



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 8, 2020 – 01:18 PM BST

PDB ID : 2GJX  
Title : Crystallographic structure of human beta-Hexosaminidase A  
Authors : Lemieux, M.J.; Mark, B.L.; Cherney, M.M.; Withers, S.G.; Mahuran, D.J.;  
James, M.N.G.  
Deposited on : 2006-03-31  
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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.13.1  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.13.1

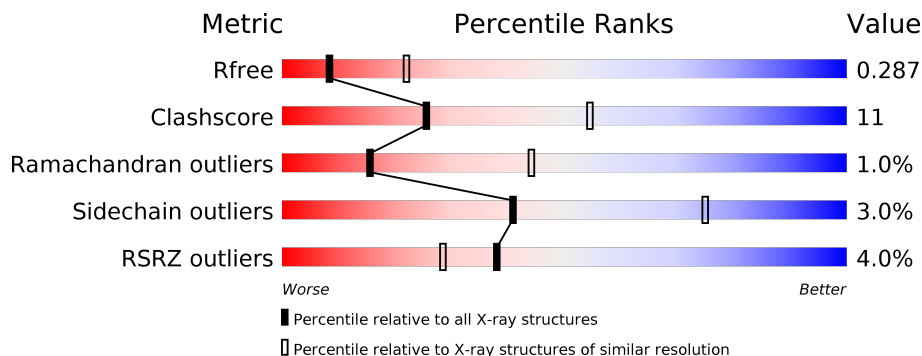
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	507	
1	D	507	
1	E	507	
1	H	507	
2	B	507	
2	C	507	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	507	 4% 73% 20% • 5%
2	G	507	 3% 76% 17% • 5%
3	I	3	 100%
3	J	3	 100%
3	K	3	 33% 67%
3	M	3	 33% 33% 33%
4	L	2	 100%
4	N	2	 50% 50%
4	O	2	 50% 50%
4	P	2	 50% 50%
4	Q	2	 50% 50%
4	R	2	 100%
4	S	2	 50% 50%

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
3	NAG	K	1	X	-	-	-
4	NAG	L	2	-	-	-	X
5	NAG	A	707	-	-	-	X
5	NAG	D	716	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 32132 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 Beta-hexosaminidase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	492	4010	2600	650	746	14	0	0	0
1	D	492	4010	2600	650	746	14	0	0	0
1	E	492	4010	2600	650	746	14	0	0	0
1	H	492	4010	2600	650	746	14	0	0	0

- Molecule 2 is a protein called Beta-hexosaminidase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	480	3877	2505	643	716	13	0	0	0
2	C	481	3885	2511	644	717	13	0	0	0
2	F	480	3877	2505	643	716	13	0	0	0
2	G	480	3877	2505	643	716	13	0	0	0

- Molecule 3 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
			Total	C	N	O			
3	I	3	39	22	2	15	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	J	3	Total	C	N	O	0	0	0
			39	22	2	15			
3	K	3	Total	C	N	O	0	0	0
			39	22	2	15			
3	M	3	Total	C	N	O	0	0	0
			39	22	2	15			

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



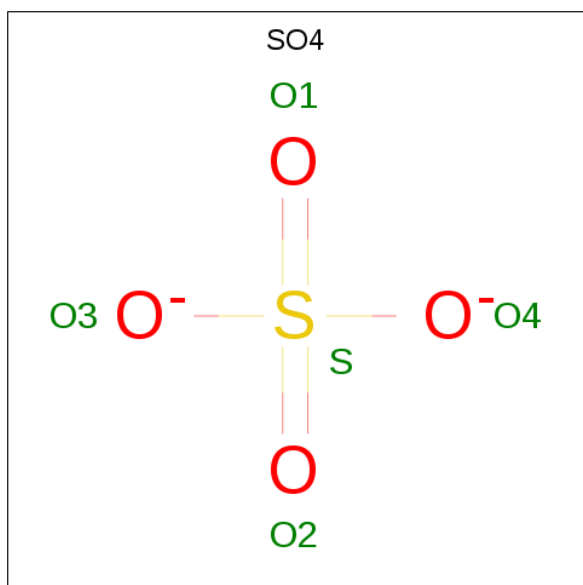
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	L	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	N	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	O	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	P	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	Q	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	R	2	Total	C	N	O	0	0	0
			28	16	2	10			
4	S	2	Total	C	N	O	0	0	0
			28	16	2	10			

- Molecule 5 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		
5	A	1	Total 14	C 8	N 1	O 5	0	0
5	D	1	Total 14	C 8	N 1	O 5	0	0
5	E	1	Total 14	C 8	N 1	O 5	0	0
5	G	1	Total 14	C 8	N 1	O 5	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	O	S	0	0
			5	4	1		
6	D	1	Total	O	S	0	0
			5	4	1		
6	H	1	Total	O	S	0	0
			5	4	1		
6	H	1	Total	O	S	0	0
			5	4	1		

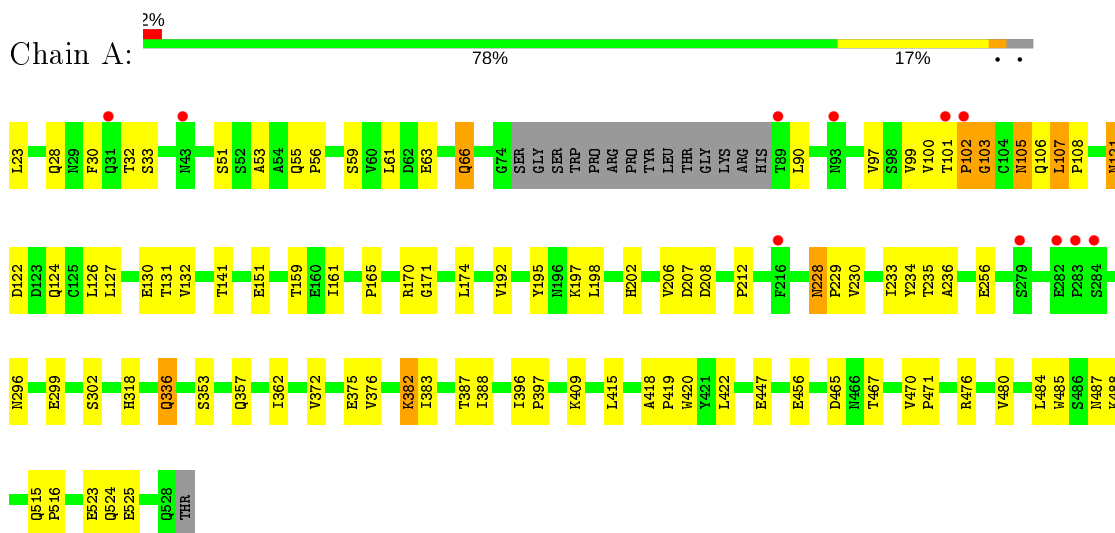
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	34	Total	O	0	0
			34	34		
7	B	20	Total	O	0	0
			20	20		
7	C	22	Total	O	0	0
			22	22		
7	D	18	Total	O	0	0
			18	18		
7	E	20	Total	O	0	0
			20	20		
7	F	11	Total	O	0	0
			11	11		
7	G	17	Total	O	0	0
			17	17		
7	H	6	Total	O	0	0
			6	6		

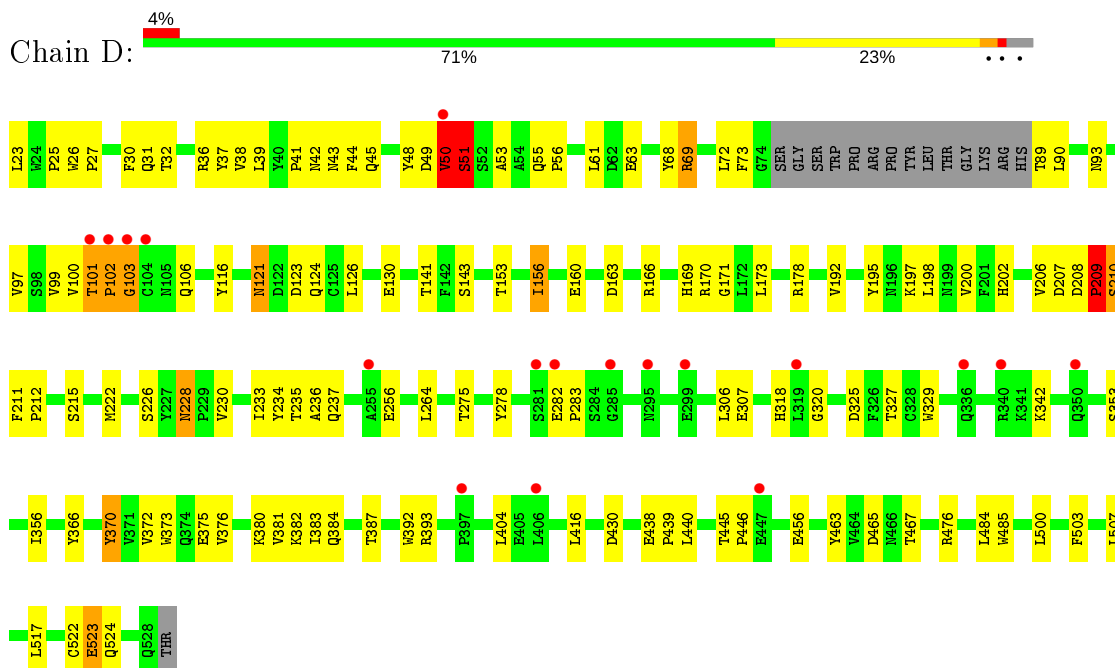
### 3 Residue-property plots [i](#)

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: Beta-hexosaminidase alpha chain

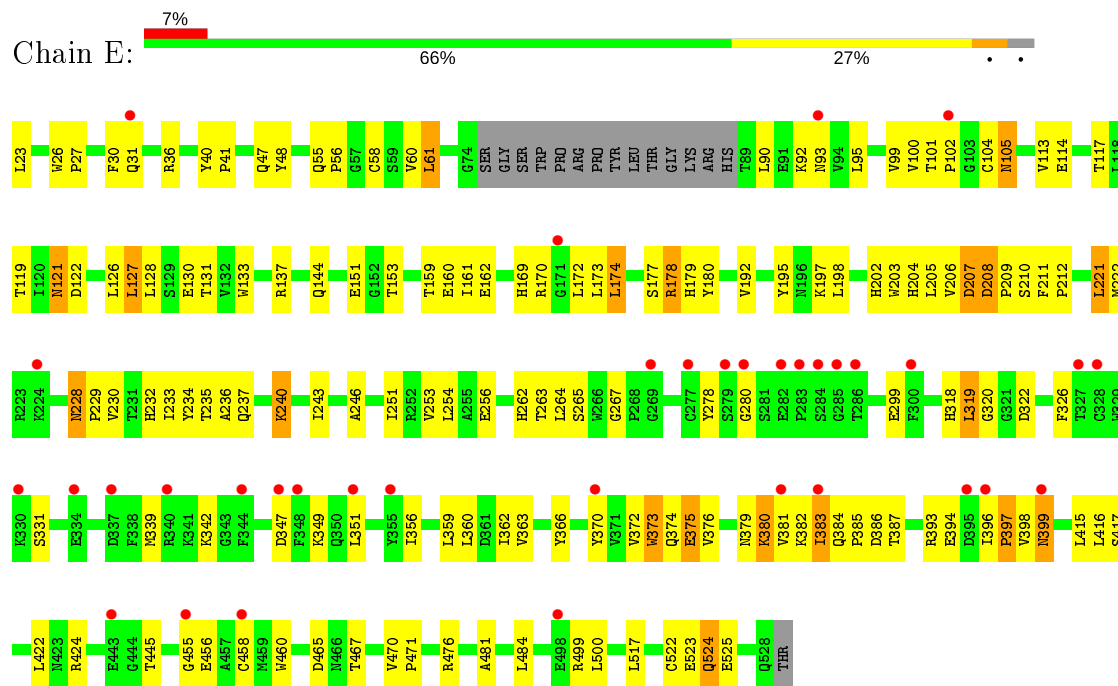


- Molecule 1: Beta-hexosaminidase alpha chain

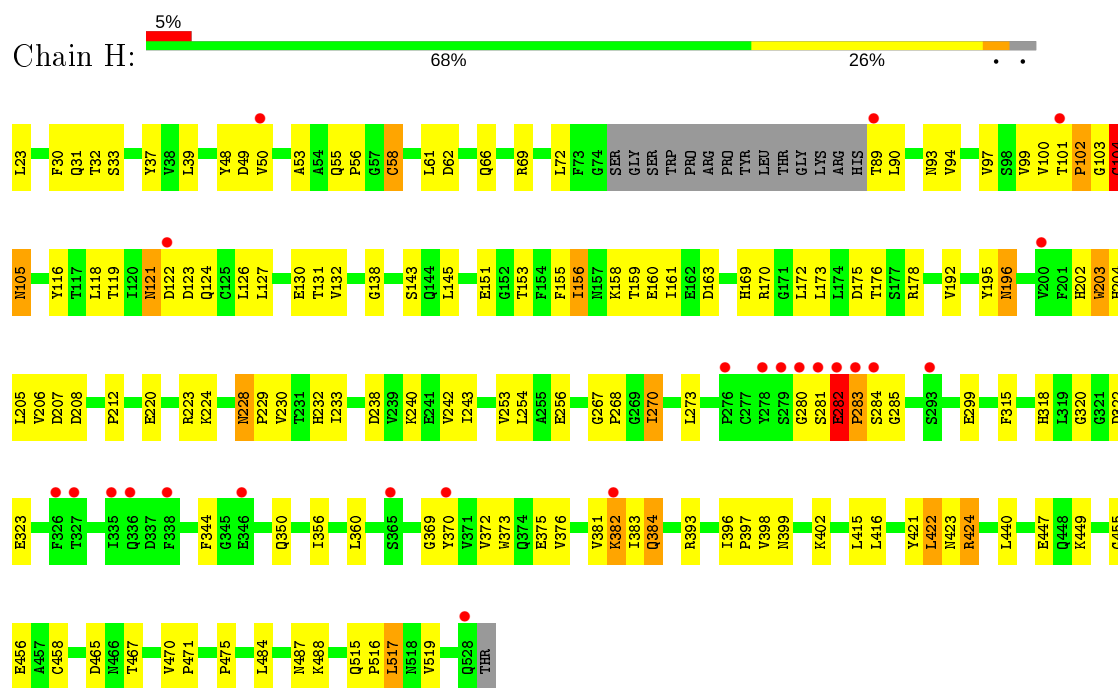




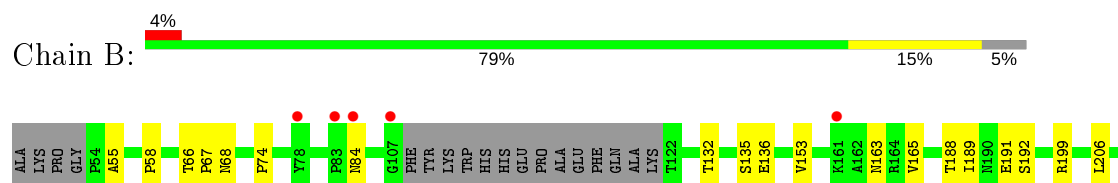
- Molecule 1: Beta-hexosaminidase alpha chain

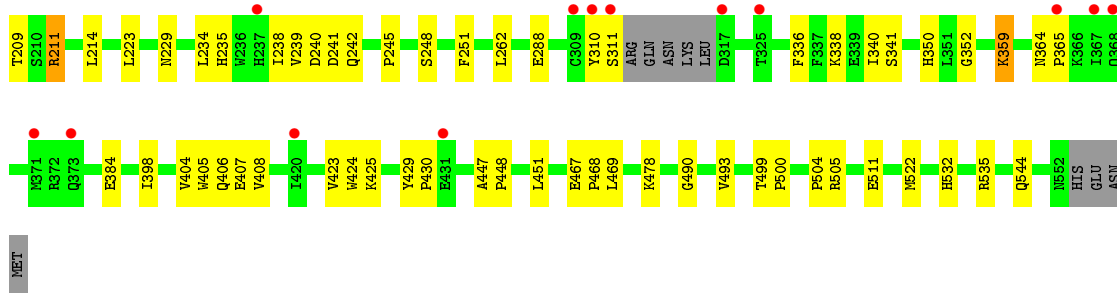


- Molecule 1: Beta-hexosaminidase alpha chain

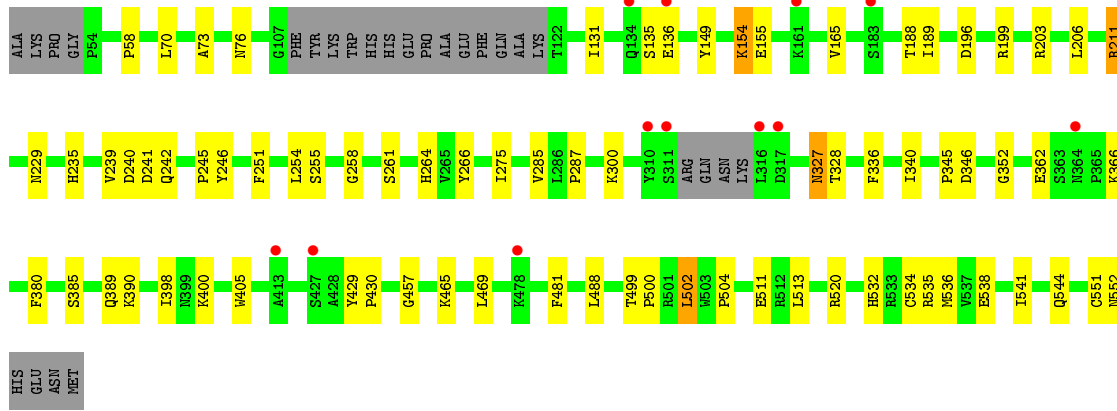
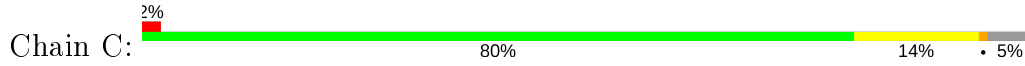


- Molecule 2: Beta-hexosaminidase beta chain

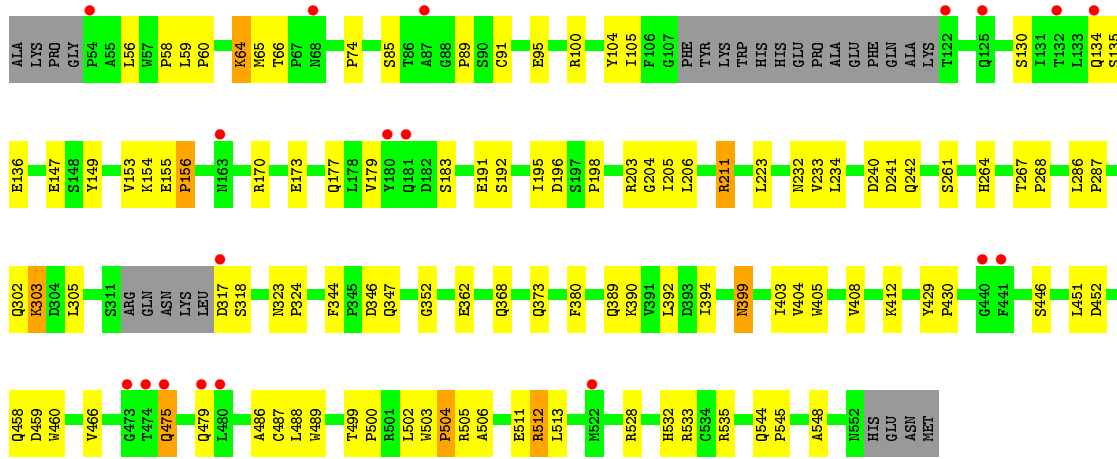




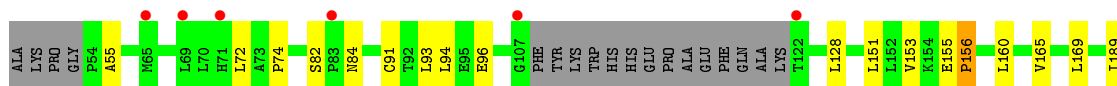
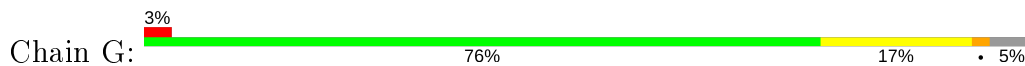
• Molecule 2: Beta-hexosaminidase beta chain

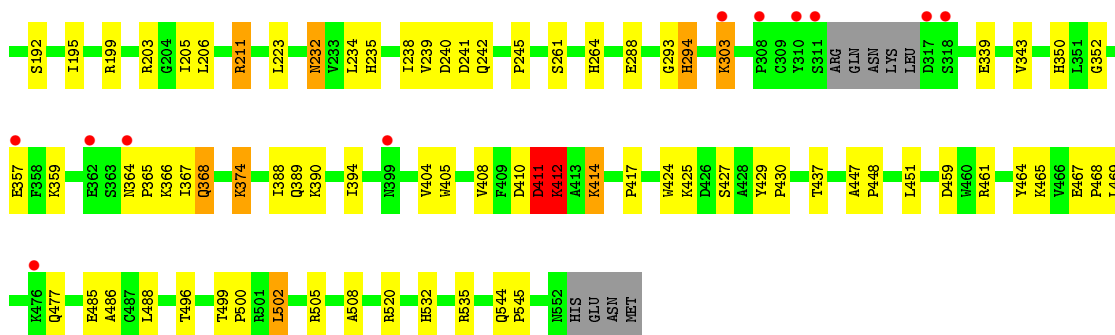


• Molecule 2: Beta-hexosaminidase beta chain



• Molecule 2: Beta-hexosaminidase beta chain





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

Chain I: 100%

HA61  
HA62  
BXA3

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

Chain J: 100%

HA61  
HA62  
BXA3

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

Chain K: 33% 67%

HA61  
HA62  
BXA3

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

Chain M: 33% 33% 33%

HA61  
HA62  
BXA3

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

Chain L: 100%

HA61  
HA62

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

Chain N:  50% 50%

 MAG1  
MAG2

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

Chain O:  50% 50%

 MAG1  
MAG2

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

Chain P:  50% 50%

 MAG1  
MAG2

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

Chain Q:  50% 50%


 MAG1  
MAG2

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

Chain R:  100%

 MAG1  
MAG2

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

Chain S:  50% 50%

 MAG1  
MAG2

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	321.09Å 110.54Å 129.67Å 90.00° 90.90° 90.00°	Depositor
Resolution (Å)	39.87 – 2.80 39.87 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.7 (39.87-2.80) 99.7 (39.87-2.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.92 (at 2.81Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.269 , 0.288 0.267 , 0.287	Depositor DCC
$R_{free}$ test set	5525 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	61.9	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 39.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.009 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	32132	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.62% 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: BMA, NAG, SO4

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	0.35	1/4132 (0.0%)	0.43	0/5633
1	D	0.43	4/4132 (0.1%)	0.48	2/5633 (0.0%)
1	E	0.31	0/4132	0.43	0/5633
1	H	0.38	2/4132 (0.0%)	0.47	3/5633 (0.1%)
2	B	0.27	0/3988	0.39	0/5423
2	C	0.27	0/3996	0.40	0/5434
2	F	0.28	0/3988	0.41	1/5423 (0.0%)
2	G	0.38	3/3988 (0.1%)	0.42	0/5423
All	All	0.34	10/32488 (0.0%)	0.43	6/44235 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	3
2	G	0	2
All	All	0	5

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	411	ASP	CA-C	12.40	1.85	1.52
1	D	210	SER	N-CA	11.89	1.70	1.46
1	A	51	SER	CB-OG	11.22	1.56	1.42
1	H	422	LEU	C-O	9.77	1.42	1.23
1	D	209	PRO	CA-C	8.49	1.69	1.52

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	51	SER	N-CA-C	-8.43	88.24	111.00
1	H	282	GLU	C-N-CD	-8.15	102.66	120.60
1	H	422	LEU	O-C-N	-6.93	111.62	122.70
2	F	503	TRP	N-CA-C	5.53	125.94	111.00
1	D	51	SER	N-CA-CB	5.50	118.75	110.50

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	209	PRO	Mainchain,Peptide
1	D	50	VAL	Peptide
2	G	411	ASP	Mainchain,Peptide

## 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	4010	0	3840	73	0
1	D	4010	0	3840	106	0
1	E	4010	0	3841	136	0
1	H	4010	0	3841	115	0
2	B	3877	0	3783	55	0
2	C	3885	0	3794	49	0
2	F	3877	0	3783	83	0
2	G	3877	0	3782	74	0
3	I	39	0	34	0	0
3	J	39	0	34	0	0
3	K	39	0	34	0	0
3	M	39	0	34	1	0
4	L	28	0	25	0	0
4	N	28	0	25	1	0
4	O	28	0	25	0	0
4	P	28	0	25	1	0
4	Q	28	0	25	0	0
4	R	28	0	25	0	0
4	S	28	0	25	2	0
5	A	14	0	13	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	14	0	13	0	0
5	E	14	0	13	0	0
5	G	14	0	13	0	0
6	A	5	0	0	0	0
6	D	5	0	0	0	0
6	H	10	0	0	1	0
7	A	34	0	0	0	0
7	B	20	0	0	2	0
7	C	22	0	0	1	0
7	D	18	0	0	0	0
7	E	20	0	0	2	0
7	F	11	0	0	0	0
7	G	17	0	0	0	0
7	H	6	0	0	0	0
All	All	32132	0	30867	663	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 11.

The worst 5 of 663 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:D:210:SER:N	1:D:210:SER:CA	1.70	1.53
2:G:411:ASP:CA	2:G:411:ASP:C	1.85	1.46
1:E:55:GLN:HG3	1:E:56:PRO:HD2	1.29	1.13
1:H:424:ARG:HH11	1:H:424:ARG:HG2	1.08	1.12
1:H:53:ALA:HB1	1:H:101:THR:HA	1.33	1.10

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	488/507 (96%)	453 (93%)	33 (7%)	2 (0%)	34	66
1	D	488/507 (96%)	449 (92%)	36 (7%)	3 (1%)	25	56
1	E	488/507 (96%)	435 (89%)	39 (8%)	14 (3%)	4	15
1	H	488/507 (96%)	447 (92%)	31 (6%)	10 (2%)	7	24
2	B	474/507 (94%)	447 (94%)	26 (6%)	1 (0%)	47	78
2	C	475/507 (94%)	450 (95%)	25 (5%)	0	100	100
2	F	474/507 (94%)	434 (92%)	36 (8%)	4 (1%)	19	49
2	G	474/507 (94%)	442 (93%)	26 (6%)	6 (1%)	12	36
All	All	3849/4056 (95%)	3557 (92%)	252 (6%)	40 (1%)	15	44

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	PRO
1	A	103	GLY
1	D	50	VAL
1	D	103	GLY
1	E	382	LYS

### 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	441/454 (97%)	424 (96%)	17 (4%)	32	66
1	D	441/454 (97%)	428 (97%)	13 (3%)	42	76
1	E	441/454 (97%)	417 (95%)	24 (5%)	22	53
1	H	441/454 (97%)	422 (96%)	19 (4%)	29	62
2	B	426/449 (95%)	420 (99%)	6 (1%)	67	90
2	C	427/449 (95%)	422 (99%)	5 (1%)	71	92
2	F	426/449 (95%)	417 (98%)	9 (2%)	53	84
2	G	426/449 (95%)	416 (98%)	10 (2%)	50	82

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	3469/3612 (96%)	3366 (97%)	103 (3%)	41 75

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	122	ASP
1	E	370	TYR
1	H	299	GLU
1	E	127	LEU
1	E	221	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 108 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	296	ASN
1	E	318	HIS
1	H	228	ASN
1	D	487	ASN
1	E	55	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](#)

26 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	1,3	14,14,15	0.63	0	17,19,21	0.70	0
3	NAG	I	2	3	14,14,15	0.54	0	17,19,21	0.70	0
3	BMA	I	3	3	11,11,12	0.57	0	15,15,17	0.68	0
3	NAG	J	1	1,3	14,14,15	0.60	0	17,19,21	0.94	1 (5%)
3	NAG	J	2	3	14,14,15	0.68	0	17,19,21	0.87	1 (5%)
3	BMA	J	3	3	11,11,12	0.52	0	15,15,17	1.36	2 (13%)
3	NAG	K	1	3,2	14,14,15	0.74	0	17,19,21	1.01	1 (5%)
3	NAG	K	2	3	14,14,15	0.62	0	17,19,21	0.74	0
3	BMA	K	3	3	11,11,12	0.60	0	15,15,17	0.96	1 (6%)
4	NAG	L	1	2,4	14,14,15	0.47	0	17,19,21	0.72	0
4	NAG	L	2	4	14,14,15	0.48	0	17,19,21	0.74	0
3	NAG	M	1	1,3	14,14,15	0.83	0	17,19,21	1.09	2 (11%)
3	NAG	M	2	3	14,14,15	0.61	0	17,19,21	0.78	0
3	BMA	M	3	3	11,11,12	0.59	0	15,15,17	0.70	0
4	NAG	N	1	1,4	14,14,15	0.70	0	17,19,21	1.08	1 (5%)
4	NAG	N	2	4	14,14,15	0.50	0	17,19,21	0.75	0
4	NAG	O	1	1,4	14,14,15	0.54	0	17,19,21	0.84	1 (5%)
4	NAG	O	2	4	14,14,15	0.57	0	17,19,21	0.58	0
4	NAG	P	1	1,4	14,14,15	0.64	0	17,19,21	0.77	0
4	NAG	P	2	4	14,14,15	0.53	0	17,19,21	0.67	0
4	NAG	Q	1	2,4	14,14,15	0.55	0	17,19,21	0.84	1 (5%)
4	NAG	Q	2	4	14,14,15	0.50	0	17,19,21	0.61	0
4	NAG	R	1	1,4	14,14,15	0.56	0	17,19,21	0.68	0
4	NAG	R	2	4	14,14,15	0.49	0	17,19,21	0.68	0
4	NAG	S	1	1,4	14,14,15	0.65	0	17,19,21	1.00	2 (11%)
4	NAG	S	2	4	14,14,15	0.49	0	17,19,21	0.64	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	1,3	-	4/6/23/26	0/1/1/1
3	NAG	I	2	3	-	5/6/23/26	0/1/1/1
3	BMA	I	3	3	-	0/2/19/22	0/1/1/1
3	NAG	J	1	1,3	-	4/6/23/26	0/1/1/1
3	NAG	J	2	3	-	3/6/23/26	0/1/1/1
3	BMA	J	3	3	-	0/2/19/22	0/1/1/1

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	NAG	K	1	3,2	1/1/5/7	4/6/23/26	0/1/1/1
3	NAG	K	2	3	-	3/6/23/26	0/1/1/1
3	BMA	K	3	3	-	0/2/19/22	0/1/1/1
4	NAG	L	1	2,4	-	2/6/23/26	0/1/1/1
4	NAG	L	2	4	-	3/6/23/26	0/1/1/1
3	NAG	M	1	1,3	-	2/6/23/26	0/1/1/1
3	NAG	M	2	3	-	3/6/23/26	0/1/1/1
3	BMA	M	3	3	-	1/2/19/22	0/1/1/1
4	NAG	N	1	1,4	-	3/6/23/26	0/1/1/1
4	NAG	N	2	4	-	4/6/23/26	0/1/1/1
4	NAG	O	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	O	2	4	-	2/6/23/26	0/1/1/1
4	NAG	P	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	P	2	4	-	3/6/23/26	0/1/1/1
4	NAG	Q	1	2,4	-	4/6/23/26	0/1/1/1
4	NAG	Q	2	4	-	3/6/23/26	0/1/1/1
4	NAG	R	1	1,4	-	2/6/23/26	0/1/1/1
4	NAG	R	2	4	-	3/6/23/26	0/1/1/1
4	NAG	S	1	1,4	-	0/6/23/26	0/1/1/1
4	NAG	S	2	4	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	3	BMA	C1-C2-C3	3.77	114.31	109.67
3	K	1	NAG	C4-C3-C2	3.22	115.74	111.02
4	N	1	NAG	C4-C3-C2	2.75	115.05	111.02
3	K	3	BMA	C1-C2-C3	2.74	113.03	109.67
4	Q	1	NAG	O5-C1-C2	-2.39	107.52	111.29

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	K	1	NAG	C1

5 of 66 torsion outliers are listed below:

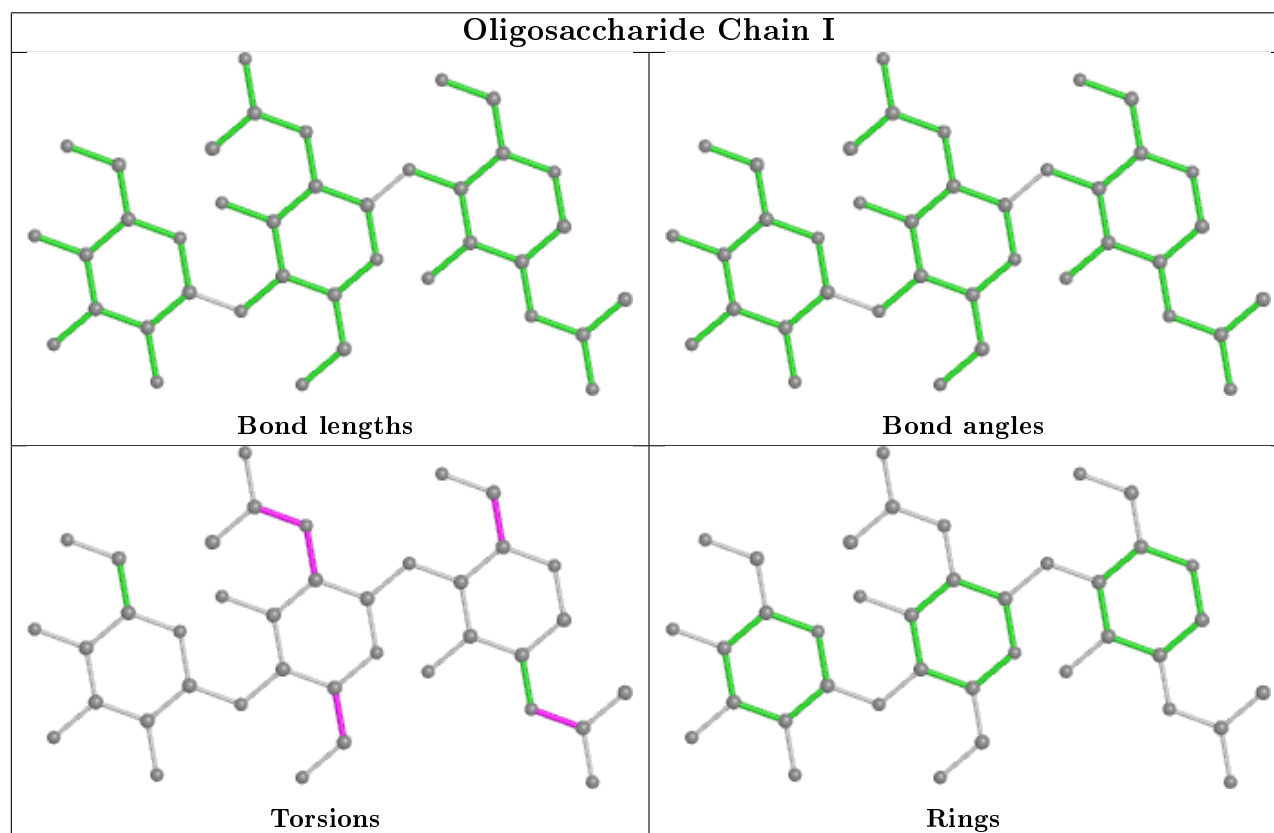
Mol	Chain	Res	Type	Atoms
4	P	2	NAG	C3-C2-N2-C7
3	J	2	NAG	C3-C2-N2-C7
4	R	2	NAG	C8-C7-N2-C2
4	R	2	NAG	O7-C7-N2-C2
4	Q	2	NAG	C3-C2-N2-C7

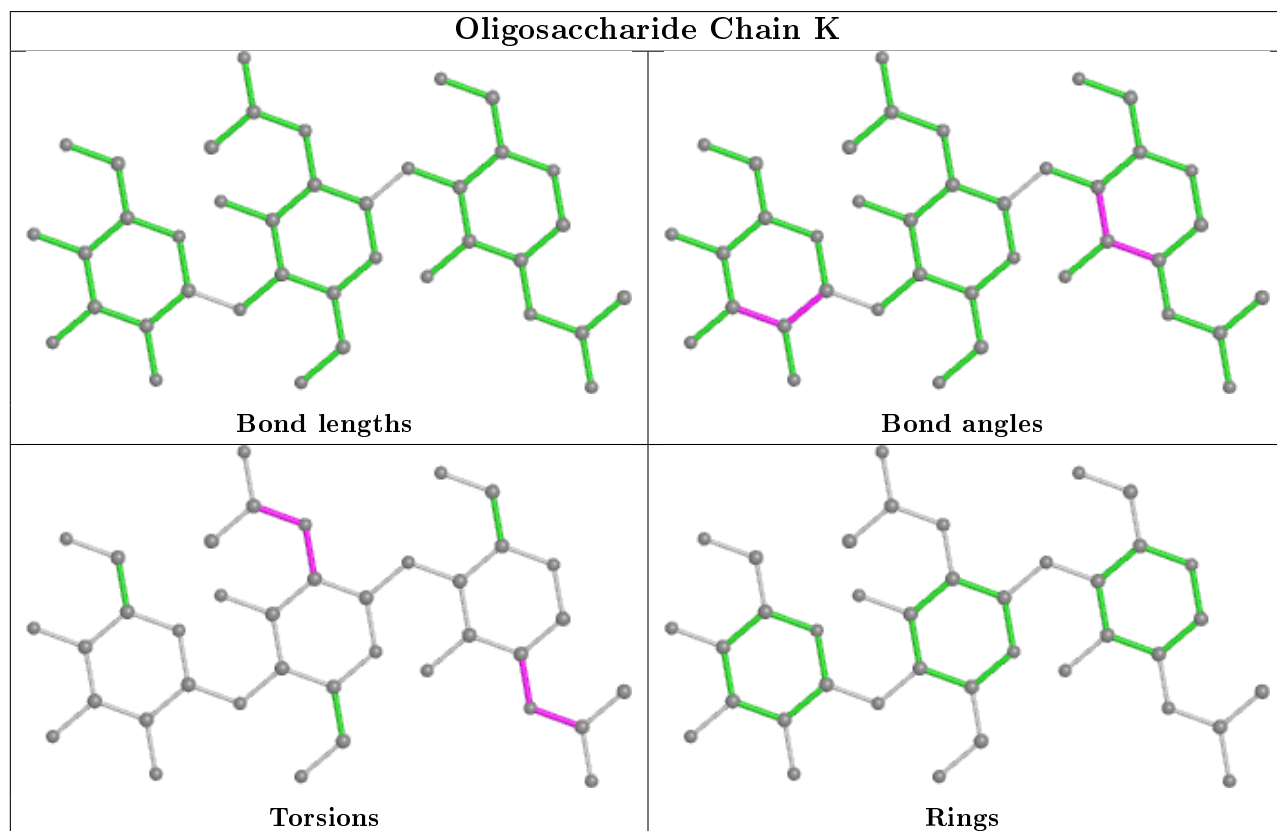
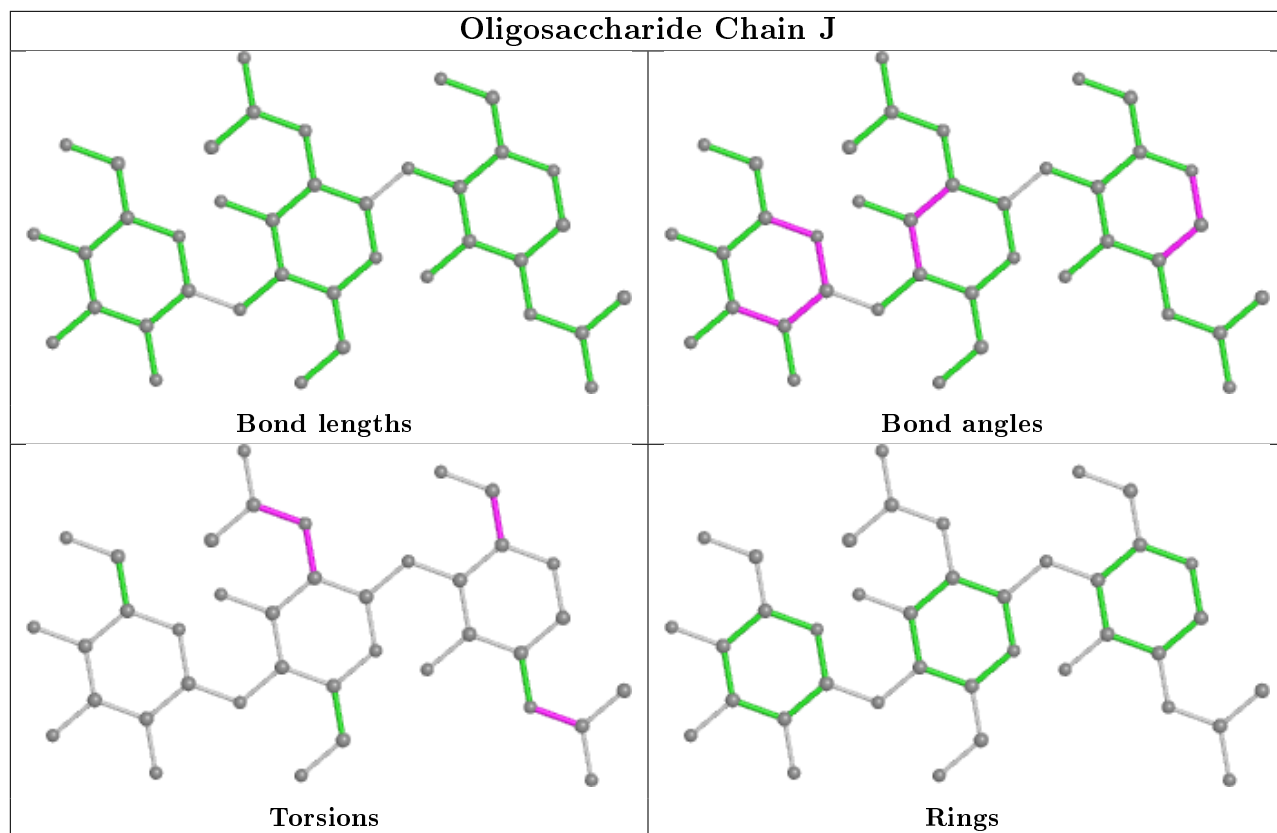
There are no ring outliers.

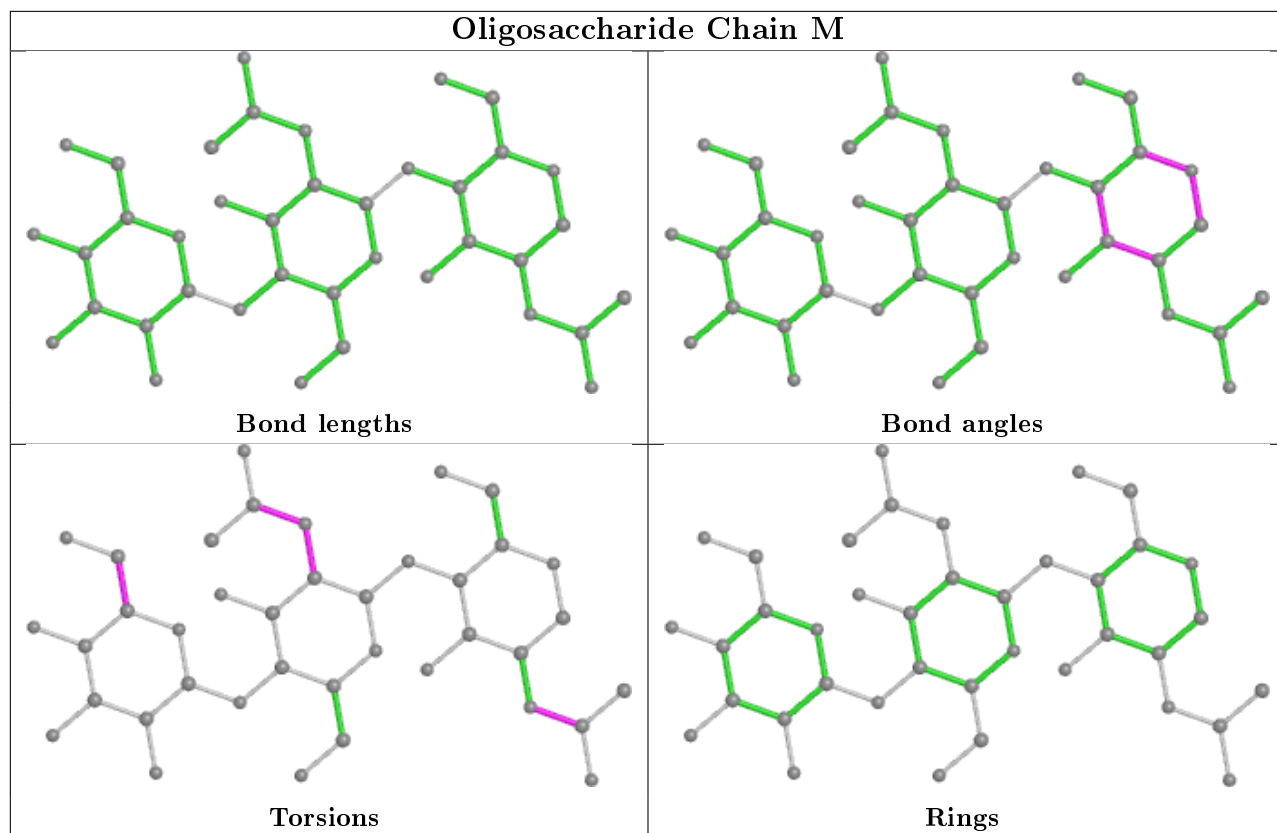
6 monomers are involved in 5 short contacts:

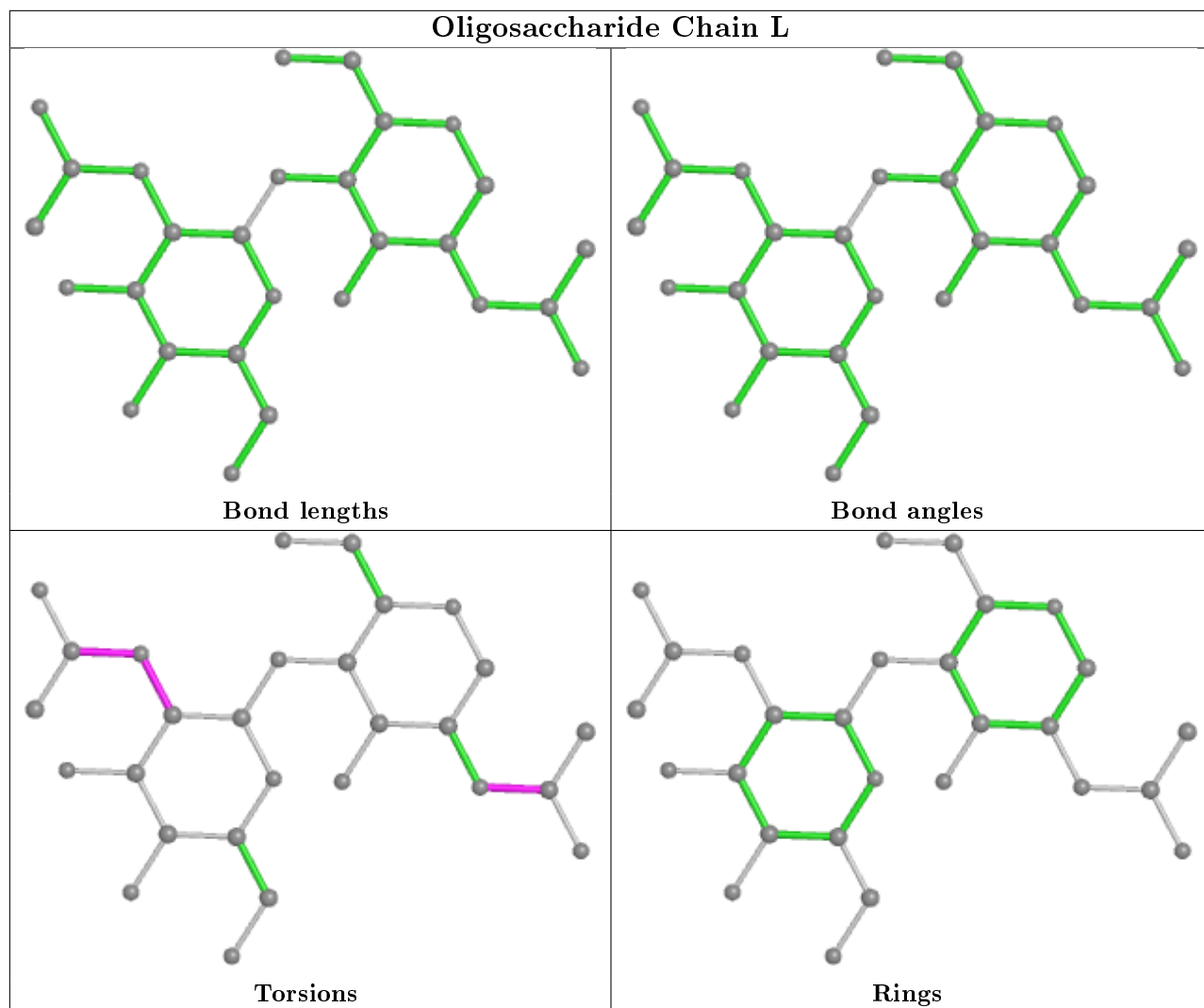
Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	N	1	NAG	1	0
4	P	1	NAG	1	0
3	M	1	NAG	1	0
4	S	2	NAG	2	0
4	S	1	NAG	2	0
3	M	2	NAG	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for oligosaccharide.

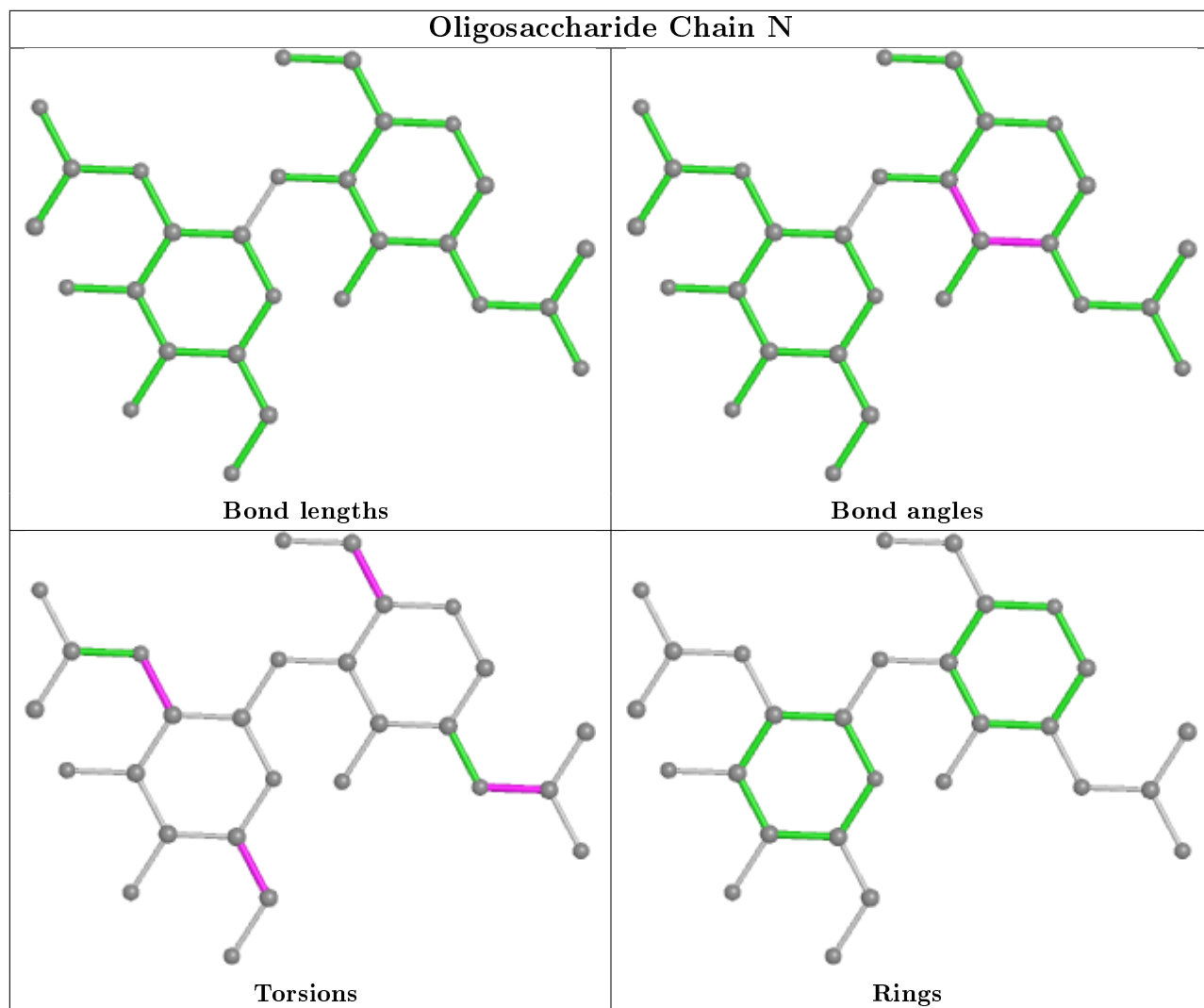


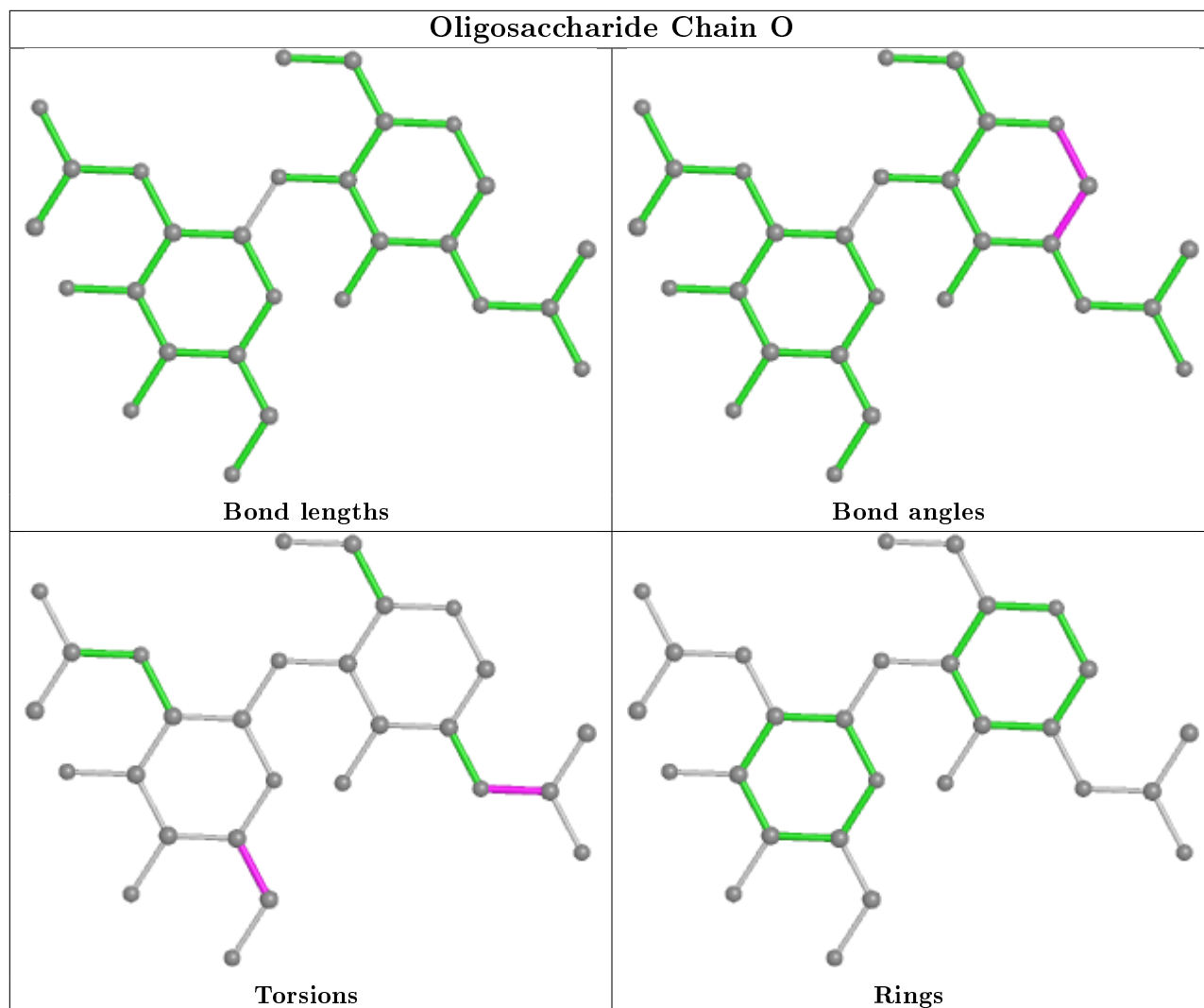


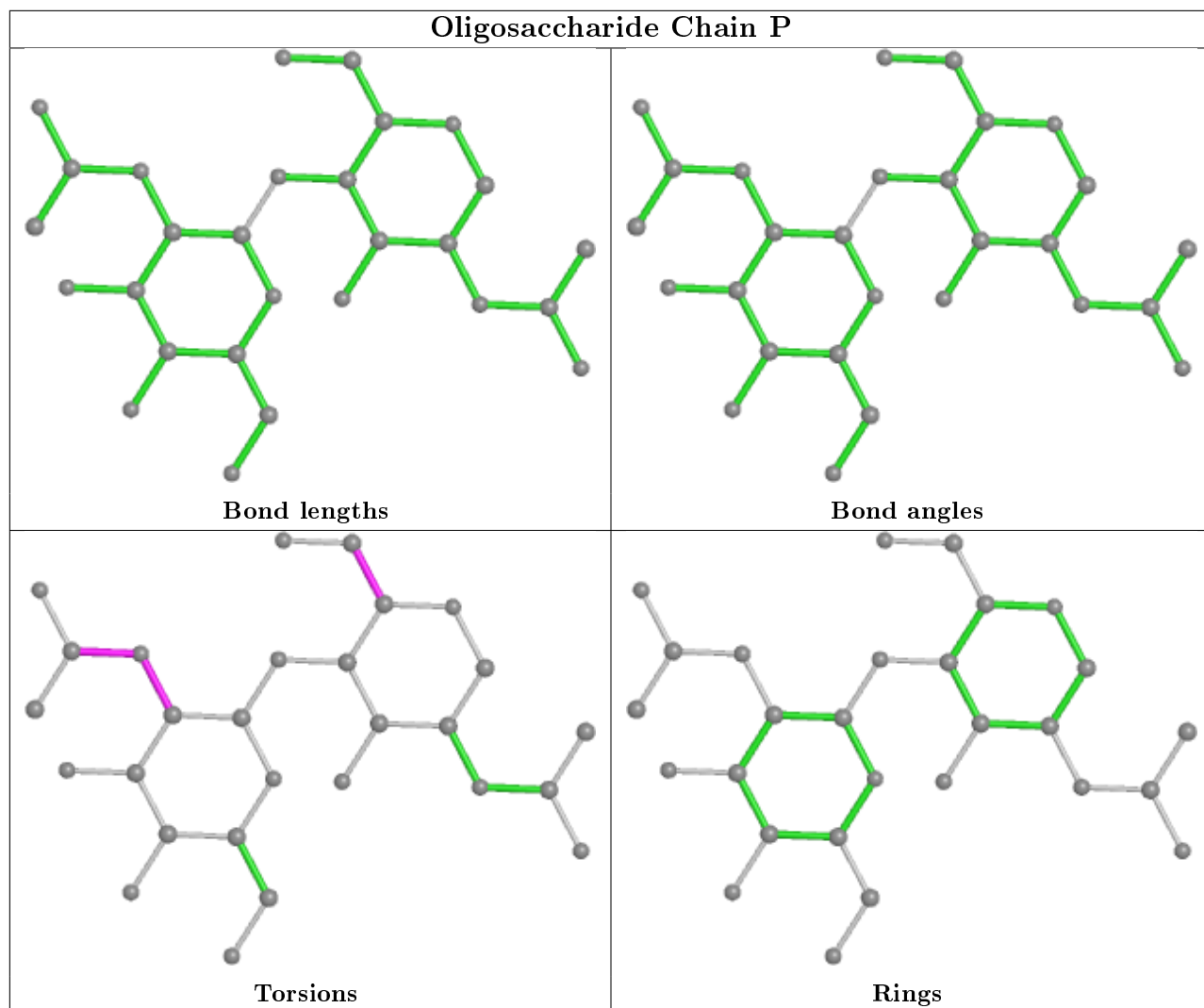


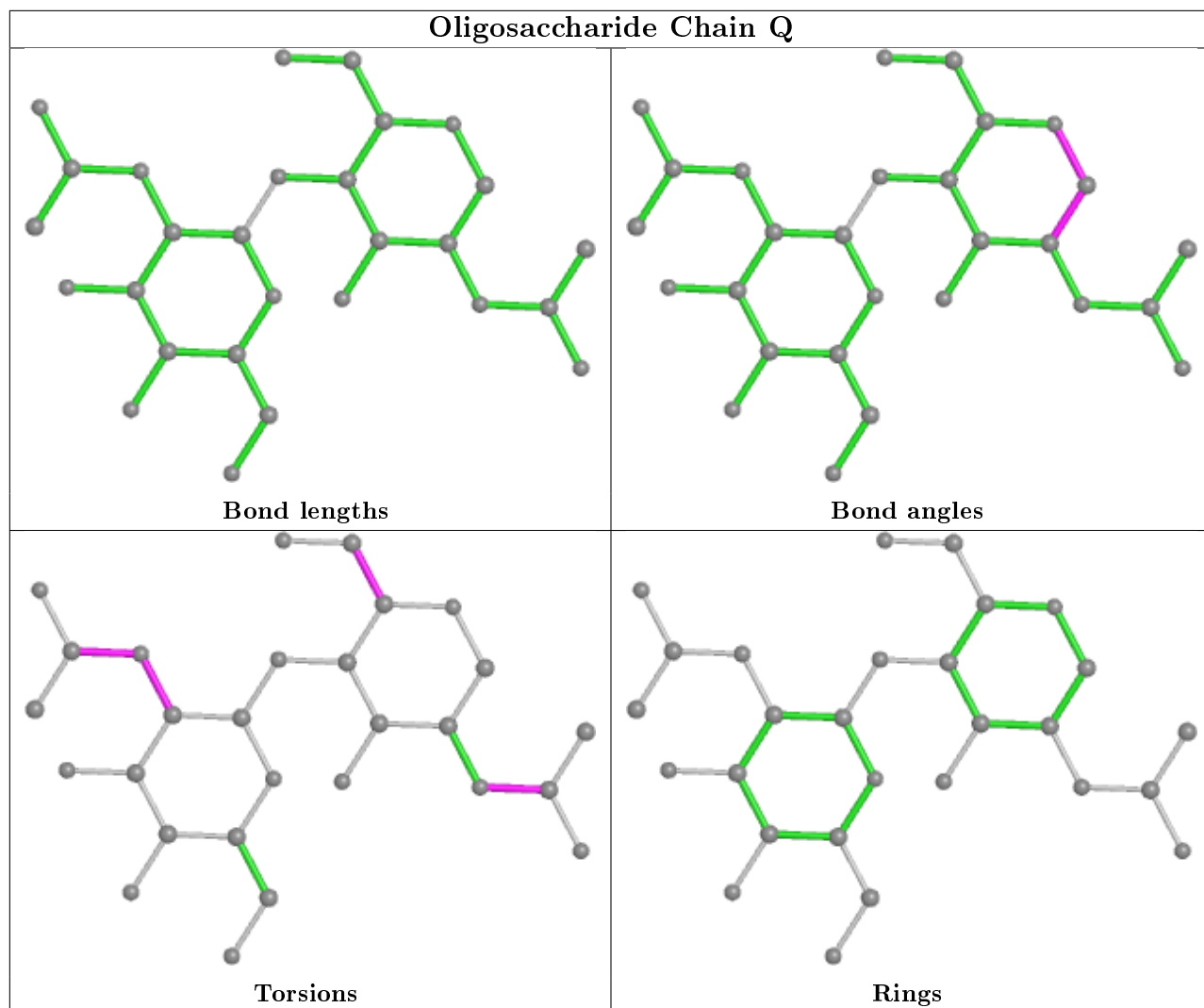


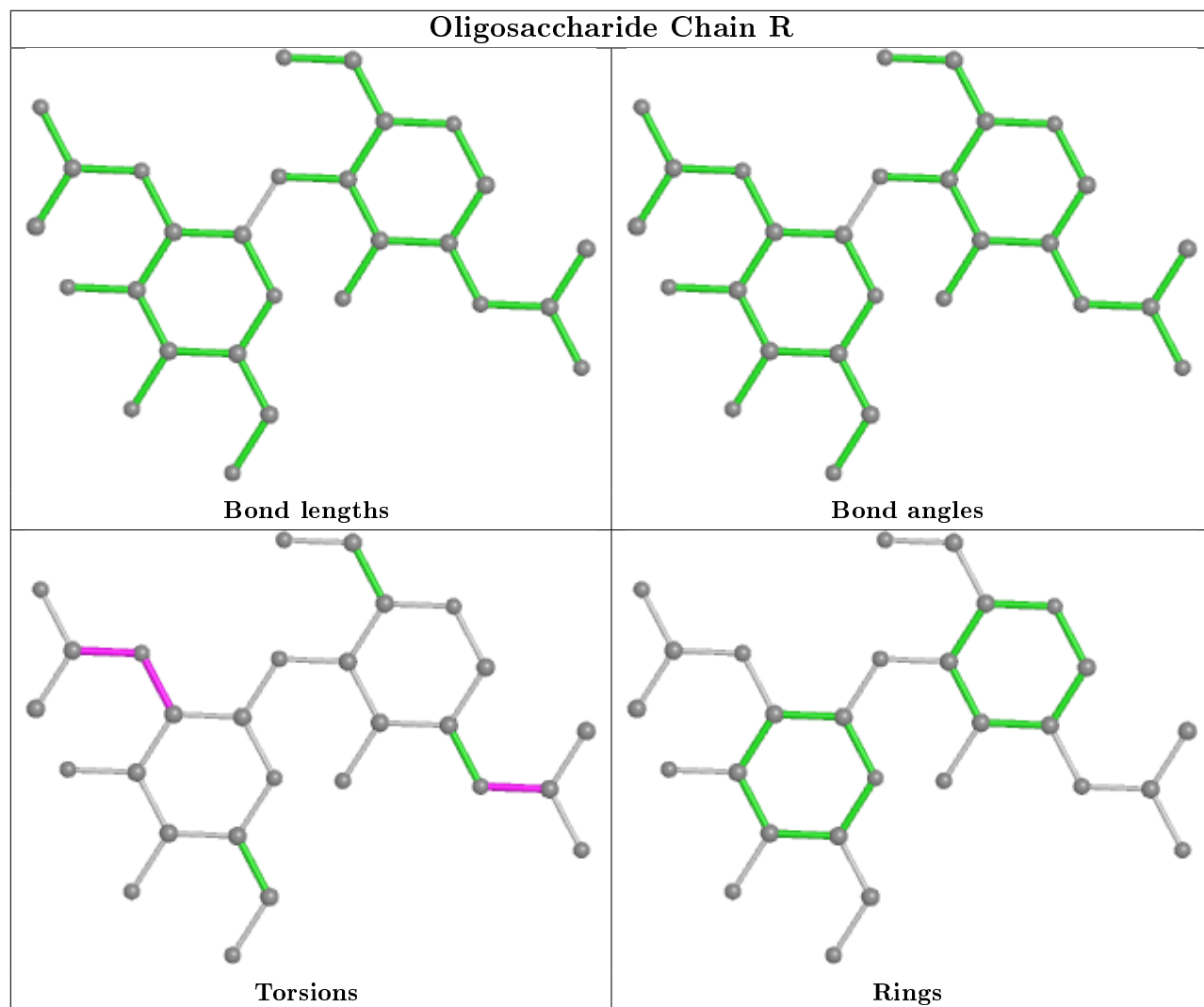


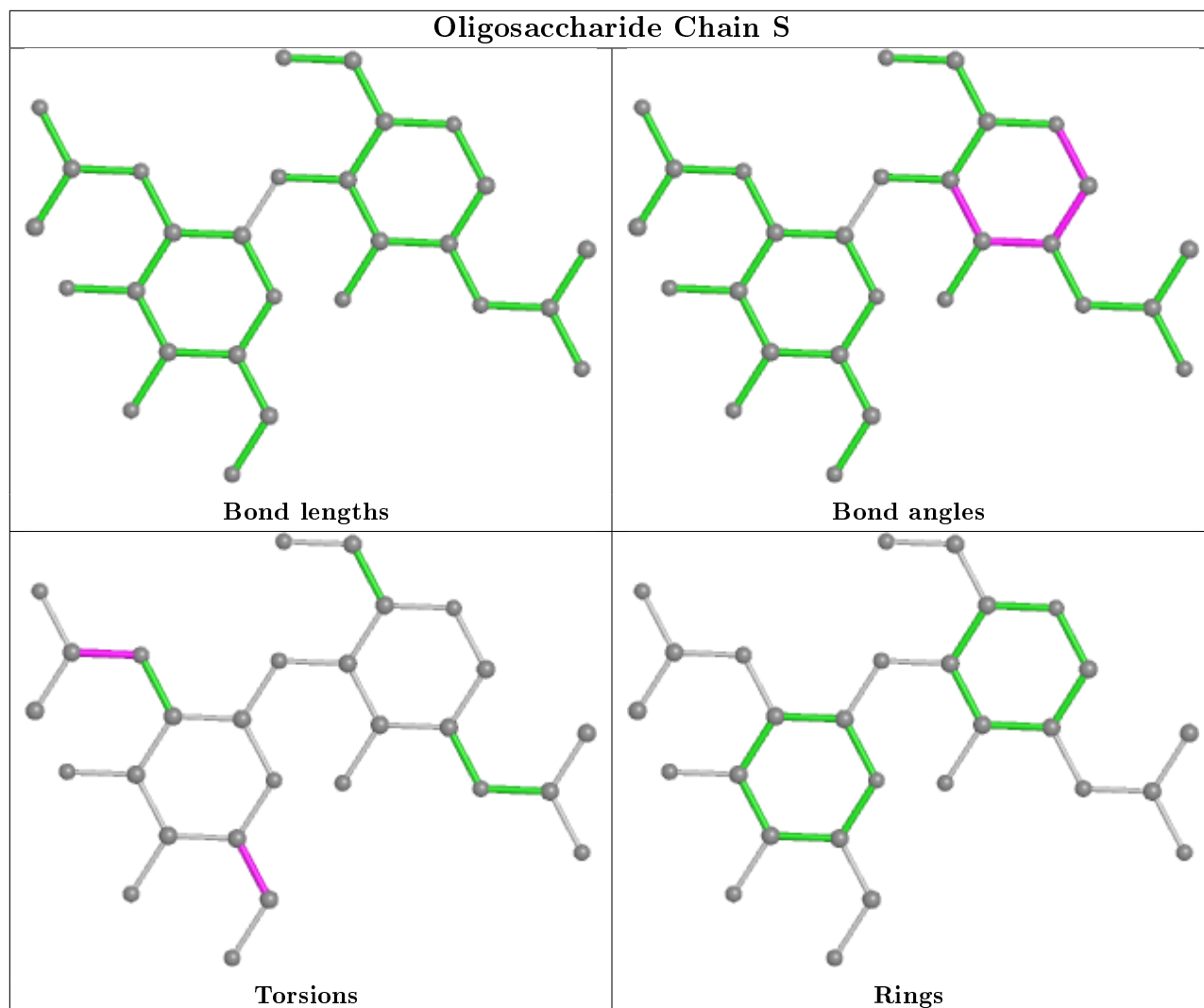












## 5.6 Ligand geometry [i](#)

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
6	SO4	A	708	-	4,4,4	0.16	0	6,6,6	0.05	0
6	SO4	H	534	-	4,4,4	0.14	0	6,6,6	0.05	0
5	NAG	A	707	1	14,14,15	0.50	0	17,19,21	0.69	0
6	SO4	H	535	-	4,4,4	0.13	0	6,6,6	0.06	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	G	726	2	14,14,15	0.52	0	17,19,21	0.64	0
5	NAG	D	716	1	14,14,15	0.55	0	17,19,21	0.77	0
6	SO4	D	717	-	4,4,4	0.15	0	6,6,6	0.05	0
5	NAG	E	723	2	14,14,15	0.48	0	17,19,21	0.75	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	D	716	1	-	3/6/23/26	0/1/1/1
5	NAG	A	707	1	-	4/6/23/26	0/1/1/1
5	NAG	G	726	2	-	4/6/23/26	0/1/1/1
5	NAG	E	723	2	-	1/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	707	NAG	C8-C7-N2-C2
5	A	707	NAG	O7-C7-N2-C2
5	D	716	NAG	C8-C7-N2-C2
5	D	716	NAG	O7-C7-N2-C2
5	G	726	NAG	C8-C7-N2-C2

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	H	534	SO4	1	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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	492/507 (97%)	0.30	11 (2%) 62 52	17, 59, 80, 87	0
1	D	492/507 (97%)	0.28	18 (3%) 41 31	21, 61, 81, 89	0
1	E	492/507 (97%)	0.47	36 (7%) 15 8	17, 61, 82, 90	0
1	H	492/507 (97%)	0.35	24 (4%) 29 20	17, 61, 81, 88	0
2	B	480/507 (94%)	0.33	18 (3%) 40 30	26, 63, 76, 84	0
2	C	481/507 (94%)	0.27	12 (2%) 57 47	26, 61, 75, 85	0
2	F	480/507 (94%)	0.40	19 (3%) 38 28	27, 64, 76, 83	0
2	G	480/507 (94%)	0.30	17 (3%) 44 34	17, 62, 77, 82	0
All	All	3889/4056 (95%)	0.34	155 (3%) 38 28	17, 62, 78, 90	0

The worst 5 of 155 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	317	ASP	6.4
2	C	316	LEU	6.3
2	C	311	SER	5.7
1	E	279	SER	5.6
2	G	317	ASP	5.5

### 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,

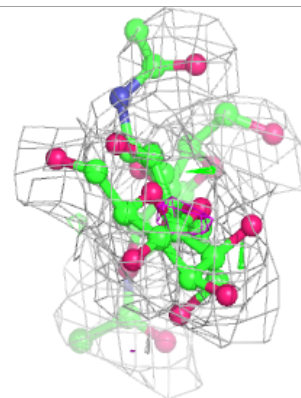
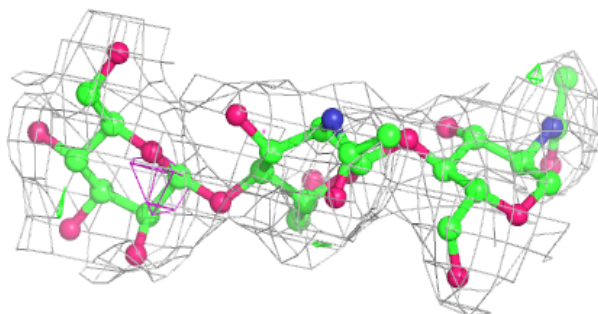
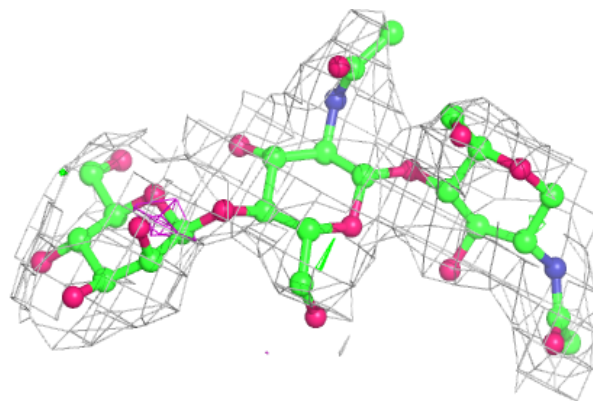
median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
4	NAG	Q	2	14/15	0.59	0.33	110,117,118,119	0
3	BMA	J	3	11/12	0.60	0.35	55,58,64,65	0
3	NAG	J	2	14/15	0.66	0.28	110,117,118,119	0
4	NAG	L	2	14/15	0.71	0.48	110,117,118,119	0
4	NAG	R	2	14/15	0.71	0.26	110,117,118,119	0
3	NAG	I	2	14/15	0.72	0.21	110,117,118,119	0
3	BMA	M	3	11/12	0.72	0.36	55,58,64,65	0
3	BMA	K	3	11/12	0.72	0.38	55,58,64,65	0
4	NAG	P	2	14/15	0.73	0.29	110,117,118,119	0
3	NAG	K	2	14/15	0.73	0.31	110,117,118,119	0
4	NAG	N	2	14/15	0.75	0.37	110,117,118,119	0
4	NAG	O	2	14/15	0.76	0.24	110,117,118,119	0
3	NAG	M	2	14/15	0.80	0.27	110,117,118,119	0
3	BMA	I	3	11/12	0.80	0.32	55,58,64,65	0
4	NAG	S	2	14/15	0.82	0.21	91,94,95,96	0
4	NAG	R	1	14/15	0.84	0.25	70,74,80,89	0
4	NAG	O	1	14/15	0.84	0.23	70,74,80,89	0
3	NAG	K	1	14/15	0.86	0.23	70,74,80,89	0
4	NAG	N	1	14/15	0.86	0.18	70,74,80,89	0
3	NAG	M	1	14/15	0.86	0.22	70,74,80,89	0
3	NAG	I	1	14/15	0.87	0.19	70,74,80,89	0
4	NAG	P	1	14/15	0.88	0.25	70,74,80,89	0
3	NAG	J	1	14/15	0.88	0.23	70,74,80,89	0
4	NAG	L	1	14/15	0.90	0.28	70,74,80,89	0
4	NAG	Q	1	14/15	0.90	0.23	70,74,80,89	0
4	NAG	S	1	14/15	0.93	0.13	66,71,74,79	0

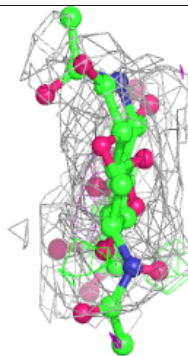
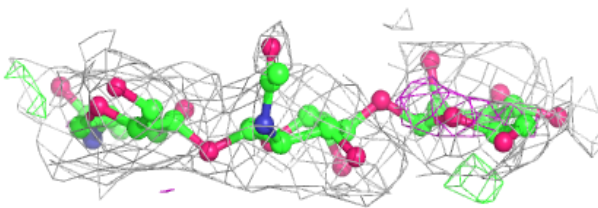
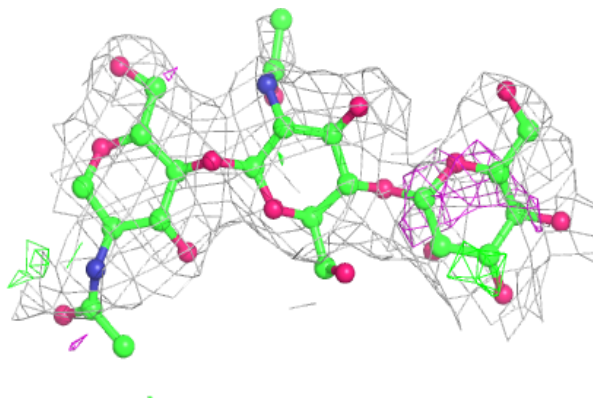
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

**Electron density around Chain I:**

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

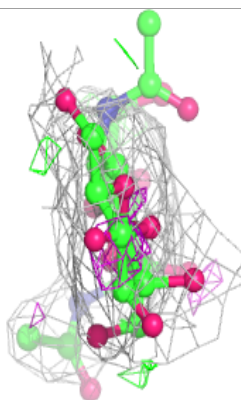
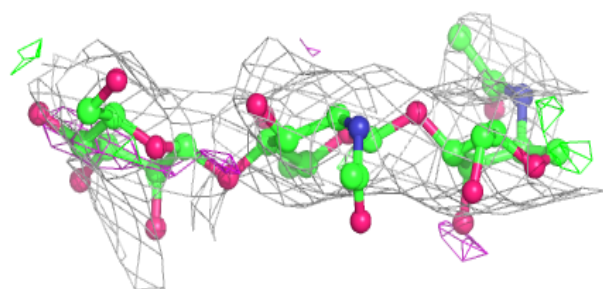
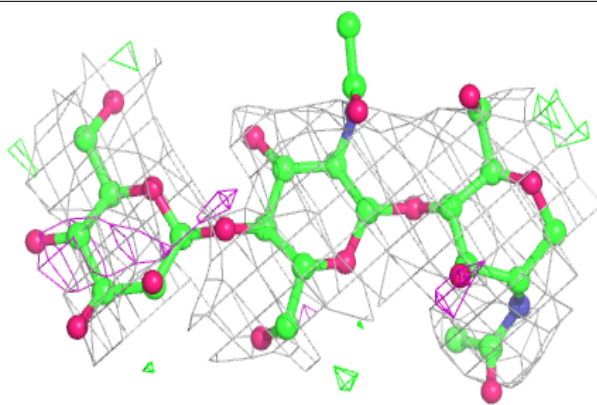
**Electron density around Chain J:**

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

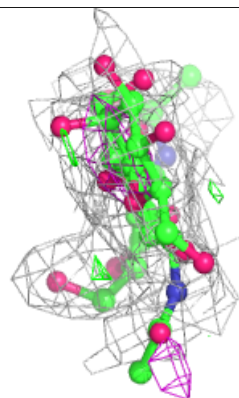
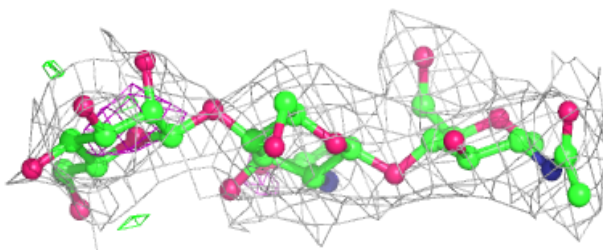
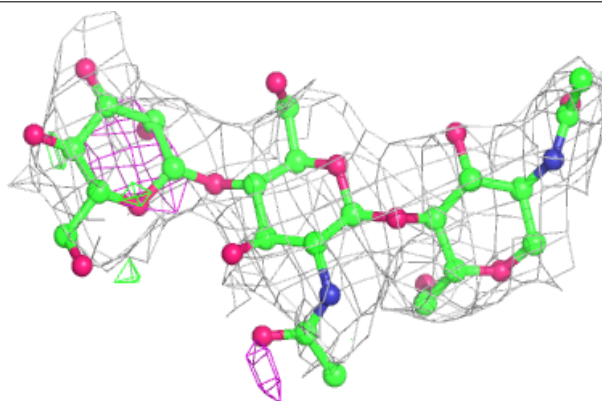


**Electron density around Chain K:**

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

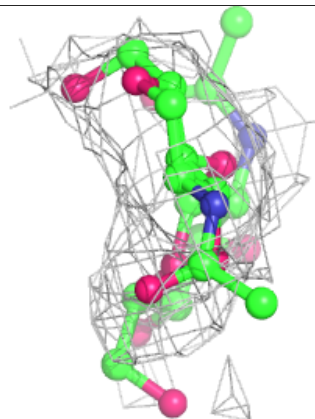
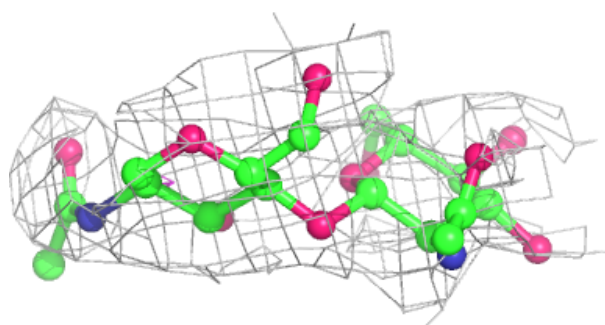
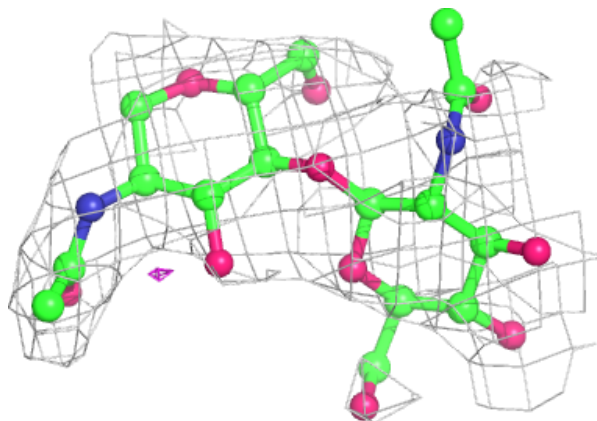
**Electron density around Chain M:**

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

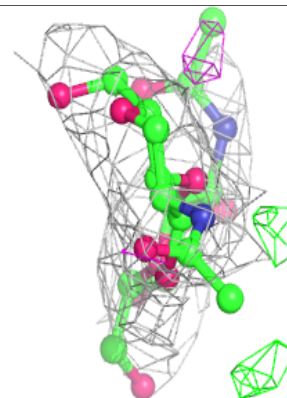
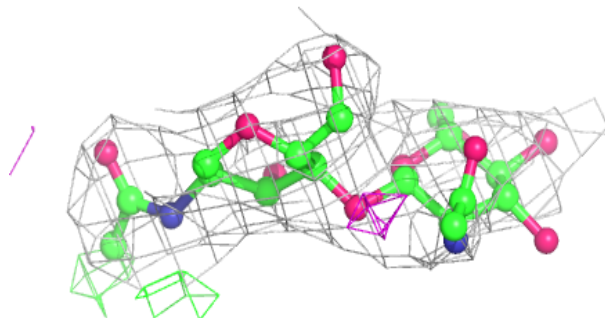
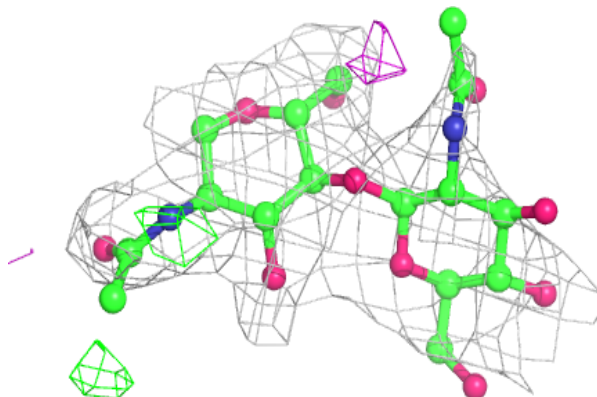


**Electron density around Chain L:**

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

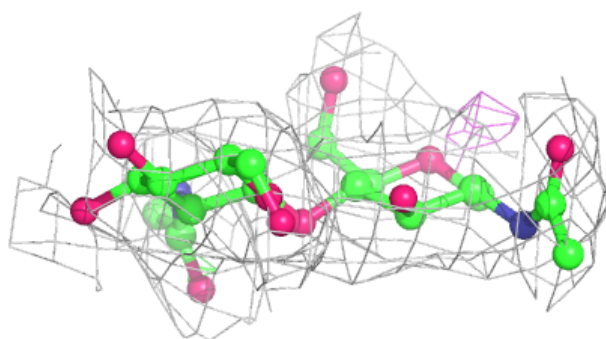
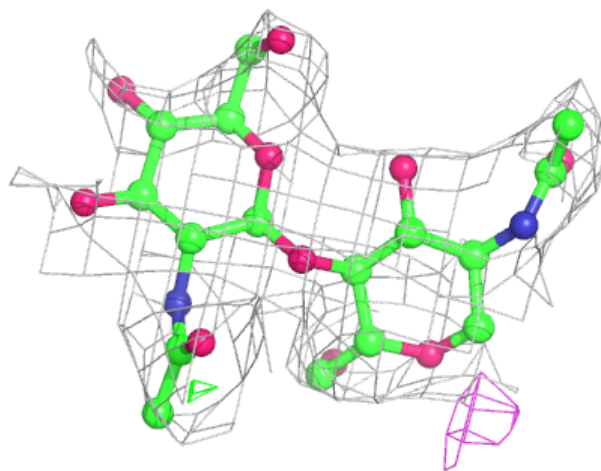
**Electron density around Chain N:**

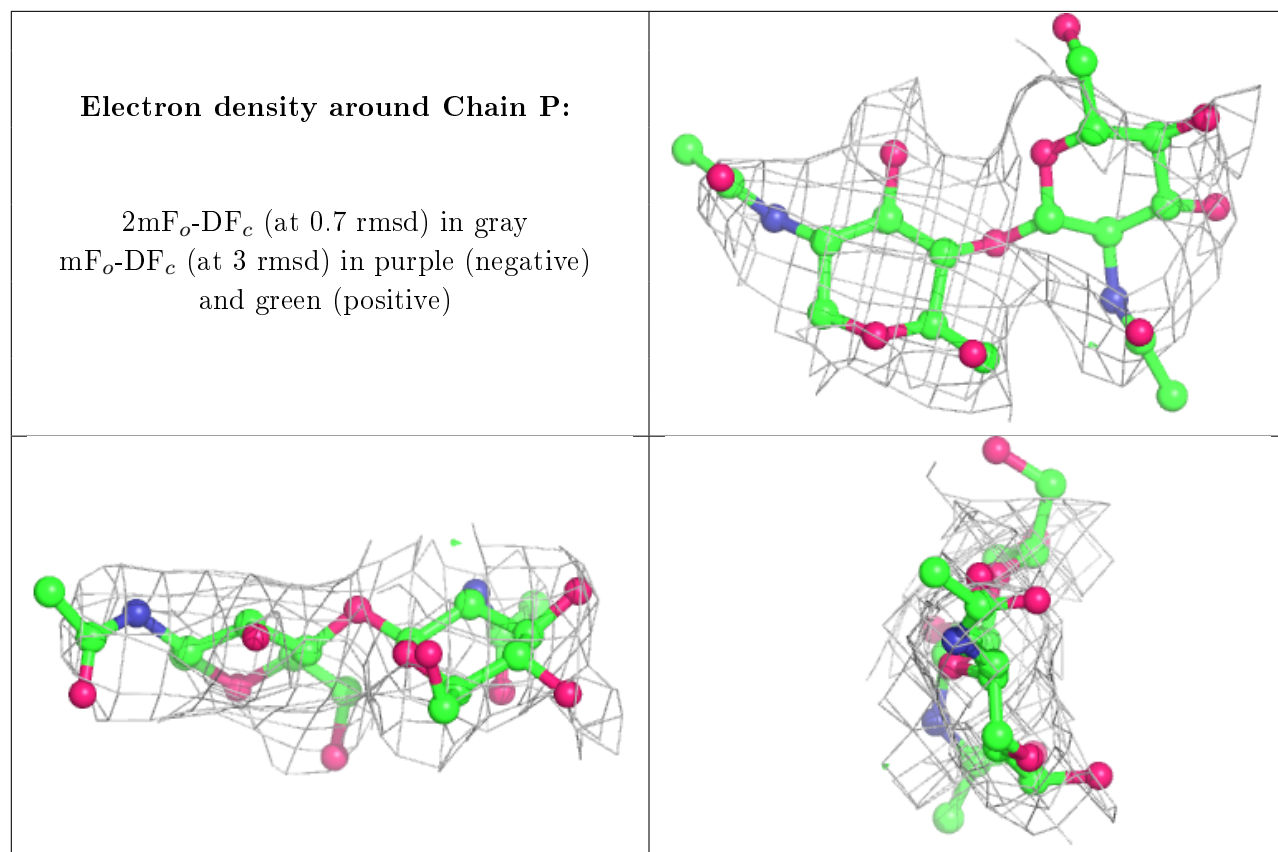
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around Chain O:**

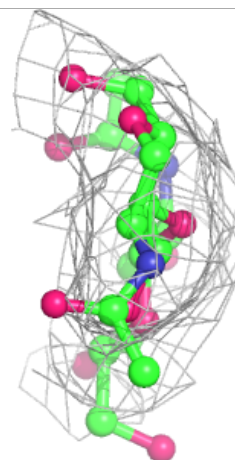
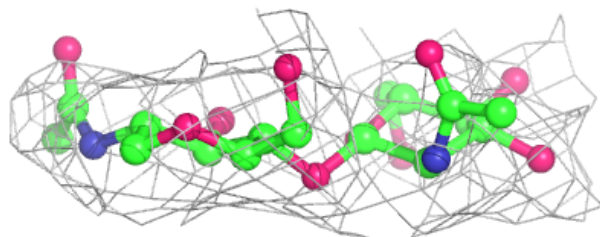
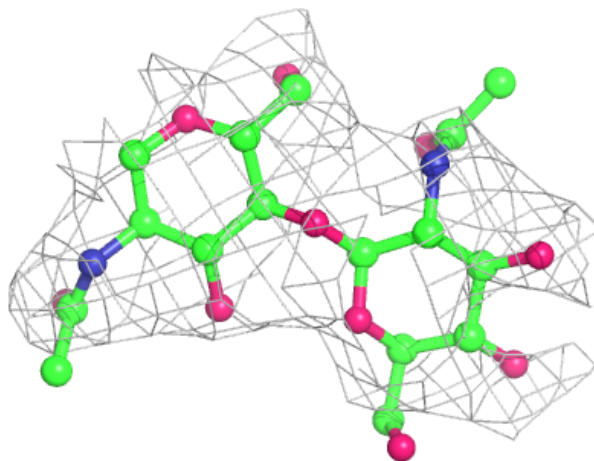
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





**Electron density around Chain Q:**

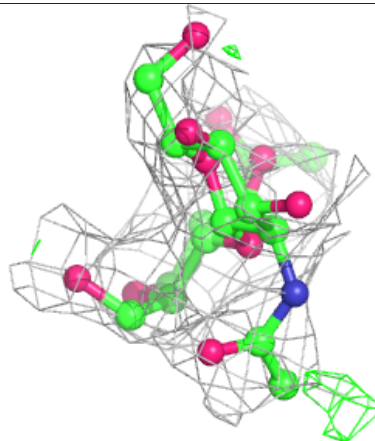
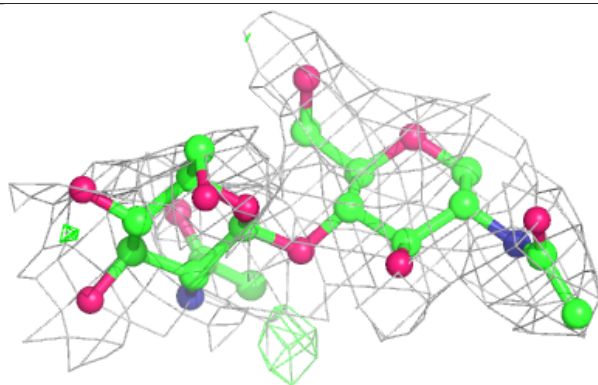
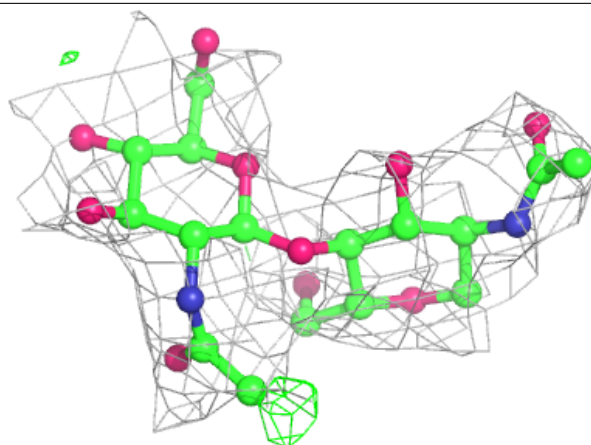
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



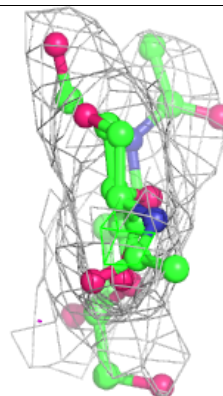
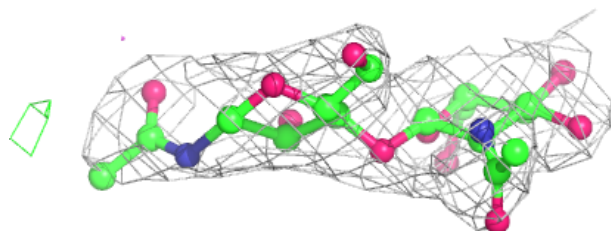
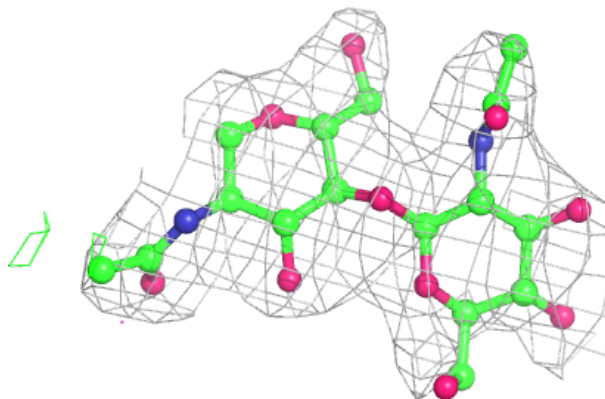


**Electron density around Chain R:**

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

**Electron density around Chain S:**

$2mF_o-DF_c$  (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	NAG	D	716	14/15	0.61	0.42	120,122,122,122	0
5	NAG	A	707	14/15	0.72	0.41	90,94,95,95	0
5	NAG	G	726	14/15	0.81	0.26	70,74,80,89	0
6	SO4	H	535	5/5	0.87	0.25	104,104,105,106	0
5	NAG	E	723	14/15	0.87	0.21	70,74,80,89	0
6	SO4	D	717	5/5	0.90	0.31	104,104,105,106	0
6	SO4	A	708	5/5	0.93	0.34	104,104,105,106	0
6	SO4	H	534	5/5	0.93	0.25	104,104,105,106	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.