (2%)	65	89
5	T	205/206 (100%)	202 (98%)	3 (2%)	65	89
All	All	2849/2920 (98%)	2813 (99%)	36 (1%)	69	91

5 of 36 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	O	45	ASP
5	T	206	ARG
5	O	159	LEU
4	S	106	THR
1	P	154	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
5	J	193	GLN
4	N	38	ASN
5	T	98	GLN
2	G	64	GLN
2	G	92	GLN

### 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](#)

6 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	0.40	0	17,19,21	1.43	2 (11%)
6	NAG	U	2	6	14,14,15	0.29	0	17,19,21	0.76	1 (5%)
6	NAG	V	1	6,1	14,14,15	0.96	1 (7%)	17,19,21	1.39	3 (17%)
6	NAG	V	2	6	14,14,15	0.47	0	17,19,21	0.37	0
6	NAG	W	1	6,1	14,14,15	0.39	0	17,19,21	1.48	2 (11%)
6	NAG	W	2	6	14,14,15	0.33	0	17,19,21	0.58	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	-	3/6/23/26	0/1/1/1
6	NAG	V	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	V	2	6	-	0/6/23/26	0/1/1/1
6	NAG	W	1	6,1	-	2/6/23/26	0/1/1/1
6	NAG	W	2	6	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	V	1	NAG	O5-C1	2.95	1.48	1.43

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	U	1	NAG	C1-O5-C5	4.51	118.31	112.19
6	W	1	NAG	C1-O5-C5	4.27	117.98	112.19
6	V	1	NAG	C1-O5-C5	3.93	117.51	112.19
6	W	1	NAG	O4-C4-C5	3.66	118.37	109.30
6	U	1	NAG	O4-C4-C5	3.17	117.16	109.30

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

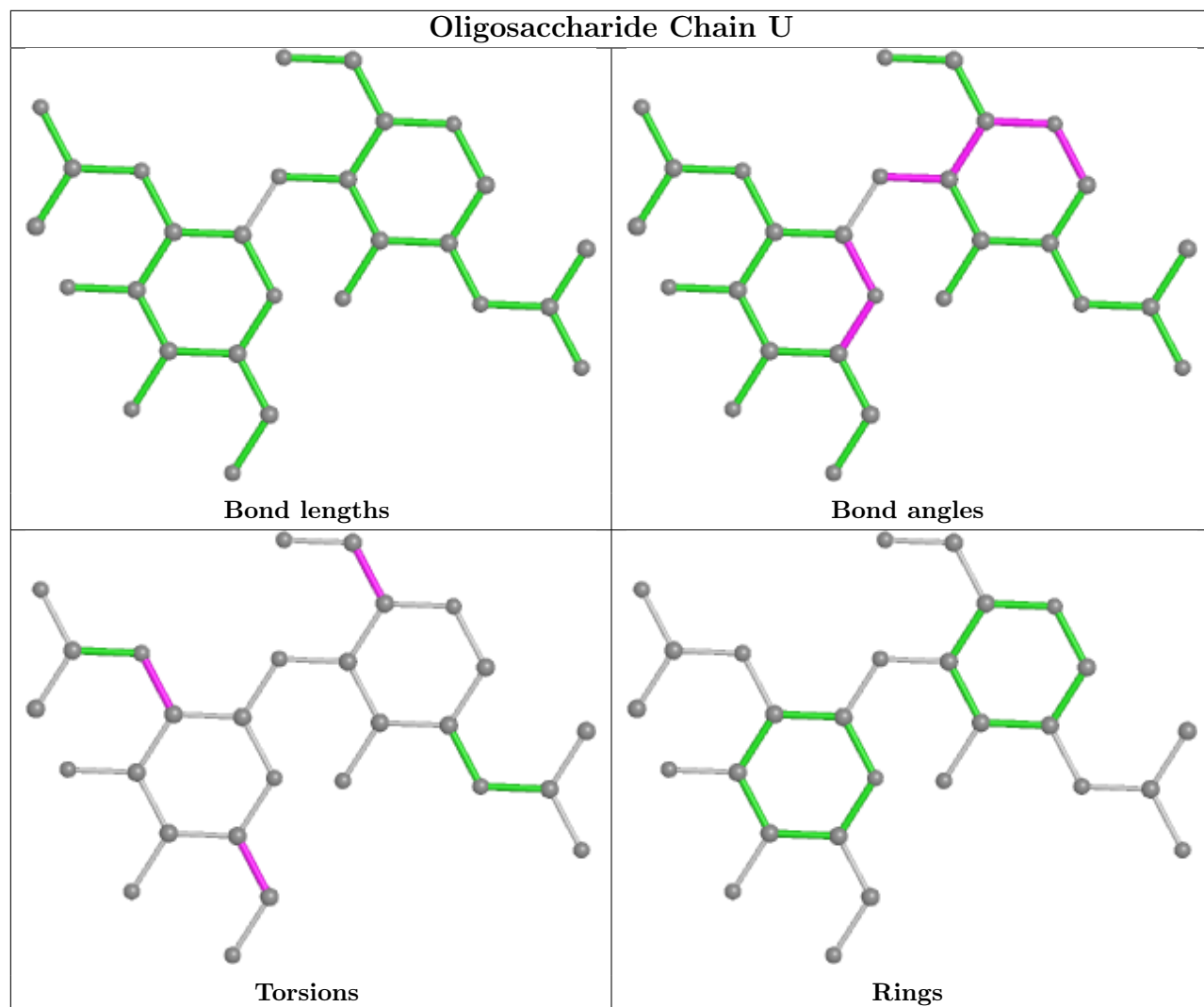
Mol	Chain	Res	Type	Atoms
6	W	1	NAG	O5-C5-C6-O6
6	W	1	NAG	C4-C5-C6-O6
6	W	2	NAG	C8-C7-N2-C2
6	W	2	NAG	O7-C7-N2-C2
6	V	1	NAG	O5-C5-C6-O6

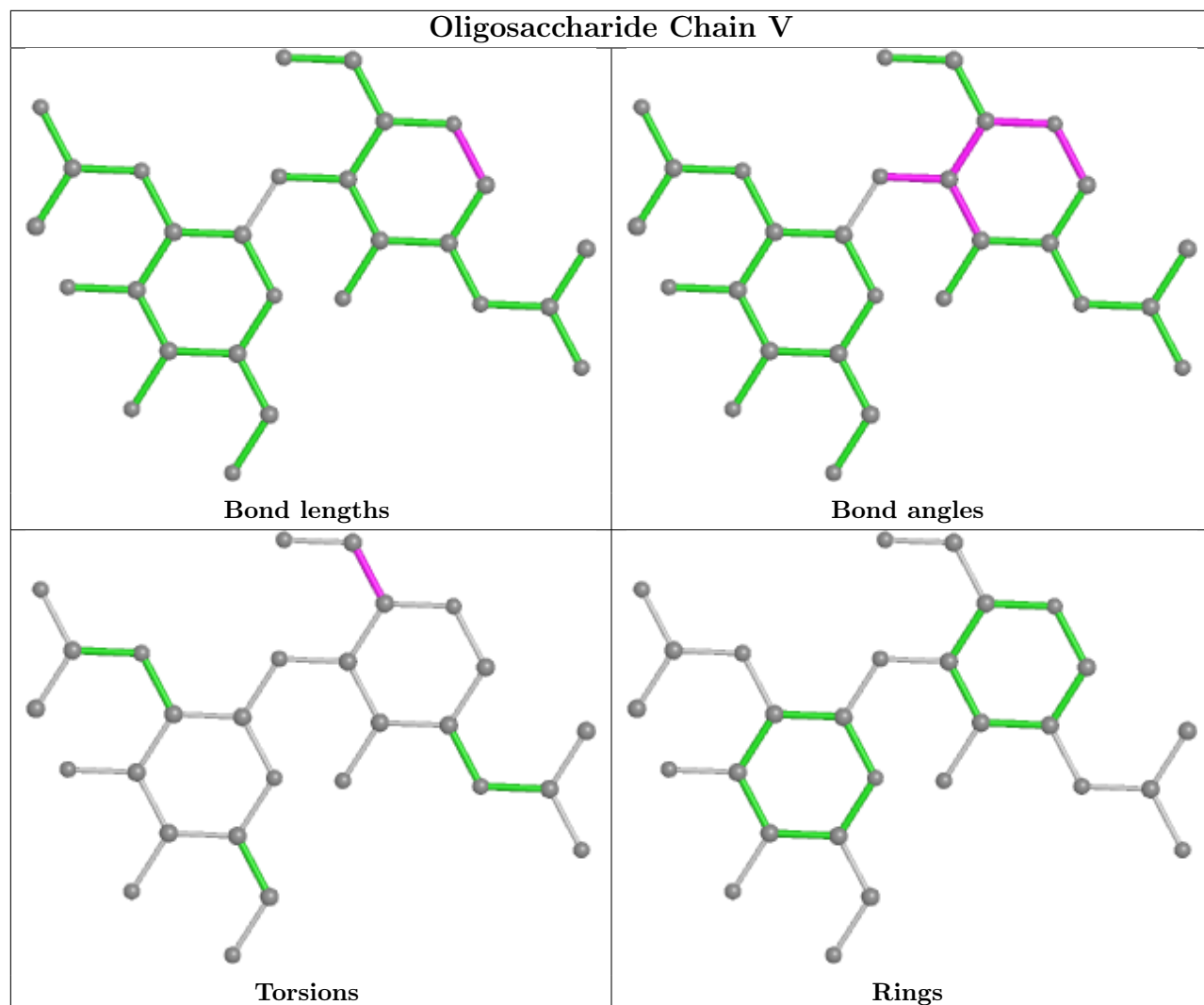
There are no ring outliers.

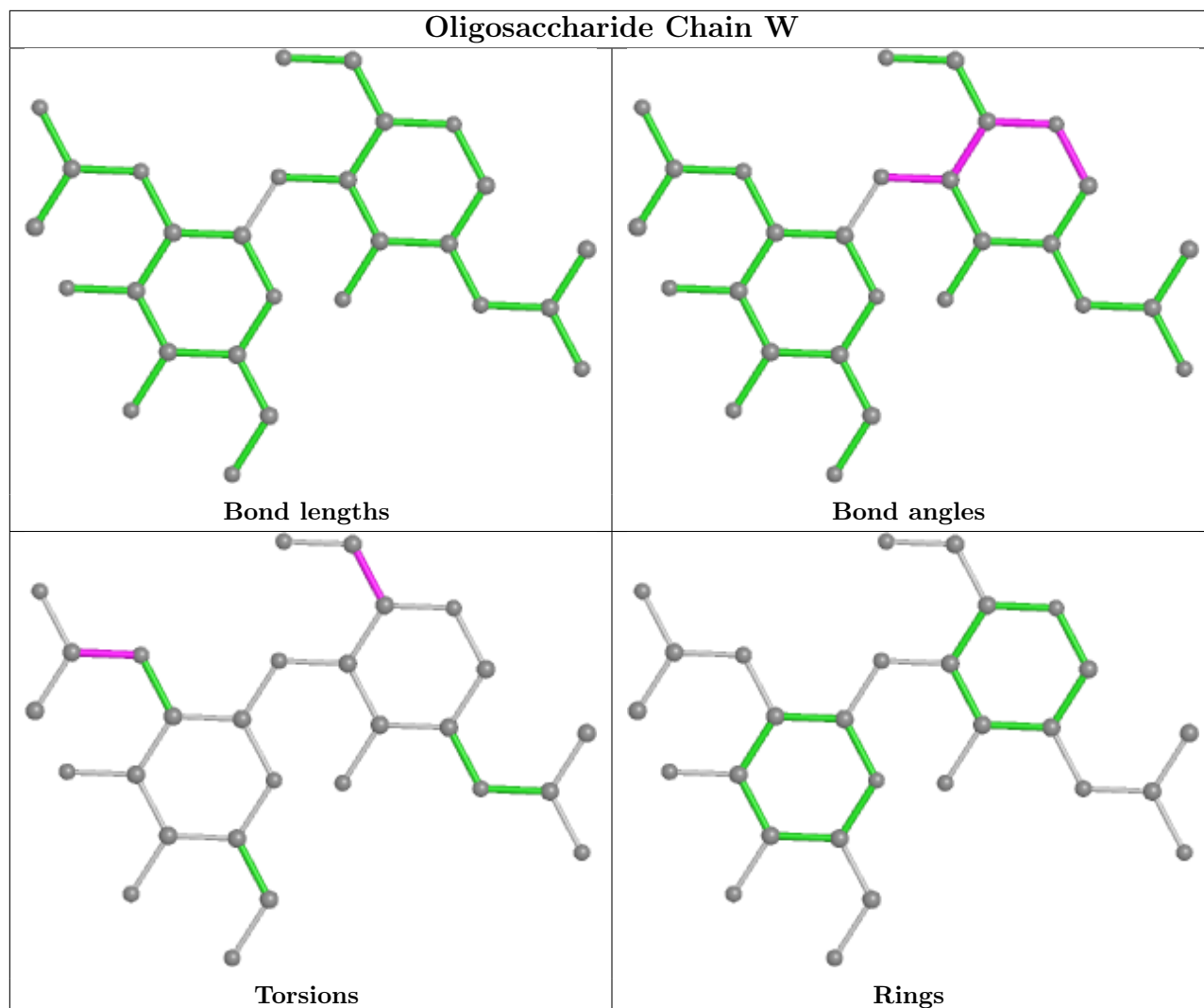
3 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	U	2	NAG	1	0
6	V	1	NAG	1	0
6	U	1	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](#)

17 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	$\# Z  > 2$
8	EDO	P	202	-	3,3,3	0.45	0	2,2,2	0.37	0
7	NAG	F	201	1	14,14,15	0.30	0	17,19,21	0.47	0
7	NAG	K	201	1	14,14,15	0.24	0	17,19,21	0.42	0
8	EDO	N	301	-	3,3,3	0.44	0	2,2,2	0.38	0



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
8	EDO	B	202	-	3,3,3	0.10	0	2,2,2	0.20	0
7	NAG	A	201	1	14,14,15	0.32	0	17,19,21	0.49	0
8	EDO	S	301	-	3,3,3	0.08	0	2,2,2	0.36	0
8	EDO	F	202	-	3,3,3	0.45	0	2,2,2	0.33	0
8	EDO	N	302	-	3,3,3	0.08	0	2,2,2	0.14	0
7	NAG	B	201	2	14,14,15	0.28	0	17,19,21	0.45	0
8	EDO	E	301	-	3,3,3	0.42	0	2,2,2	0.43	0
7	NAG	G	201	2	14,14,15	0.26	0	17,19,21	0.35	0
7	NAG	K	202	1	14,14,15	0.32	0	17,19,21	0.48	0
8	EDO	J	301	-	3,3,3	0.42	0	2,2,2	0.40	0
7	NAG	P	201	1	14,14,15	0.33	0	17,19,21	0.45	0
7	NAG	Q	201	2	14,14,15	0.28	0	17,19,21	0.38	0
8	EDO	A	202	-	3,3,3	0.45	0	2,2,2	0.36	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
8	EDO	P	202	-	-	0/1/1/1	-
7	NAG	F	201	1	-	1/6/23/26	0/1/1/1
7	NAG	K	201	1	-	0/6/23/26	0/1/1/1
8	EDO	N	301	-	-	0/1/1/1	-
8	EDO	B	202	-	-	0/1/1/1	-
7	NAG	A	201	1	-	0/6/23/26	0/1/1/1
8	EDO	S	301	-	-	0/1/1/1	-
8	EDO	F	202	-	-	0/1/1/1	-
8	EDO	N	302	-	-	0/1/1/1	-
7	NAG	B	201	2	-	0/6/23/26	0/1/1/1
8	EDO	E	301	-	-	0/1/1/1	-
7	NAG	G	201	2	-	0/6/23/26	0/1/1/1
7	NAG	K	202	1	-	1/6/23/26	0/1/1/1
8	EDO	J	301	-	-	0/1/1/1	-
7	NAG	P	201	1	-	0/6/23/26	0/1/1/1
7	NAG	Q	201	2	-	0/6/23/26	0/1/1/1
8	EDO	A	202	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	K	202	NAG	O5-C5-C6-O6
7	F	201	NAG	C4-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	K	201	NAG	1	0
8	F	202	EDO	1	0
8	N	302	EDO	2	0
8	J	301	EDO	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%)	0.12	3 (1%) 70 63	40, 57, 96, 121	0
1	F	180/181 (99%)	0.09	0 100 100	40, 51, 87, 138	0
1	K	180/181 (99%)	0.15	2 (1%) 80 75	54, 73, 106, 146	0
1	P	180/181 (99%)	0.06	3 (1%) 70 63	40, 59, 105, 147	0
2	B	180/190 (94%)	0.69	26 (14%) 2 1	42, 78, 152, 174	0
2	G	185/190 (97%)	0.19	6 (3%) 47 37	40, 69, 116, 220	0
2	L	179/190 (94%)	0.36	11 (6%) 21 13	59, 88, 133, 171	0
2	Q	189/190 (99%)	0.27	8 (4%) 36 26	43, 74, 127, 172	0
3	C	14/15 (93%)	0.19	0 100 100	43, 52, 91, 98	0
3	H	14/15 (93%)	0.10	0 100 100	45, 54, 102, 104	0
3	M	13/15 (86%)	0.33	0 100 100	60, 66, 113, 123	0
3	R	14/15 (93%)	0.09	0 100 100	44, 51, 96, 104	0
4	D	196/207 (94%)	0.18	6 (3%) 49 39	40, 60, 111, 132	0
4	I	190/207 (91%)	0.67	22 (11%) 4 2	46, 91, 134, 184	0
4	N	197/207 (95%)	-0.02	1 (0%) 91 88	44, 66, 91, 114	0
4	S	197/207 (95%)	0.20	5 (2%) 57 47	46, 68, 109, 140	0
5	E	239/241 (99%)	0.04	4 (1%) 70 63	40, 57, 104, 146	0
5	J	239/241 (99%)	0.10	2 (0%) 86 81	46, 69, 112, 178	0
5	O	239/241 (99%)	0.25	5 (2%) 63 54	44, 69, 119, 136	0
5	T	239/241 (99%)	-0.04	2 (0%) 86 81	45, 62, 102, 128	0
All	All	3244/3336 (97%)	0.20	106 (3%) 46 36	40, 68, 119, 220	0

The worst 5 of 106 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	141	GLY	4.7
2	G	108	PRO	4.5
2	B	168	GLY	4.3
4	I	151	CYS	4.0
2	Q	108	PRO	3.9

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

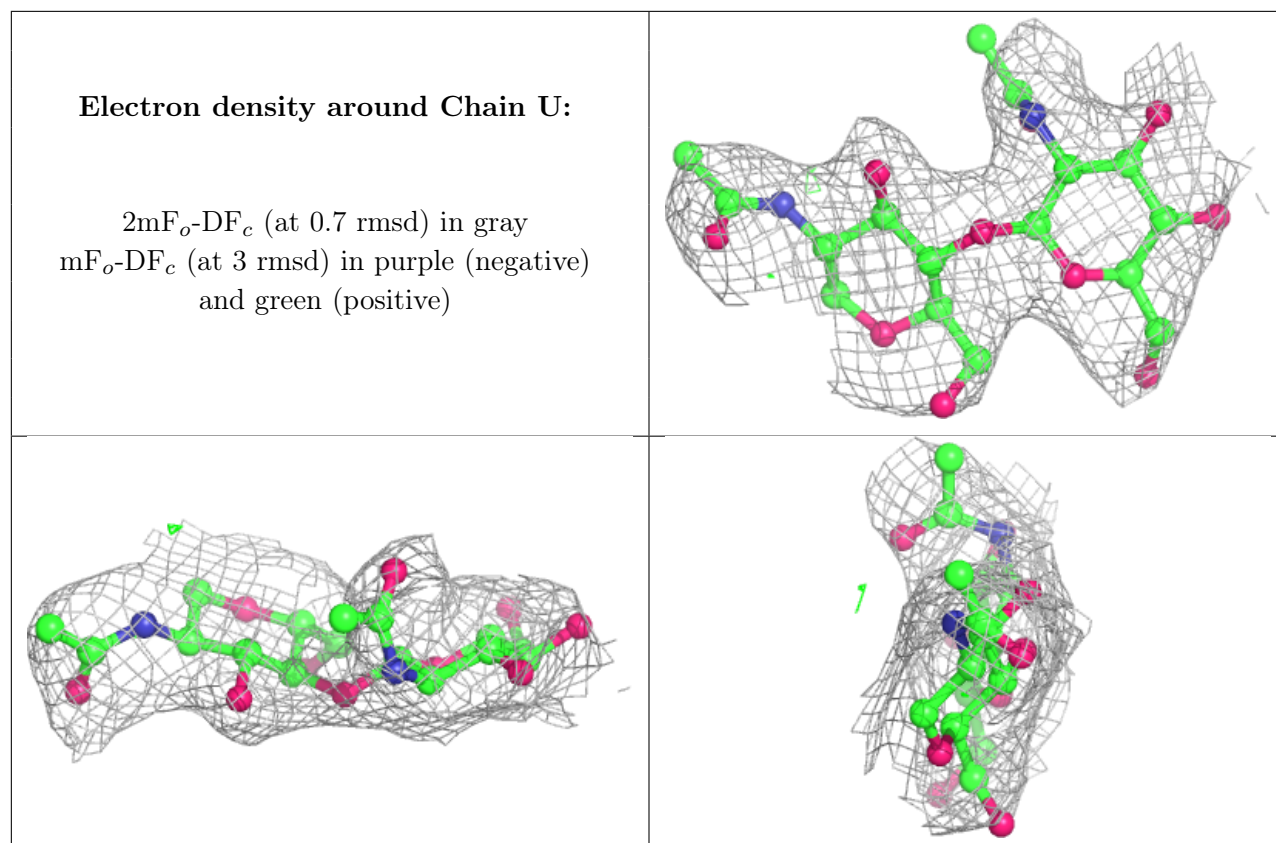
There are no non-standard protein/DNA/RNA residues in this entry.

## 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.

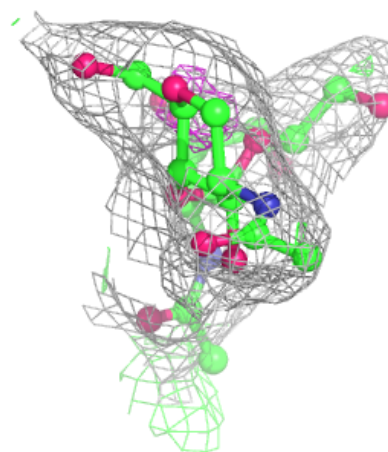
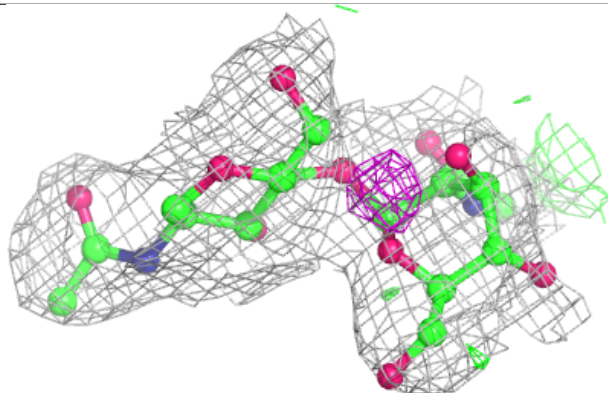
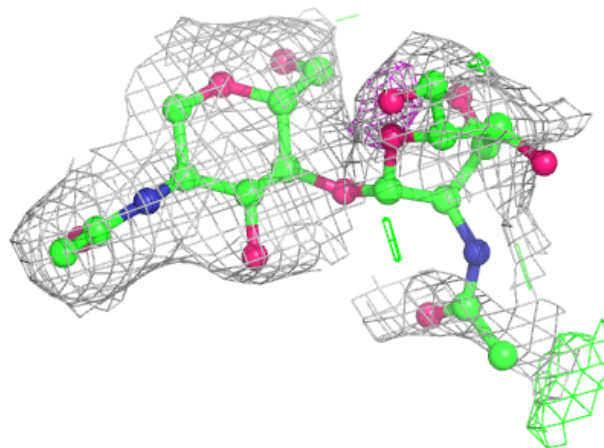
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
6	NAG	V	2	14/15	0.79	0.20	97,116,126,128	0
6	NAG	W	2	14/15	0.82	0.24	96,119,134,136	0
6	NAG	U	2	14/15	0.86	0.20	96,108,113,119	0
6	NAG	U	1	14/15	0.93	0.17	54,79,98,104	0
6	NAG	V	1	14/15	0.93	0.18	49,70,91,101	0
6	NAG	W	1	14/15	0.95	0.17	59,71,94,103	0

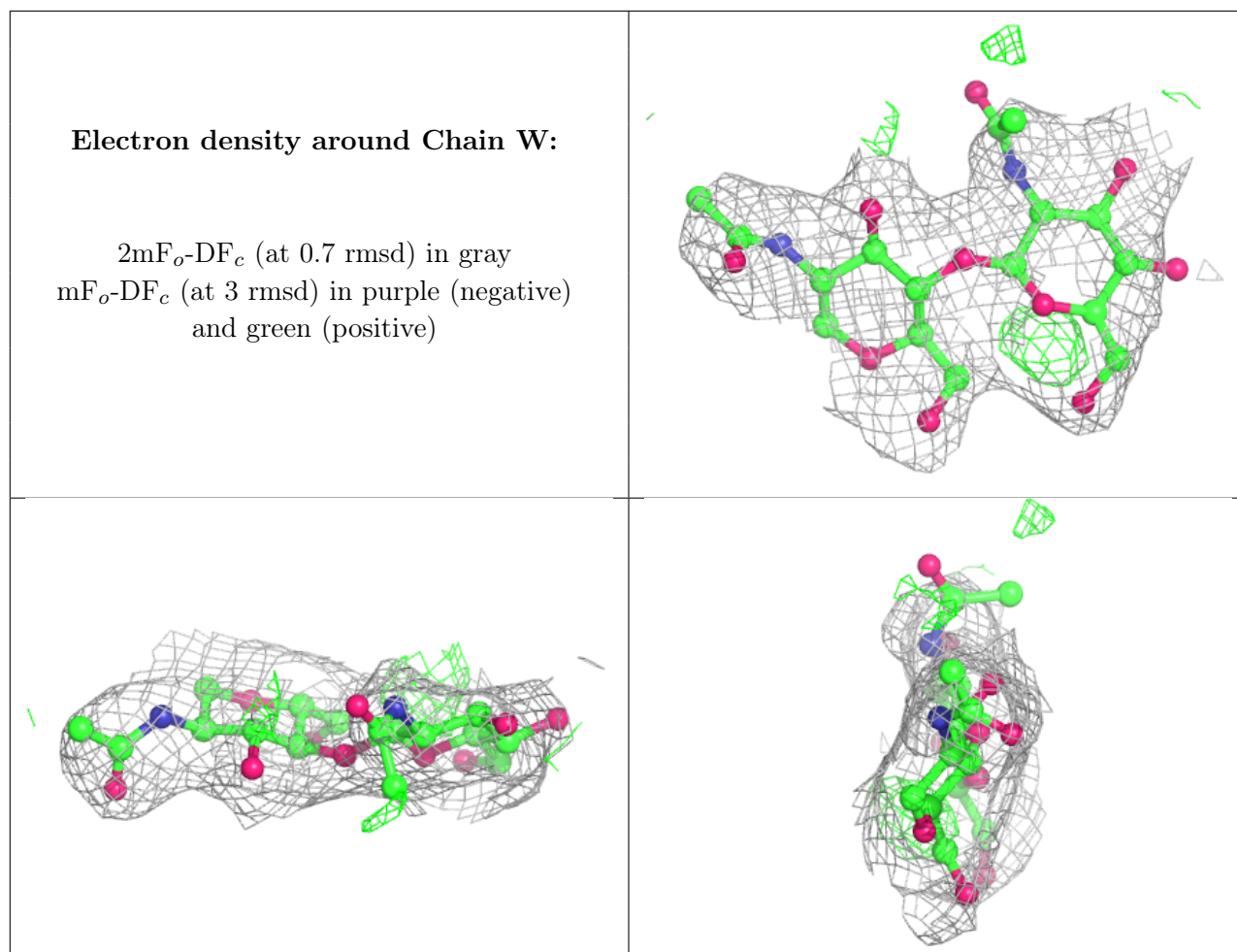
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.



**Electron density around Chain V:**

$2mF_o-DF_c$  (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
7	NAG	K	201	14/15	0.65	0.22	122,156,168,176	0
7	NAG	G	201	14/15	0.69	0.32	102,131,139,145	0
7	NAG	F	201	14/15	0.72	0.23	98,113,125,133	0
7	NAG	Q	201	14/15	0.77	0.24	82,111,122,133	0
7	NAG	B	201	14/15	0.80	0.27	80,112,120,122	0
8	EDO	E	301	4/4	0.81	0.22	56,56,63,75	0
8	EDO	A	202	4/4	0.82	0.26	68,73,76,77	0
8	EDO	P	202	4/4	0.85	0.26	61,62,66,70	0
8	EDO	S	301	4/4	0.86	0.44	20,20,20,20	0
8	EDO	J	301	4/4	0.87	0.24	61,62,63,69	0
8	EDO	N	302	4/4	0.88	0.45	20,20,20,20	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
7	NAG	A	201	14/15	0.89	0.19	89,97,102,103	0
7	NAG	K	202	14/15	0.90	0.19	71,90,109,118	0
8	EDO	B	202	4/4	0.93	0.46	20,20,20,20	0
7	NAG	P	201	14/15	0.93	0.13	79,94,101,101	0
8	EDO	F	202	4/4	0.94	0.22	59,59,62,70	0
8	EDO	N	301	4/4	0.95	0.20	52,55,66,74	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.