



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2024 – 10:27 am GMT

PDB ID : 8RKG  
Title : Crystal structure of tetrameric collagenase-cleaved *Xenopus* ZP2-N2N3 (cleaved xZP2-N2N3)  
Authors : Nishio, S.; de Sanctis, D.; Jovine, L.  
Deposited on : 2023-12-25  
Resolution : 2.90 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 : 1.8.4, 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

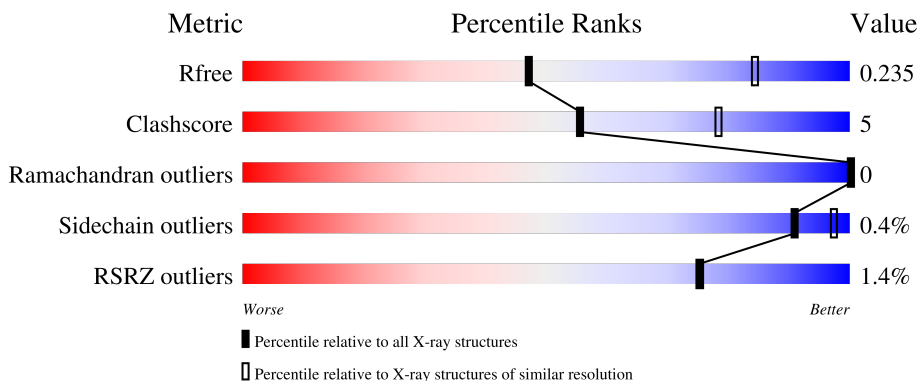
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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	1957 (2.90-2.90)
Clashscore	141614	2172 (2.90-2.90)
Ramachandran outliers	138981	2115 (2.90-2.90)
Sidechain outliers	138945	2117 (2.90-2.90)
RSRZ outliers	127900	1906 (2.90-2.90)

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	L	31	
1	M	31	
1	N	31	
1	O	31	
1	P	31	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	Q	31	<p>3% 65% 10% 26%</p>
1	R	31	<p>74% 6% 19%</p>
1	S	31	<p>81% 10% 10%</p>
2	A	188	<p>81% 9% 10%</p>
2	B	188	<p>75% 14% 11%</p>
2	C	188	<p>% 80% 10% 10%</p>
2	D	188	<p>3% 77% 11% 13%</p>
2	E	188	<p>% 80% 10% 10%</p>
2	F	188	<p>2% 78% 11% 11%</p>
2	G	188	<p>3% 76% 15% 10%</p>
2	H	188	<p>% 71% 16% 13%</p>
3	I	2	<p>50% 50%</p>
3	J	2	<p>100%</p>
3	K	2	<p>100%</p>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 12316 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 XIZPA protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	L	25	Total 184	C 111	N 32	O 39	S 2	0	0	0
1	M	22	Total 166	C 102	N 29	O 33	S 2	0	0	0
1	N	22	Total 166	C 102	N 29	O 33	S 2	0	0	0
1	O	28	Total 208	C 127	N 35	O 44	S 2	0	0	0
1	P	25	Total 189	C 114	N 32	O 41	S 2	0	0	0
1	Q	23	Total 174	C 106	N 30	O 36	S 2	0	0	0
1	R	25	Total 191	C 117	N 32	O 40	S 2	0	0	0
1	S	28	Total 208	C 127	N 35	O 44	S 2	0	0	0

- Molecule 2 is a protein called XIZPA protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	170	Total 1337	C 853	N 227	O 251	S 6	0	0	0
2	B	168	Total 1317	C 839	N 224	O 248	S 6	0	0	0
2	C	169	Total 1333	C 851	N 226	O 250	S 6	0	0	0
2	D	164	Total 1292	C 826	N 217	O 243	S 6	0	0	0
2	E	170	Total 1337	C 853	N 227	O 251	S 6	0	0	0
2	F	168	Total 1324	C 846	N 225	O 247	S 6	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	G	170	1339	854	227	252	6	0	0	0
2	H	164	1297	829	220	242	6	0	0	0

There are 80 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	339	LEU	-	expression tag	UNP A1L3D9
A	340	GLU	-	expression tag	UNP A1L3D9
A	341	HIS	-	expression tag	UNP A1L3D9
A	342	HIS	-	expression tag	UNP A1L3D9
A	343	HIS	-	expression tag	UNP A1L3D9
A	344	HIS	-	expression tag	UNP A1L3D9
A	345	HIS	-	expression tag	UNP A1L3D9
A	346	HIS	-	expression tag	UNP A1L3D9
A	347	HIS	-	expression tag	UNP A1L3D9
A	348	HIS	-	expression tag	UNP A1L3D9
B	339	LEU	-	expression tag	UNP A1L3D9
B	340	GLU	-	expression tag	UNP A1L3D9
B	341	HIS	-	expression tag	UNP A1L3D9
B	342	HIS	-	expression tag	UNP A1L3D9
B	343	HIS	-	expression tag	UNP A1L3D9
B	344	HIS	-	expression tag	UNP A1L3D9
B	345	HIS	-	expression tag	UNP A1L3D9
B	346	HIS	-	expression tag	UNP A1L3D9
B	347	HIS	-	expression tag	UNP A1L3D9
B	348	HIS	-	expression tag	UNP A1L3D9
C	339	LEU	-	expression tag	UNP A1L3D9
C	340	GLU	-	expression tag	UNP A1L3D9
C	341	HIS	-	expression tag	UNP A1L3D9
C	342	HIS	-	expression tag	UNP A1L3D9
C	343	HIS	-	expression tag	UNP A1L3D9
C	344	HIS	-	expression tag	UNP A1L3D9
C	345	HIS	-	expression tag	UNP A1L3D9
C	346	HIS	-	expression tag	UNP A1L3D9
C	347	HIS	-	expression tag	UNP A1L3D9
C	348	HIS	-	expression tag	UNP A1L3D9
D	339	LEU	-	expression tag	UNP A1L3D9
D	340	GLU	-	expression tag	UNP A1L3D9
D	341	HIS	-	expression tag	UNP A1L3D9
D	342	HIS	-	expression tag	UNP A1L3D9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	343	HIS	-	expression tag	UNP A1L3D9
D	344	HIS	-	expression tag	UNP A1L3D9
D	345	HIS	-	expression tag	UNP A1L3D9
D	346	HIS	-	expression tag	UNP A1L3D9
D	347	HIS	-	expression tag	UNP A1L3D9
D	348	HIS	-	expression tag	UNP A1L3D9
E	339	LEU	-	expression tag	UNP A1L3D9
E	340	GLU	-	expression tag	UNP A1L3D9
E	341	HIS	-	expression tag	UNP A1L3D9
E	342	HIS	-	expression tag	UNP A1L3D9
E	343	HIS	-	expression tag	UNP A1L3D9
E	344	HIS	-	expression tag	UNP A1L3D9
E	345	HIS	-	expression tag	UNP A1L3D9
E	346	HIS	-	expression tag	UNP A1L3D9
E	347	HIS	-	expression tag	UNP A1L3D9
E	348	HIS	-	expression tag	UNP A1L3D9
F	339	LEU	-	expression tag	UNP A1L3D9
F	340	GLU	-	expression tag	UNP A1L3D9
F	341	HIS	-	expression tag	UNP A1L3D9
F	342	HIS	-	expression tag	UNP A1L3D9
F	343	HIS	-	expression tag	UNP A1L3D9
F	344	HIS	-	expression tag	UNP A1L3D9
F	345	HIS	-	expression tag	UNP A1L3D9
F	346	HIS	-	expression tag	UNP A1L3D9
F	347	HIS	-	expression tag	UNP A1L3D9
F	348	HIS	-	expression tag	UNP A1L3D9
G	339	LEU	-	expression tag	UNP A1L3D9
G	340	GLU	-	expression tag	UNP A1L3D9
G	341	HIS	-	expression tag	UNP A1L3D9
G	342	HIS	-	expression tag	UNP A1L3D9
G	343	HIS	-	expression tag	UNP A1L3D9
G	344	HIS	-	expression tag	UNP A1L3D9
G	345	HIS	-	expression tag	UNP A1L3D9
G	346	HIS	-	expression tag	UNP A1L3D9
G	347	HIS	-	expression tag	UNP A1L3D9
G	348	HIS	-	expression tag	UNP A1L3D9
H	339	LEU	-	expression tag	UNP A1L3D9
H	340	GLU	-	expression tag	UNP A1L3D9
H	341	HIS	-	expression tag	UNP A1L3D9
H	342	HIS	-	expression tag	UNP A1L3D9
H	343	HIS	-	expression tag	UNP A1L3D9
H	344	HIS	-	expression tag	UNP A1L3D9

*Continued on next page...*

Continued from previous page...

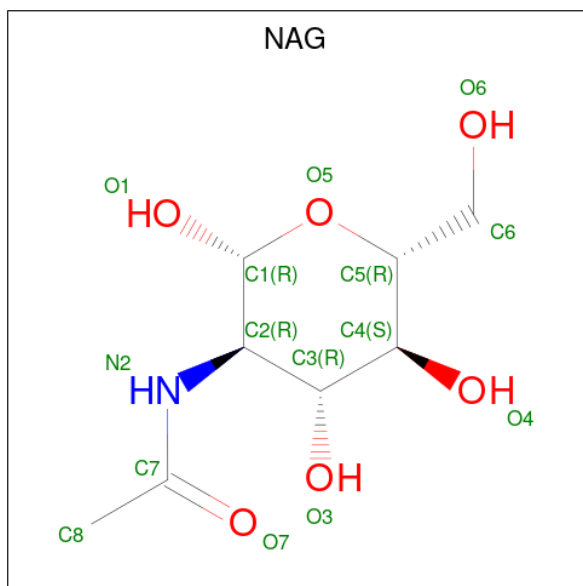
Chain	Residue	Modelled	Actual	Comment	Reference
H	345	HIS	-	expression tag	UNP A1L3D9
H	346	HIS	-	expression tag	UNP A1L3D9
H	347	HIS	-	expression tag	UNP A1L3D9
H	348	HIS	-	expression tag	UNP A1L3D9

- Molecule 3 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			
3	I	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	J	2	Total	C	N	O	0	0	0
			28	16	2	10			
3	K	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 4 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>).



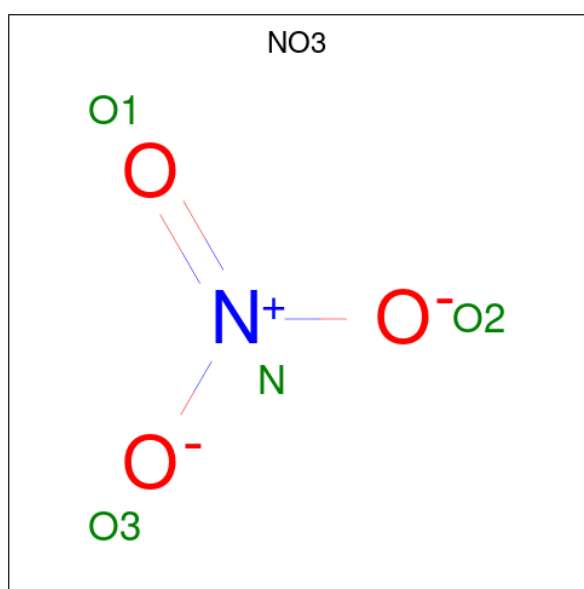
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	A	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	B	1	Total 14	C 8	N 1	O 5	0	0
4	D	1	Total 14	C 8	N 1	O 5	0	0
4	E	1	Total 14	C 8	N 1	O 5	0	0
4	H	1	Total 14	C 8	N 1	O 5	0	0

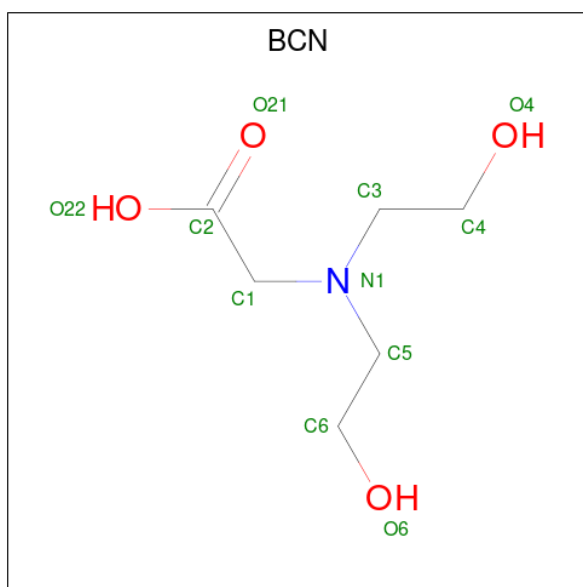
- Molecule 5 is NITRATE ION (three-letter code: NO3) (formula: NO<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	N	O		
5	C	1	Total 4	N 1	O 3	0	0
5	O	1	Total 4	N 1	O 3	0	0
5	E	1	Total 4	N 1	O 3	0	0
5	G	1	Total 4	N 1	O 3	0	0
5	H	1	Total 4	N 1	O 3	0	0
5	H	1	Total 4	N 1	O 3	0	0

- Molecule 6 is BICINE (three-letter code: BCN) (formula: C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>).





Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	C	1	Total	C	N	O	0	0
			11	6	1	4		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	L	1	Total	O	0	0
			1	1		
7	A	10	Total	O	0	0
			10	10		
7	M	1	Total	O	0	0
			1	1		
7	B	4	Total	O	0	0
			4	4		
7	C	5	Total	O	0	0
			5	5		
7	D	8	Total	O	0	0
			8	8		
7	P	1	Total	O	0	0
			1	1		
7	E	7	Total	O	0	0
			7	7		
7	F	7	Total	O	0	0
			7	7		
7	R	1	Total	O	0	0
			1	1		
7	G	9	Total	O	0	0
			9	9		

*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>
7	H	11	Total	O	0	0
			11	11		

### 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: XIZPA protein

Chain L:  58% 19% 1% 19%



- Molecule 1: XIZPA protein

Chain M:  58% 13% 29%



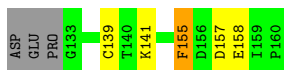
- Molecule 1: XIZPA protein

Chain N:  3% 65% 6% 29%




- Molecule 1: XIZPA protein

Chain O:  74% 13% 1% 10%



- Molecule 1: XIZPA protein

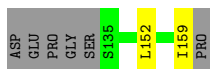
Chain P:  3% 74% 6% 19%



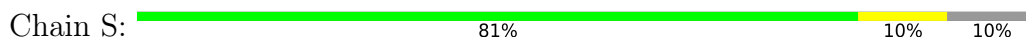
- Molecule 1: XIZPA protein



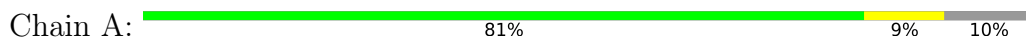
- Molecule 1: XIZPA protein



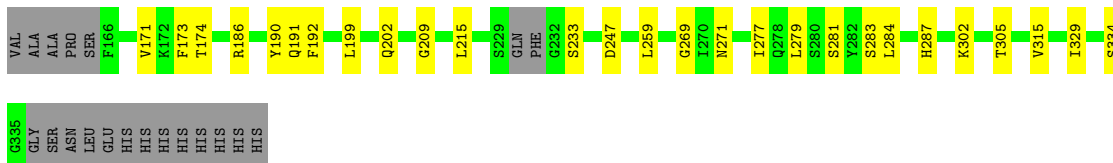
- Molecule 1: XIZPA protein



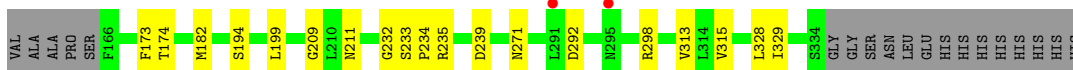
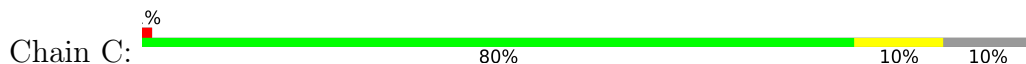
- Molecule 2: XIZPA protein



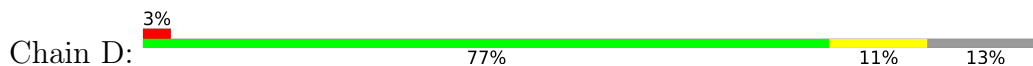
- Molecule 2: XIZPA protein

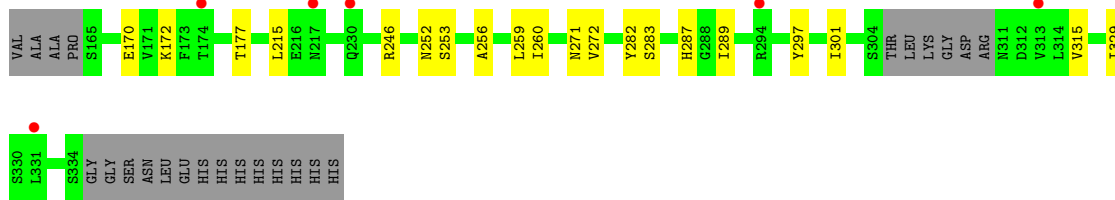


- Molecule 2: XIZPA protein

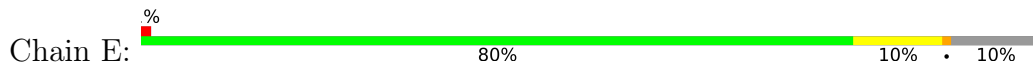


- Molecule 2: XIZPA protein

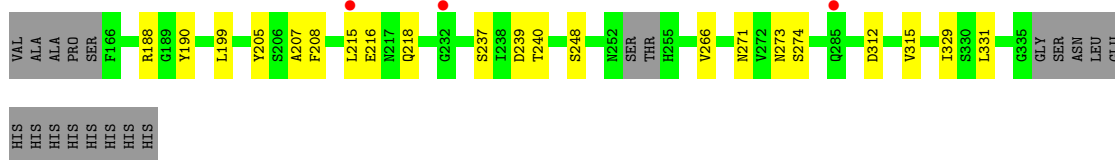
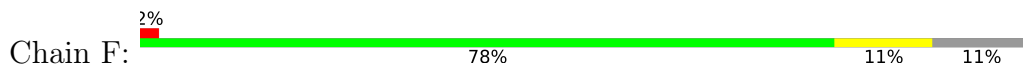




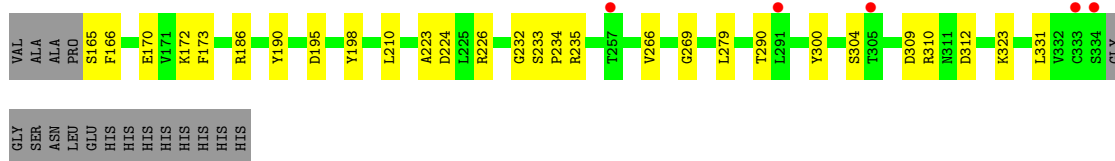
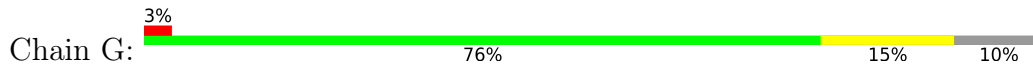
• Molecule 2: XIZPA protein



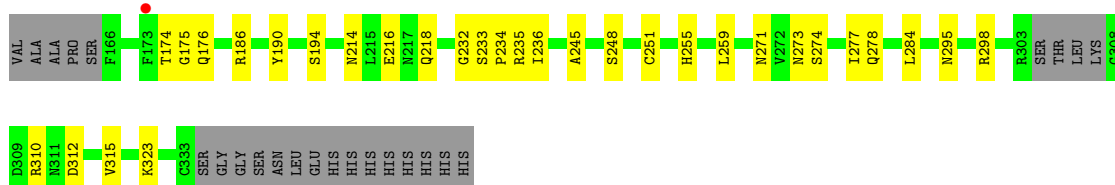
• Molecule 2: XIZPA protein



• Molecule 2: XIZPA protein



• Molecule 2: XIZPA protein



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



MAG1  
MAG2

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

Chain J:  100%

MAG1  
MAG2

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

Chain K:  100%

MAG1  
MAG2

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	88.21Å 70.31Å 149.86Å 90.00° 90.04° 90.00°	Depositor
Resolution (Å)	38.00 – 2.90 38.00 – 2.90	Depositor EDS
% Data completeness (in resolution range)	99.1 (38.00-2.90) 99.0 (38.00-2.90)	Depositor EDS
$R_{merge}$	0.23	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.29 (at 2.90Å)	Xtrriage
Refinement program	PHENIX 1.19.2_4158+SVN	Depositor
R, $R_{free}$	0.212 , 0.247 0.213 , 0.235	Depositor DCC
$R_{free}$ test set	1670 reflections (4.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	52.9	Xtrriage
Anisotropy	0.218	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 40.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.358 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12316	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	48.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, NO3, BCN

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	L	0.29	0/185	0.58	0/249
1	M	0.25	0/167	0.60	0/225
1	N	0.29	0/167	0.58	0/225
1	O	0.31	0/210	0.59	0/284
1	P	0.28	0/190	0.58	0/256
1	Q	0.28	0/175	0.59	0/236
1	R	0.28	0/192	0.58	0/259
1	S	0.28	0/210	0.57	0/284
2	A	0.32	0/1365	0.57	0/1852
2	B	0.33	0/1343	0.56	0/1821
2	C	0.30	0/1361	0.55	0/1847
2	D	0.31	0/1319	0.53	0/1790
2	E	0.31	0/1365	0.55	0/1852
2	F	0.31	0/1351	0.53	0/1831
2	G	0.30	0/1367	0.53	0/1855
2	H	0.31	0/1324	0.55	0/1796
All	All	0.31	0/12291	0.55	0/16662

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	L	184	0	182	6	0
1	M	166	0	170	6	0
1	N	166	0	170	3	0
1	O	208	0	206	5	0
1	P	189	0	185	2	0
1	Q	174	0	174	4	0
1	R	191	0	191	2	0
1	S	208	0	206	2	0
2	A	1337	0	1325	14	0
2	B	1317	0	1307	18	0
2	C	1333	0	1322	15	0
2	D	1292	0	1275	13	0
2	E	1337	0	1325	13	0
2	F	1324	0	1312	14	0
2	G	1339	0	1327	16	0
2	H	1297	0	1280	21	0
3	I	28	0	25	3	0
3	J	28	0	25	0	0
3	K	28	0	25	0	0
4	A	14	0	13	1	0
4	B	14	0	13	0	0
4	D	14	0	13	1	0
4	E	14	0	13	0	0
4	H	14	0	13	0	0
5	C	4	0	0	0	0
5	E	4	0	0	0	0
5	G	4	0	0	0	0
5	H	8	0	0	1	0
5	O	4	0	0	0	0
6	C	11	0	12	1	0
7	A	10	0	0	1	0
7	B	4	0	0	0	0
7	C	5	0	0	0	0
7	D	8	0	0	0	0
7	E	7	0	0	0	0
7	F	7	0	0	0	0
7	G	9	0	0	0	0
7	H	11	0	0	0	0
7	L	1	0	0	0	0
7	M	1	0	0	0	0
7	P	1	0	0	0	0
7	R	1	0	0	2	0
All	All	12316	0	12109	123	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 5.

The worst 5 of 123 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:L:136:VAL:HG23	2:A:326:PRO:HB2	1.62	0.80
2:H:310:ARG:NH1	2:H:312:ASP:OD1	2.24	0.70
2:A:166:PHE:N	1:M:156:ASP:OD1	2.25	0.69
2:B:283:SER:O	2:B:287:HIS:ND1	2.21	0.65
2:H:278:GLN:NE2	5:H:902:NO3:O1	2.29	0.65

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	L	23/31 (74%)	23 (100%)	0	0	100	100
1	M	20/31 (64%)	19 (95%)	1 (5%)	0	100	100
1	N	20/31 (64%)	20 (100%)	0	0	100	100
1	O	26/31 (84%)	25 (96%)	1 (4%)	0	100	100
1	P	23/31 (74%)	23 (100%)	0	0	100	100
1	Q	21/31 (68%)	21 (100%)	0	0	100	100
1	R	23/31 (74%)	23 (100%)	0	0	100	100
1	S	26/31 (84%)	23 (88%)	3 (12%)	0	100	100
2	A	168/188 (89%)	163 (97%)	5 (3%)	0	100	100
2	B	164/188 (87%)	158 (96%)	6 (4%)	0	100	100
2	C	167/188 (89%)	162 (97%)	5 (3%)	0	100	100
2	D	160/188 (85%)	149 (93%)	11 (7%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	E	168/188 (89%)	161 (96%)	7 (4%)	0	100	100
2	F	164/188 (87%)	155 (94%)	9 (6%)	0	100	100
2	G	168/188 (89%)	163 (97%)	5 (3%)	0	100	100
2	H	160/188 (85%)	151 (94%)	9 (6%)	0	100	100
All	All	1501/1752 (86%)	1439 (96%)	62 (4%)	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	L	23/29 (79%)	22 (96%)	1 (4%)	29	62
1	M	21/29 (72%)	21 (100%)	0	100	100
1	N	21/29 (72%)	21 (100%)	0	100	100
1	O	26/29 (90%)	25 (96%)	1 (4%)	33	67
1	P	24/29 (83%)	24 (100%)	0	100	100
1	Q	22/29 (76%)	22 (100%)	0	100	100
1	R	24/29 (83%)	24 (100%)	0	100	100
1	S	26/29 (90%)	26 (100%)	0	100	100
2	A	151/166 (91%)	151 (100%)	0	100	100
2	B	149/166 (90%)	149 (100%)	0	100	100
2	C	151/166 (91%)	150 (99%)	1 (1%)	84	95
2	D	147/166 (89%)	147 (100%)	0	100	100
2	E	151/166 (91%)	150 (99%)	1 (1%)	84	95
2	F	149/166 (90%)	149 (100%)	0	100	100
2	G	152/166 (92%)	150 (99%)	2 (1%)	69	90
2	H	146/166 (88%)	146 (100%)	0	100	100
All	All	1383/1560 (89%)	1377 (100%)	6 (0%)	91	97

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

Mol	Chain	Res	Type
2	E	290	THR
2	G	173	PHE
2	G	304	SER
2	C	239	ASP
1	L	135	SER

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

### 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	$\# Z  > 2$
3	NAG	I	1	3,2	14,14,15	0.23	0	17,19,21	0.60	0
3	NAG	I	2	3	14,14,15	0.35	0	17,19,21	0.85	1 (5%)
3	NAG	J	1	3,2	14,14,15	0.20	0	17,19,21	0.54	0
3	NAG	J	2	3	14,14,15	0.27	0	17,19,21	0.53	0
3	NAG	K	1	3,2	14,14,15	0.15	0	17,19,21	0.59	0
3	NAG	K	2	3	14,14,15	0.26	0	17,19,21	0.55	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	I	1	3,2	-	3/6/23/26	0/1/1/1
3	NAG	I	2	3	-	4/6/23/26	0/1/1/1
3	NAG	J	1	3,2	-	2/6/23/26	0/1/1/1
3	NAG	J	2	3	-	0/6/23/26	0/1/1/1
3	NAG	K	1	3,2	-	0/6/23/26	0/1/1/1
3	NAG	K	2	3	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	2	NAG	C2-N2-C7	2.67	126.71	122.90

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

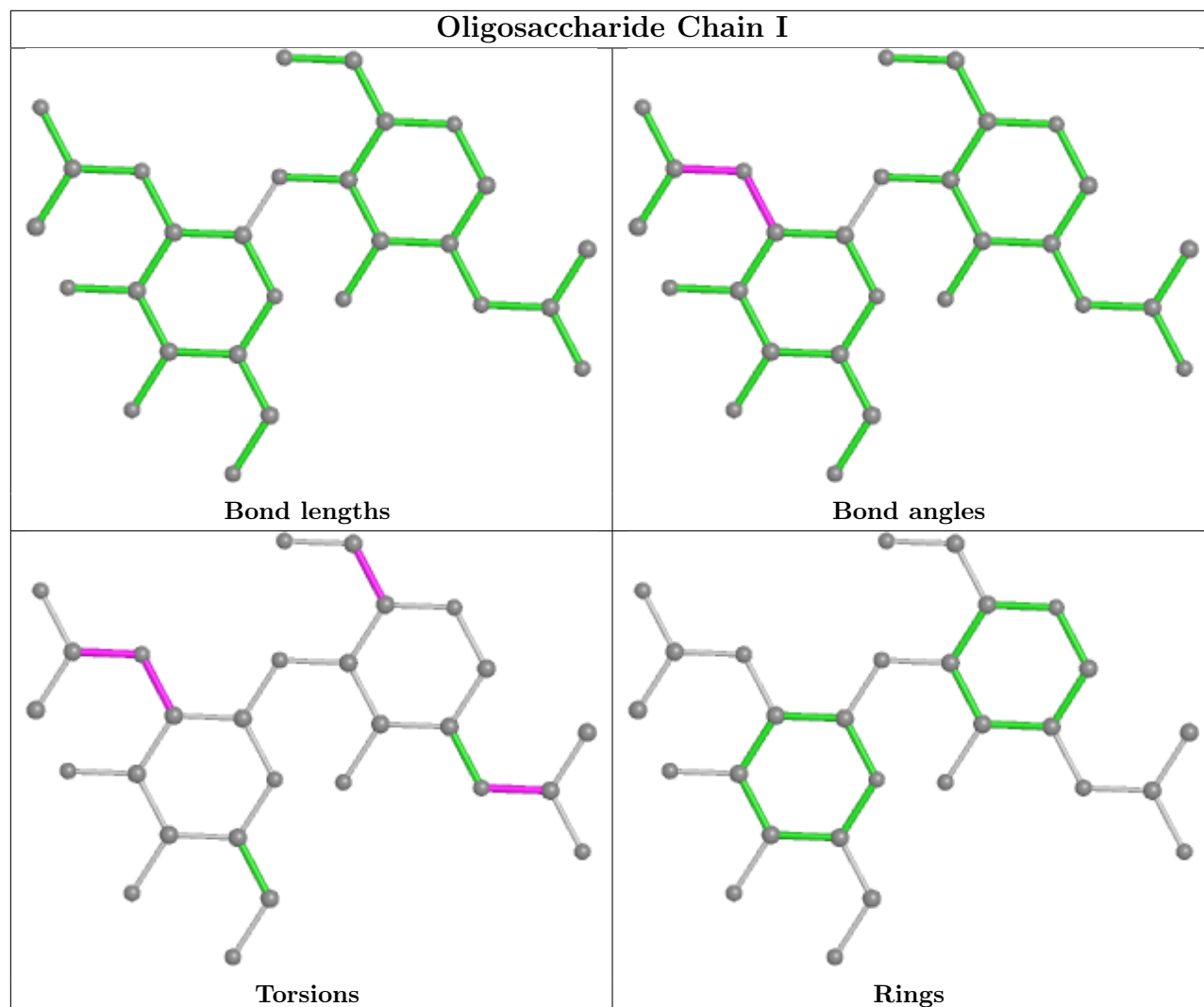
Mol	Chain	Res	Type	Atoms
3	I	1	NAG	C8-C7-N2-C2
3	I	1	NAG	O7-C7-N2-C2
3	I	2	NAG	C8-C7-N2-C2
3	I	2	NAG	O7-C7-N2-C2
3	J	1	NAG	C8-C7-N2-C2

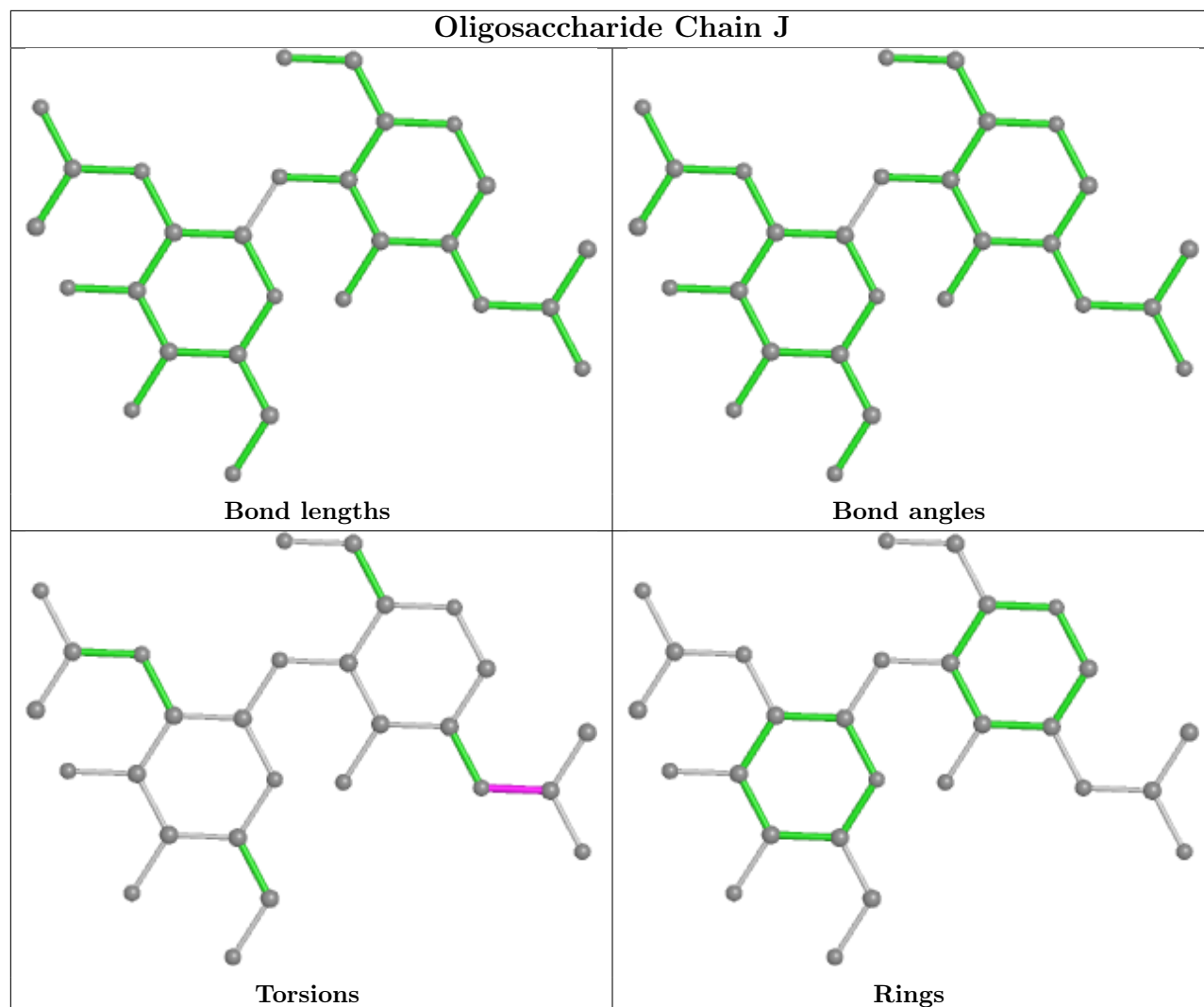
There are no ring outliers.

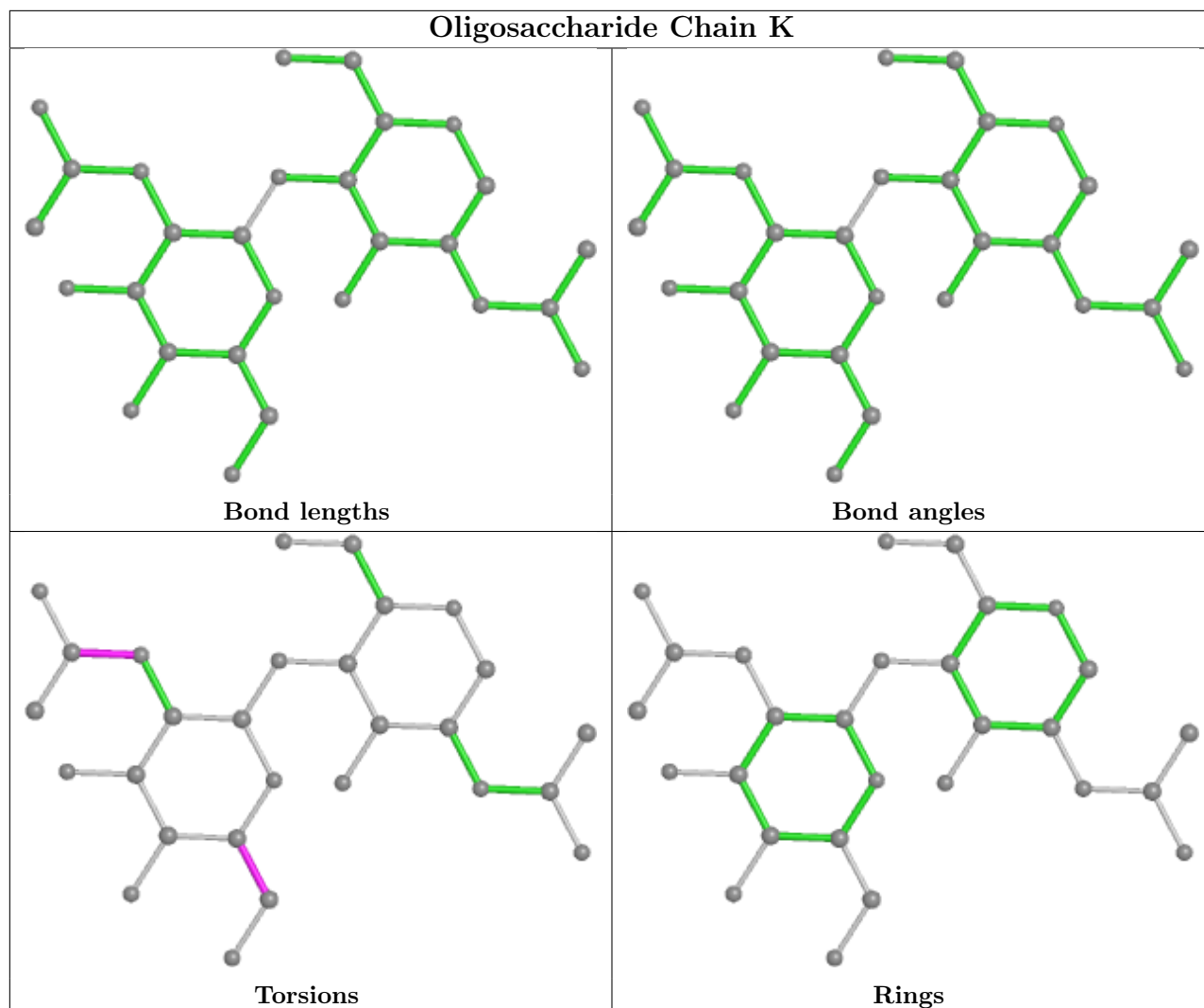
1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	I	2	NAG	3	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](#)

12 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
5	NO3	O	201	-	1,3,3	0.69	0	0,3,3	-	-
5	NO3	C	401	-	1,3,3	0.69	0	0,3,3	-	-
6	BCN	C	402	-	10,10,10	0.66	0	11,11,11	0.68	0
5	NO3	H	902	-	1,3,3	0.71	0	0,3,3	-	-



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	NAG	A	901	2	14,14,15	0.41	0	17,19,21	0.84	1 (5%)
4	NAG	H	901	2	14,14,15	0.29	0	17,19,21	0.72	1 (5%)
5	NO3	H	903	-	1,3,3	0.70	0	0,3,3	-	-
4	NAG	D	901	2	14,14,15	0.17	0	17,19,21	0.42	0
5	NO3	E	902	-	1,3,3	0.78	0	0,3,3	-	-
4	NAG	B	901	2	14,14,15	0.15	0	17,19,21	0.55	0
5	NO3	G	401	-	1,3,3	0.72	0	0,3,3	-	-
4	NAG	E	901	2	14,14,15	0.41	0	17,19,21	0.68	1 (5%)

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	BCN	C	402	-	-	2/10/10/10	-
4	NAG	A	901	2	-	4/6/23/26	0/1/1/1
4	NAG	H	901	2	-	0/6/23/26	0/1/1/1
4	NAG	D	901	2	-	2/6/23/26	0/1/1/1
4	NAG	B	901	2	-	4/6/23/26	0/1/1/1
4	NAG	E	901	2	-	4/6/23/26	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(°)
4	A	901	NAG	C1-O5-C5	2.78	115.97	112.19
4	H	901	NAG	C1-O5-C5	2.50	115.57	112.19
4	E	901	NAG	C1-O5-C5	2.16	115.11	112.19

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	901	NAG	C4-C5-C6-O6
4	D	901	NAG	O5-C5-C6-O6
4	B	901	NAG	C4-C5-C6-O6
4	B	901	NAG	O5-C5-C6-O6
4	E	901	NAG	O5-C5-C6-O6

There are no ring outliers.

4 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	402	BCN	1	0
5	H	902	NO3	1	0
4	A	901	NAG	1	0
4	D	901	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	L	25/31 (80%)	-0.20	0 <a href="#">100</a> <a href="#">100</a>	23, 38, 62, 74	0
1	M	22/31 (70%)	0.06	0 <a href="#">100</a> <a href="#">100</a>	35, 44, 55, 61	0
1	N	22/31 (70%)	0.09	1 (4%) <a href="#">33</a> <a href="#">29</a>	35, 49, 65, 74	0
1	O	28/31 (90%)	-0.13	0 <a href="#">100</a> <a href="#">100</a>	22, 44, 62, 73	0
1	P	25/31 (80%)	0.39	1 (4%) <a href="#">38</a> <a href="#">33</a>	33, 47, 70, 72	0
1	Q	23/31 (74%)	0.34	1 (4%) <a href="#">35</a> <a href="#">31</a>	36, 52, 72, 76	0
1	R	25/31 (80%)	0.02	0 <a href="#">100</a> <a href="#">100</a>	35, 49, 61, 79	0
1	S	28/31 (90%)	0.15	0 <a href="#">100</a> <a href="#">100</a>	36, 49, 76, 86	0
2	A	170/188 (90%)	-0.18	0 <a href="#">100</a> <a href="#">100</a>	18, 36, 54, 67	0
2	B	168/188 (89%)	-0.14	0 <a href="#">100</a> <a href="#">100</a>	21, 43, 77, 96	0
2	C	169/188 (89%)	0.01	2 (1%) <a href="#">79</a> <a href="#">79</a>	23, 44, 66, 88	0
2	D	164/188 (87%)	0.13	6 (3%) <a href="#">41</a> <a href="#">37</a>	29, 46, 73, 98	0
2	E	170/188 (90%)	0.08	2 (1%) <a href="#">79</a> <a href="#">79</a>	30, 48, 69, 80	0
2	F	168/188 (89%)	0.11	3 (1%) <a href="#">68</a> <a href="#">67</a>	29, 50, 74, 96	0
2	G	170/188 (90%)	0.34	5 (2%) <a href="#">51</a> <a href="#">47</a>	29, 53, 79, 97	0
2	H	164/188 (87%)	0.07	1 (0%) <a href="#">89</a> <a href="#">89</a>	30, 45, 64, 92	0
All	All	1541/1752 (87%)	0.06	22 (1%) <a href="#">75</a> <a href="#">75</a>	18, 46, 73, 98	0

The worst 5 of 22 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	232	GLY	4.3
2	G	334	SER	4.1
2	G	291	LEU	3.0
2	E	167	TRP	2.8
1	Q	154	GLY	2.6

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

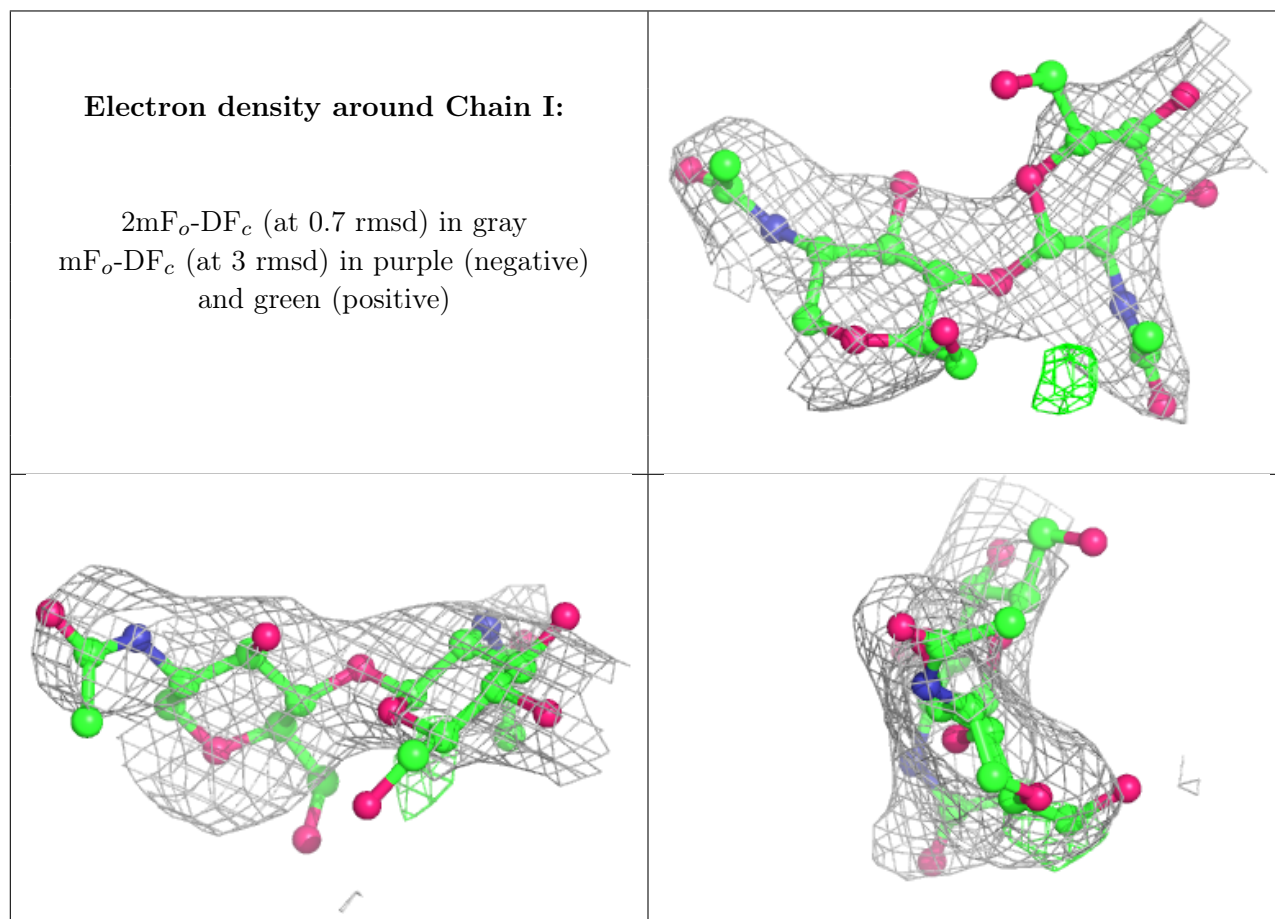
There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

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

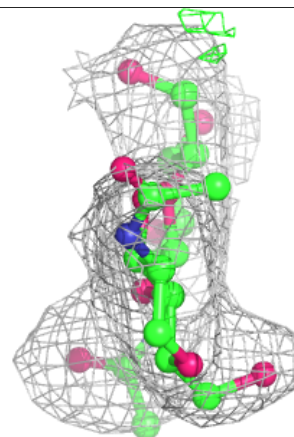
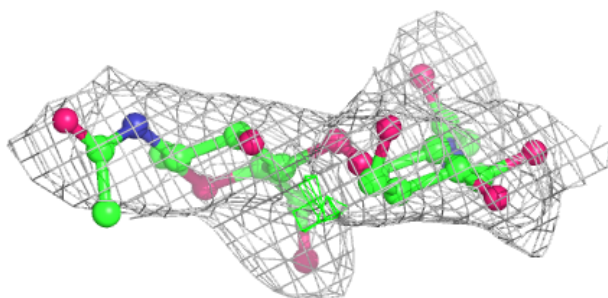
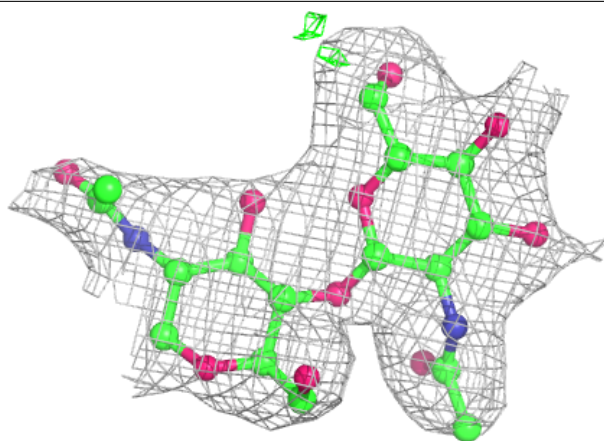
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	NAG	I	2	14/15	0.77	0.29	80,100,114,115	0
3	NAG	J	1	14/15	0.89	0.17	44,79,87,92	0
3	NAG	K	2	14/15	0.89	0.18	43,85,94,98	0
3	NAG	K	1	14/15	0.90	0.23	62,90,98,98	0
3	NAG	I	1	14/15	0.92	0.21	52,82,93,95	0
3	NAG	J	2	14/15	0.92	0.12	41,58,75,81	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.

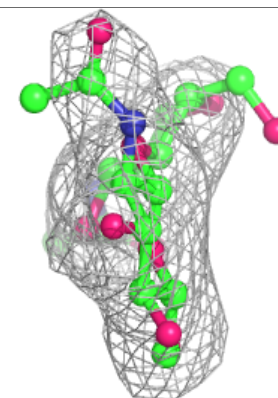
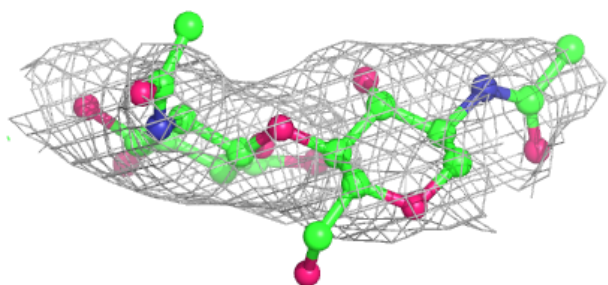
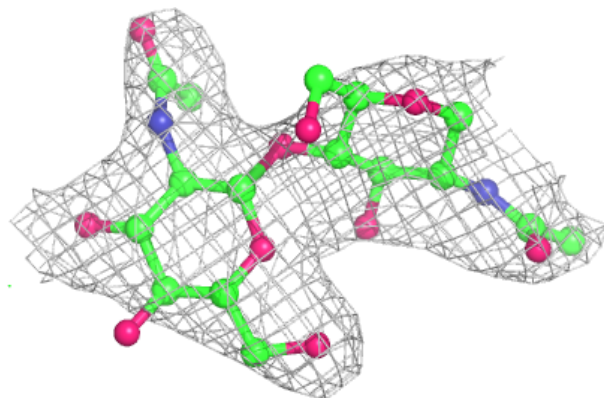


**Electron density around Chain J:**

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

**Electron density around Chain K:**

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



## 6.4 Ligands [i](#)

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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	NO3	O	201	4/4	0.87	0.15	60,61,62,76	0
4	NAG	D	901	14/15	0.90	0.16	49,69,94,104	0
6	BCN	C	402	11/11	0.90	0.27	44,52,65,70	0
4	NAG	H	901	14/15	0.92	0.13	28,61,70,89	0
4	NAG	A	901	14/15	0.93	0.16	29,48,62,66	0
4	NAG	E	901	14/15	0.94	0.14	45,58,71,75	0
5	NO3	G	401	4/4	0.94	0.15	51,53,63,71	0
5	NO3	H	903	4/4	0.94	0.11	46,46,51,57	0
4	NAG	B	901	14/15	0.94	0.17	54,61,75,119	0
5	NO3	E	902	4/4	0.95	0.17	35,36,60,84	0
5	NO3	C	401	4/4	0.98	0.10	40,47,49,52	0
5	NO3	H	902	4/4	0.98	0.20	31,47,51,64	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.