Full wwPDB X-ray Structure Validation Report ⓘ

Jun 25, 2024 – 02:19 AM EDT

PDB ID	:	6PFO
Title	:	Crystal structure of N-glycosylated human calcitonin receptor extracellular domain in complex with salmon calcitonin (16-32)
Authors	:	Lee, S.; Pioszak, A.A.
Deposited on	:	2019-06-21
Resolution	:	1.78 Å(reported)

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

We welcome your comments at *validation@mail.wwpdb.org*
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:

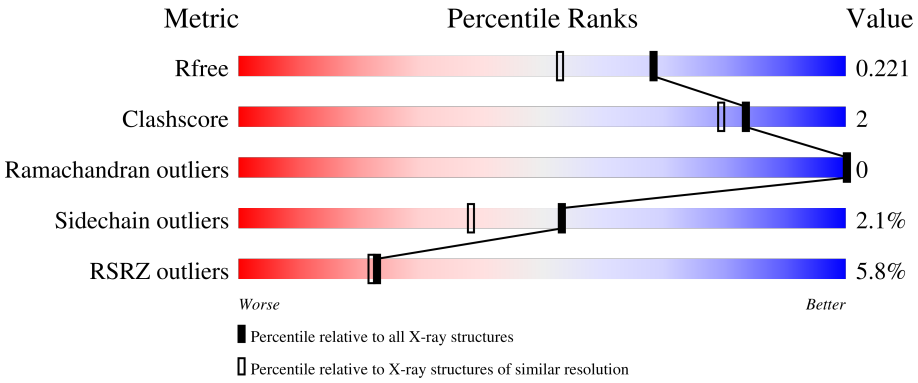
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.78 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	9185 (1.80-1.76)
Clashscore	141614	10184 (1.80-1.76)
Ramachandran outliers	138981	10051 (1.80-1.76)
Sidechain outliers	138945	10050 (1.80-1.76)
RSRZ outliers	127900	9032 (1.80-1.76)

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 >=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 <=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	484	6% 93% 5%
1	B	484	4% 90% 6%
2	C	18	17% 94% 6%
2	D	18	11% 83% 17%
3	E	2	50% 50%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	F	2	100%

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
4	NAG	A	501	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8740 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 Maltodextrin-binding protein,Calcitonin receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	474	Total	C	N	O	S	0	6	0
			3778	2435	615	710	18			
1	B	474	Total	C	N	O	S	0	26	0
			3923	2523	636	743	21			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-336	MET	-	initiating methionine	UNP A0A0A8UN35
A	32	ASN	-	linker	UNP A0A0A8UN35
A	33	ALA	-	linker	UNP A0A0A8UN35
A	34	ALA	-	linker	UNP A0A0A8UN35
A	35	ALA	-	linker	UNP A0A0A8UN35
A	36	GLU	-	linker	UNP A0A0A8UN35
A	37	PHE	-	linker	UNP A0A0A8UN35
A	142	HIS	-	expression tag	UNP P30988
A	143	HIS	-	expression tag	UNP P30988
A	144	HIS	-	expression tag	UNP P30988
A	145	HIS	-	expression tag	UNP P30988
A	146	HIS	-	expression tag	UNP P30988
A	147	HIS	-	expression tag	UNP P30988
B	-336	MET	-	initiating methionine	UNP A0A0A8UN35
B	32	ASN	-	linker	UNP A0A0A8UN35
B	33	ALA	-	linker	UNP A0A0A8UN35
B	34	ALA	-	linker	UNP A0A0A8UN35
B	35	ALA	-	linker	UNP A0A0A8UN35
B	36	GLU	-	linker	UNP A0A0A8UN35
B	37	PHE	-	linker	UNP A0A0A8UN35
B	142	HIS	-	expression tag	UNP P30988
B	143	HIS	-	expression tag	UNP P30988
B	144	HIS	-	expression tag	UNP P30988
B	145	HIS	-	expression tag	UNP P30988
B	146	HIS	-	expression tag	UNP P30988

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	147	HIS	-	expression tag	UNP P30988

- Molecule 2 is a protein called Calcitonin.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	18	Total	C	N	O	0	0	1
			132	81	26	25			
2	D	18	Total	C	N	O	0	2	1
			148	90	31	27			

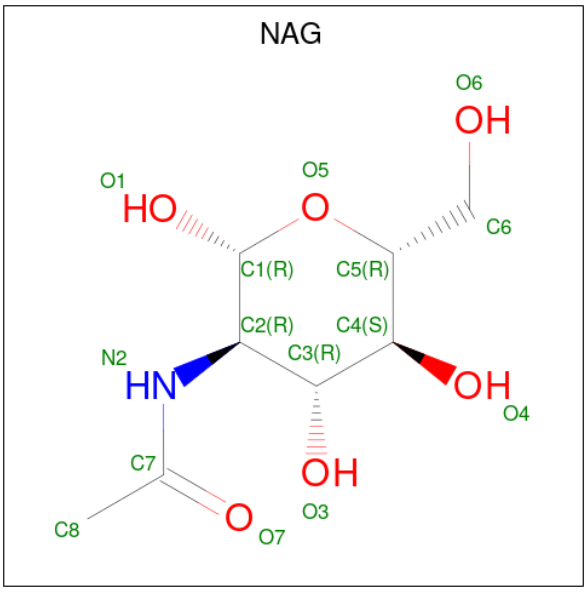
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	33	NH2	-	amidation	UNP Q92163
D	33	NH2	-	amidation	UNP Q92163

- Molecule 3 is an oligosaccharide called alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose.

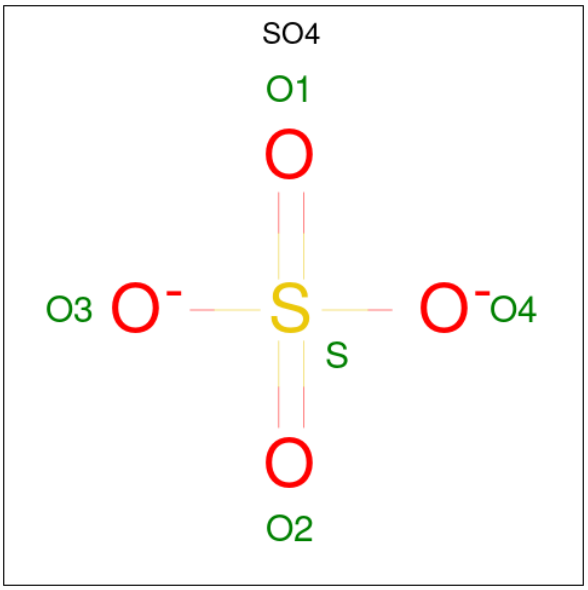
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	E	2	Total	C	O	0	0	0
			23	12	11			
3	F	2	Total	C	O	0	0	0
			23	12	11			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$) (labeled as "Ligand of Interest" by depositor).



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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total O S 5 4 1	0	0

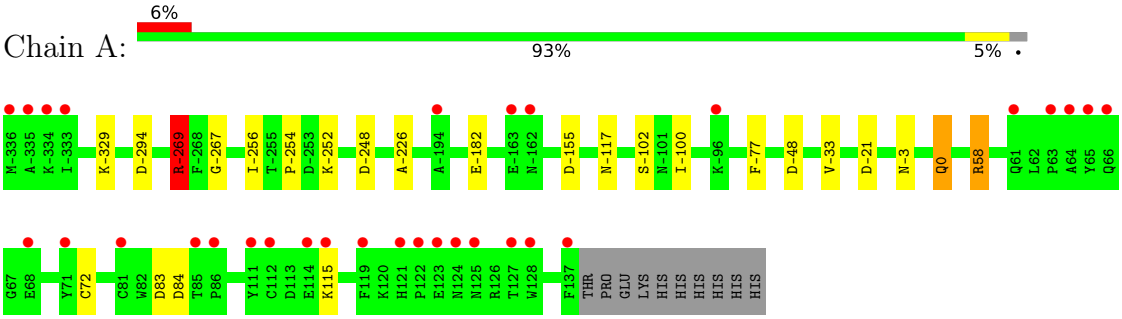
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	327	Total O 327 327	0	0
6	B	303	Total O 303 303	0	0
6	C	9	Total O 9 9	0	0
6	D	13	Total O 13 13	0	0

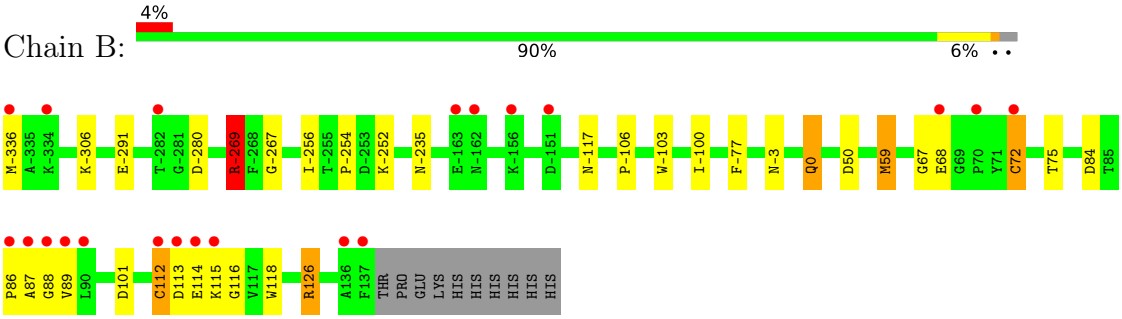
3 Residue-property plots (i)

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.

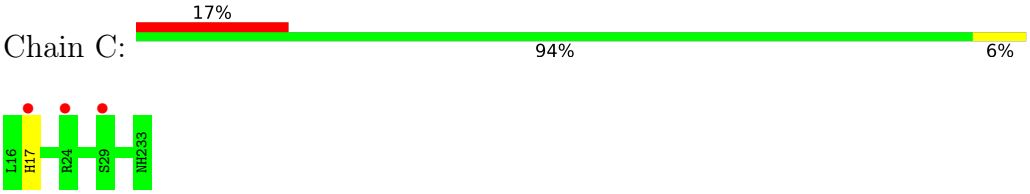
• Molecule 1: Maltodextrin-binding protein,Calcitonin receptor



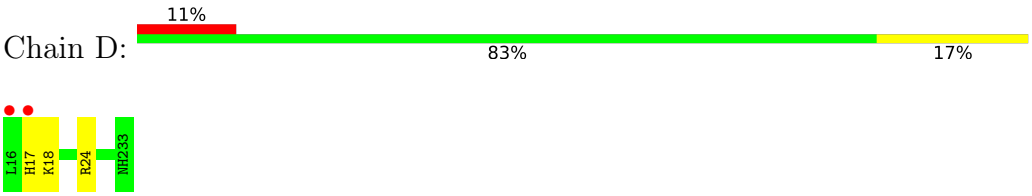
• Molecule 1: Maltodextrin-binding protein,Calcitonin receptor



• Molecule 2: Calcitonin



• Molecule 2: Calcitonin

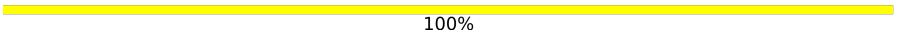


• Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain E:

• Molecule 3: alpha-D-glucopyranose-(1-4)-alpha-D-glucopyranose

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	121.03Å 121.03Å 263.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.86 – 1.78 49.86 – 1.78	Depositor EDS
% Data completeness (in resolution range)	99.1 (49.86-1.78) 99.1 (49.86-1.78)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.01 (at 1.78Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.186 , 0.214 0.195 , 0.221	Depositor DCC
R_{free} test set	5348 reflections (4.93%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.271	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	8740	wwPDB-VP
Average B, all atoms (Å^2)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.56% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: GLC, NH2, SO4, NAG

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	0.77	2/3900 (0.1%)	0.80	7/5294 (0.1%)
1	B	0.73	0/4048	0.84	6/5496 (0.1%)
2	C	0.72	0/134	0.91	0/182
2	D	0.72	0/153	0.86	0/207
All	All	0.75	2/8235 (0.0%)	0.82	13/11179 (0.1%)

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	B	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	-102	SER	CB-OG	-6.13	1.34	1.42
1	A	-182	GLU	CD-OE1	5.43	1.31	1.25

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	59	MET	CG-SD-CE	-10.18	83.91	100.20
1	A	-269[A]	ARG	NE-CZ-NH1	-7.73	116.43	120.30
1	A	-269[B]	ARG	NE-CZ-NH1	-7.73	116.43	120.30
1	B	-269[A]	ARG	NE-CZ-NH2	-7.71	116.45	120.30
1	B	-269[B]	ARG	NE-CZ-NH2	-7.71	116.45	120.30
1	B	50	ASP	CB-CG-OD1	5.59	123.33	118.30
1	B	-280	ASP	CB-CG-OD1	5.58	123.32	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	-48	ASP	CB-CG-OD1	5.55	123.30	118.30
1	B	126	ARG	NE-CZ-NH2	-5.49	117.56	120.30
1	A	-294	ASP	CB-CG-OD1	5.44	123.20	118.30
1	A	-21	ASP	CB-CG-OD1	5.15	122.93	118.30
1	A	-21	ASP	CB-CG-OD2	-5.03	113.77	118.30
1	A	-248	ASP	CB-CG-OD1	5.02	122.81	118.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	67	GLY	Mainchain
1	B	68[A]	GLU	Peptide
1	B	72[B]	CYS	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	3778	0	3683	10	1
1	B	3923	0	3809	28	1
2	C	132	0	130	0	0
2	D	148	0	148	3	0
3	E	23	0	20	1	0
3	F	23	0	20	0	0
4	A	28	0	26	0	0
4	B	28	0	26	0	0
5	B	5	0	0	0	0
6	A	327	0	0	1	0
6	B	303	0	0	1	0
6	C	9	0	0	0	0
6	D	13	0	0	0	0
All	All	8740	0	7862	39	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (39) close contacts within the same asymmetric unit are listed below, sorted by their clash

magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:87[A]:ALA:HA	1:B:112[A]:CYS:SG	1.86	1.14
1:B:112[B]:CYS:SG	1:B:113[B]:ASP:N	2.30	1.05
1:B:86[A]:PRO:O	1:B:112[A]:CYS:SG	2.16	1.03
1:B:87[A]:ALA:CA	1:B:112[A]:CYS:SG	2.53	0.97
1:B:88[A]:GLY:N	1:B:112[A]:CYS:SG	2.42	0.92
1:B:112[A]:CYS:SG	1:B:112[A]:CYS:O	2.35	0.84
1:B:113[B]:ASP:OD1	1:B:116:GLY:N	2.16	0.78
1:B:87[A]:ALA:C	1:B:112[A]:CYS:SG	2.66	0.74
1:B:72[A]:CYS:SG	1:B:112[A]:CYS:HB2	2.27	0.73
1:B:59:MET:HE2	1:B:75:THR:HA	1.77	0.65
1:B:-256:ILE:HG22	1:B:-254:PRO:HD3	1.79	0.64
1:A:-256:ILE:HD12	1:A:-254:PRO:HD3	1.86	0.57
1:B:59:MET:HE2	1:B:75:THR:CA	2.34	0.57
1:B:72[A]:CYS:O	1:B:84[A]:ASP:HA	2.06	0.56
1:A:0:GLN:H	1:A:0:GLN:NE2	2.06	0.54
1:A:-269[A]:ARG:NH2	6:A:602:HOH:O	2.40	0.53
1:B:0:GLN:NE2	1:B:0:GLN:H	2.07	0.51
1:B:0:GLN:H	1:B:0:GLN:HE21	1.59	0.51
1:B:-267:GLY:HA3	1:B:-3:ASN:O	2.12	0.49
1:B:59:MET:CE	1:B:75:THR:CA	2.90	0.49
1:B:-117:ASN:HD21	1:B:-100:ILE:HG12	1.78	0.49
1:A:72:CYS:O	1:A:84:ASP:HA	2.13	0.48
1:A:0:GLN:H	1:A:0:GLN:HE21	1.61	0.48
1:A:58:ARG:NH1	1:A:83:ASP:OD1	2.47	0.48
1:B:59:MET:HE1	1:B:75:THR:C	2.34	0.48
1:B:118:TRP:HB3	6:B:440:HOH:O	2.12	0.48
1:B:114[B]:GLU:OE1	1:B:114[B]:GLU:N	2.45	0.47
1:B:-235[A]:ASN:HD21	2:D:18:LYS:NZ	2.12	0.47
1:A:-267:GLY:HA3	1:A:-3:ASN:O	2.16	0.46
1:A:-117:ASN:HD21	1:A:-100:ILE:HG12	1.82	0.44
1:B:89[B]:VAL:O	1:B:112[B]:CYS:HB3	2.17	0.43
1:B:86[A]:PRO:C	1:B:112[A]:CYS:SG	2.92	0.43
1:B:-291:GLU:CG	1:B:-269[B]:ARG:HD3	2.49	0.43
2:D:24[B]:ARG:HD2	2:D:24[B]:ARG:HA	1.38	0.42
1:B:-106:PRO:HA	1:B:-103:TRP:CE2	2.55	0.42
1:A:-269[A]:ARG:NH2	3:E:2:GLC:O4	2.53	0.41
1:B:101:ASP:HB2	2:D:24[B]:ARG:NH1	2.36	0.41
1:B:59:MET:HE2	1:B:75:THR:N	2.37	0.40
1:A:-226:ALA:HA	1:A:-33:VAL:HA	2.02	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the sym-

metry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:-155:ASP:OD1	1:B:126:ARG:NH2[12_544]	2.19	0.01

5.3 Torsion angles

5.3.1 Protein backbone

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	478/484 (99%)	469 (98%)	9 (2%)	0	100	100
1	B	498/484 (103%)	491 (99%)	7 (1%)	0	100	100
2	C	16/18 (89%)	16 (100%)	0	0	100	100
2	D	18/18 (100%)	18 (100%)	0	0	100	100
All	All	1010/1004 (101%)	994 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	393/397 (99%)	385 (98%)	8 (2%)	55	40
1	B	410/397 (103%)	399 (97%)	11 (3%)	44	28
2	C	15/15 (100%)	14 (93%)	1 (7%)	16	4
2	D	17/15 (113%)	16 (94%)	1 (6%)	19	6
All	All	835/824 (101%)	814 (98%)	21 (2%)	53	31

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	-329	LYS
1	A	-269[A]	ARG
1	A	-269[B]	ARG
1	A	-252	LYS
1	A	-77	PHE
1	A	0	GLN
1	A	58	ARG
1	A	115	LYS
1	B	-336	MET
1	B	-306	LYS
1	B	-269[A]	ARG
1	B	-269[B]	ARG
1	B	-252	LYS
1	B	-77	PHE
1	B	0	GLN
1	B	112[A]	CYS
1	B	112[B]	CYS
1	B	115[A]	LYS
1	B	115[B]	LYS
2	C	17	HIS
2	D	17	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	-117	ASN
1	A	0	GLN
1	B	-134	ASN
1	B	-117	ASN
1	B	0	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

4 monosaccharides are modelled in this entry.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	GLC	E	1	3	12,12,12	1.16	1 (8%)	17,17,17	1.69	5 (29%)
3	GLC	E	2	3	11,11,12	0.85	0	15,15,17	1.40	2 (13%)
3	GLC	F	1	3	12,12,12	0.81	0	17,17,17	1.81	4 (23%)
3	GLC	F	2	3	11,11,12	0.92	0	15,15,17	1.60	2 (13%)

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

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	GLC	E	1	3	-	0/2/22/22	0/1/1/1
3	GLC	E	2	3	-	0/2/19/22	0/1/1/1
3	GLC	F	1	3	-	0/2/22/22	0/1/1/1
3	GLC	F	2	3	-	0/2/19/22	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	GLC	C4-C3	2.24	1.58	1.52

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	2	GLC	C1-O5-C5	5.03	119.00	112.19
3	F	1	GLC	C1-O5-C5	4.47	122.10	113.66
3	F	1	GLC	O5-C1-C2	3.51	116.55	110.28
3	E	1	GLC	C1-O5-C5	3.32	119.92	113.66
3	E	2	GLC	C1-O5-C5	3.18	116.50	112.19
3	E	1	GLC	O5-C1-C2	2.98	115.61	110.28

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	E	2	GLC	O5-C5-C6	-2.97	102.55	107.20
3	F	1	GLC	O1-C1-O5	-2.86	101.81	110.38
3	F	2	GLC	O2-C2-C3	2.52	115.18	110.14
3	E	1	GLC	O1-C1-O5	-2.33	103.39	110.38
3	E	1	GLC	C1-C2-C3	2.33	115.14	110.31
3	E	1	GLC	O2-C2-C3	-2.28	105.08	110.35
3	F	1	GLC	O2-C2-C3	-2.15	105.39	110.35

There are no chirality outliers.

There are no torsion outliers.

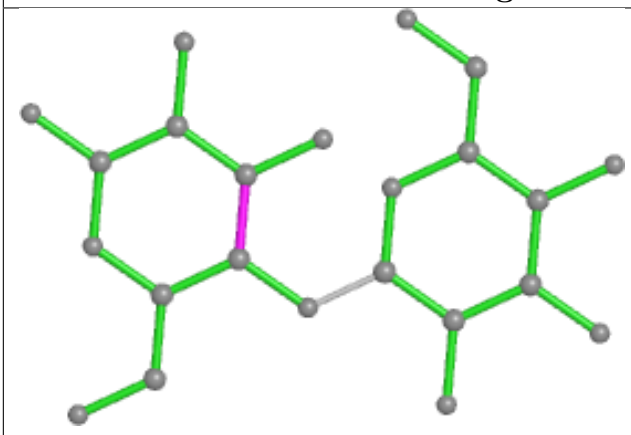
There are no ring outliers.

1 monomer is involved in 1 short contact:

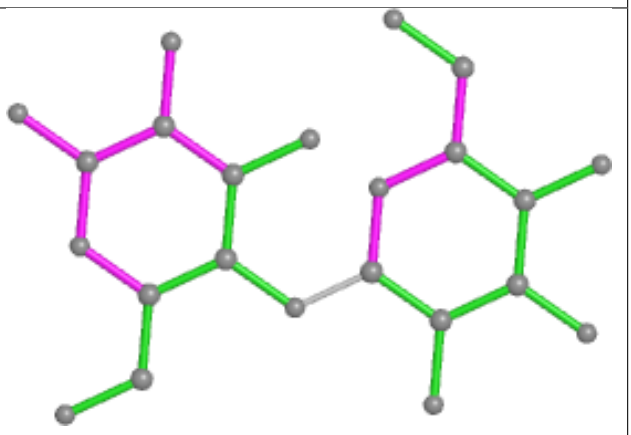
Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	E	2	GLC	1	0

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

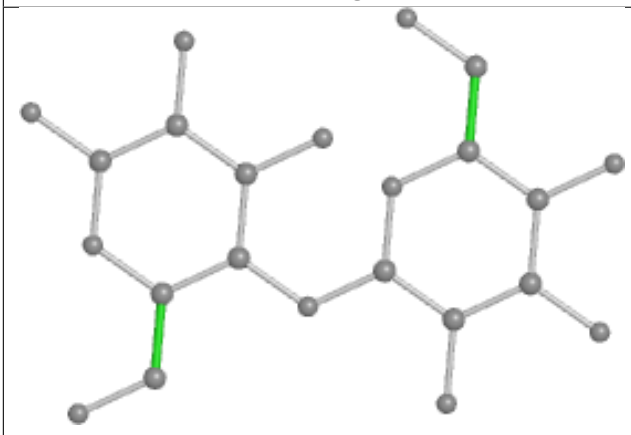
Oligosaccharide Chain E



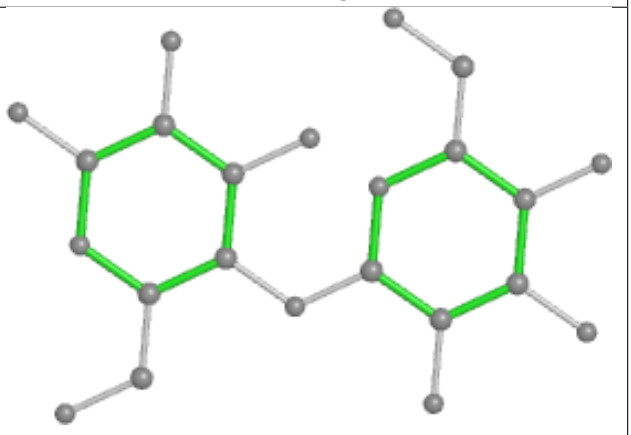
Bond lengths



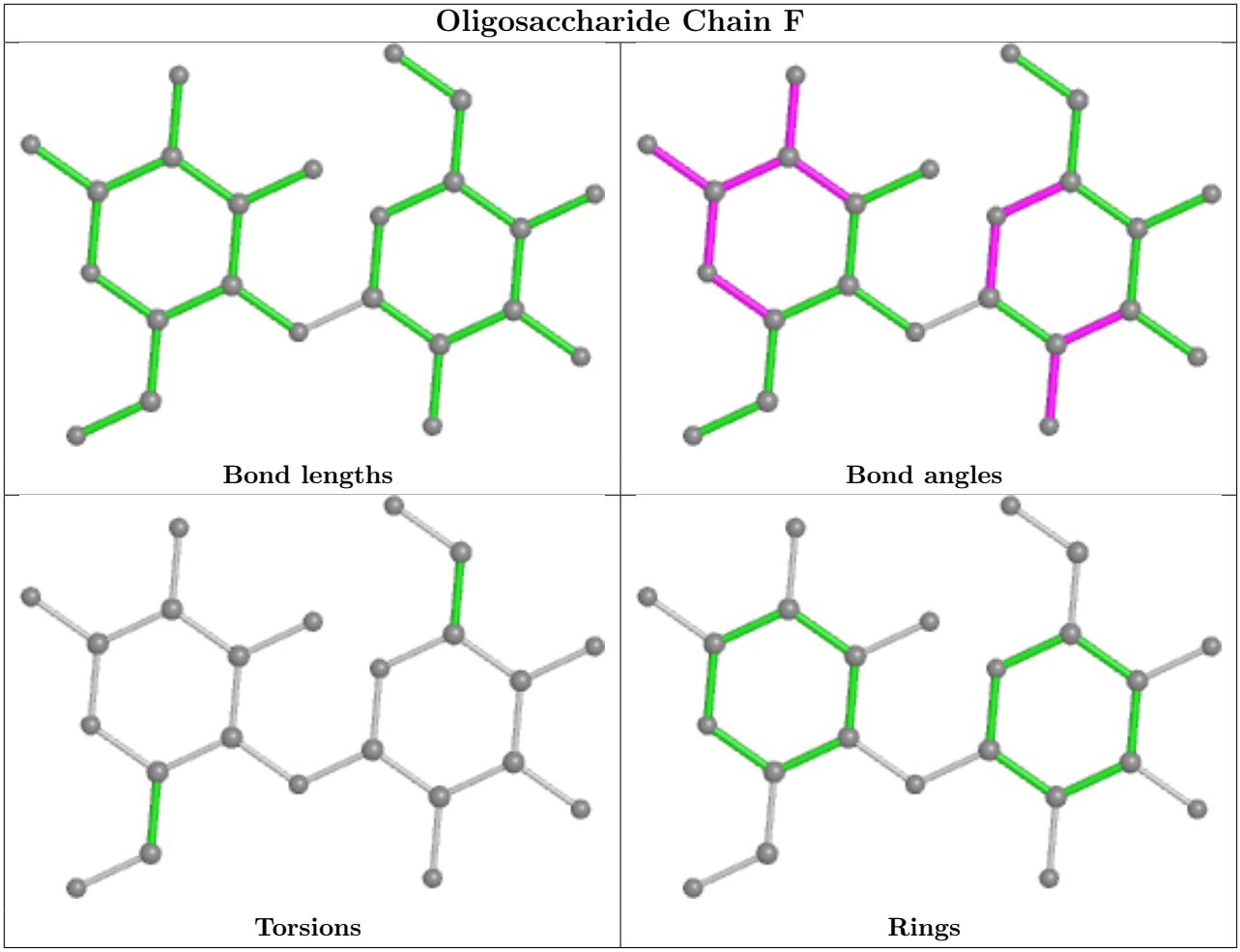
Bond angles



Torsions



Rings



5.6 Ligand geometry (i)

5 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	$\#\|Z\| > 2$
4	NAG	A	501	1	14,14,15	0.38	0	17,19,21	0.97	1 (5%)
4	NAG	A	502	1	14,14,15	0.44	0	17,19,21	0.58	0
4	NAG	B	203	1	14,14,15	0.43	0	17,19,21	1.09	1 (5%)
4	NAG	B	204	1	14,14,15	0.37	0	17,19,21	0.69	0
5	SO4	B	201	-	4,4,4	0.54	0	6,6,6	0.71	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
4	NAG	A	501	1	-	2/6/23/26	0/1/1/1
4	NAG	B	204	1	-	0/6/23/26	0/1/1/1
4	NAG	A	502	1	-	0/6/23/26	0/1/1/1
4	NAG	B	203	1	-	0/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	501	NAG	C2-N2-C7	3.05	127.25	122.90
4	B	203	NAG	C2-N2-C7	2.69	126.73	122.90

There are no chirality outliers.

All (2) torsion outliers are listed below:

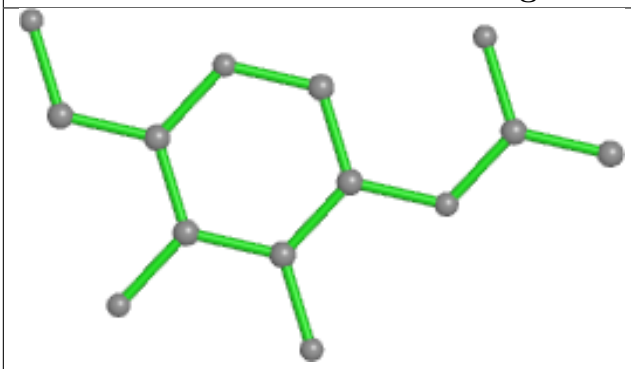
Mol	Chain	Res	Type	Atoms
4	A	501	NAG	O5-C5-C6-O6
4	A	501	NAG	C4-C5-C6-O6

There are no ring outliers.

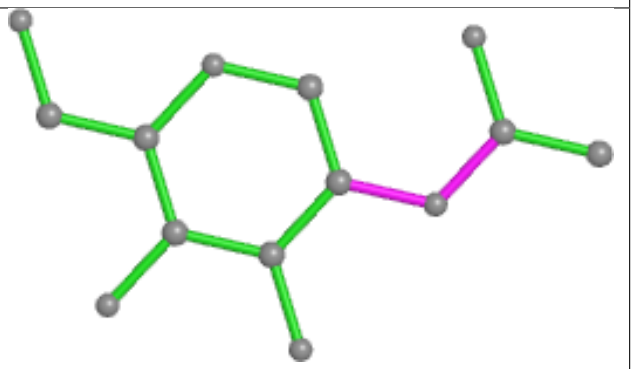
No monomer is involved in short contacts.

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

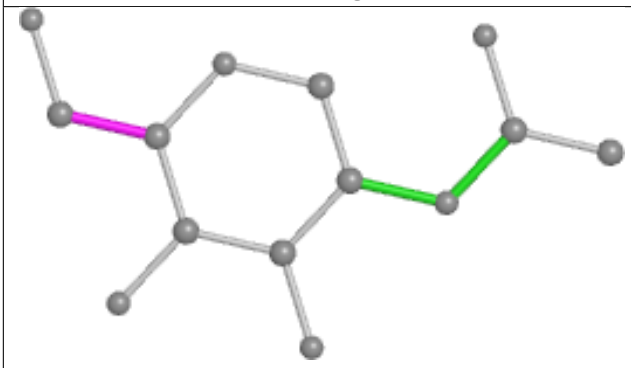
Ligand NAG A 501



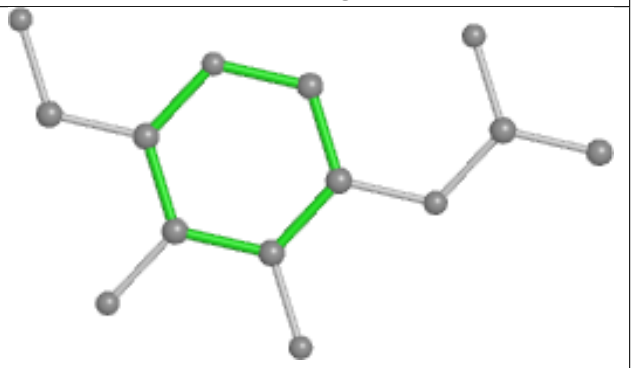
Bond lengths



Bond angles

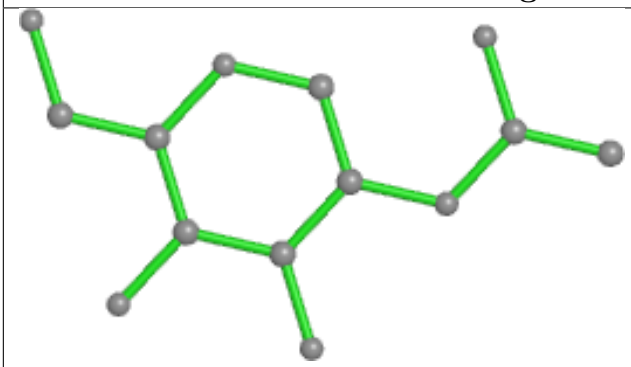


Torsions

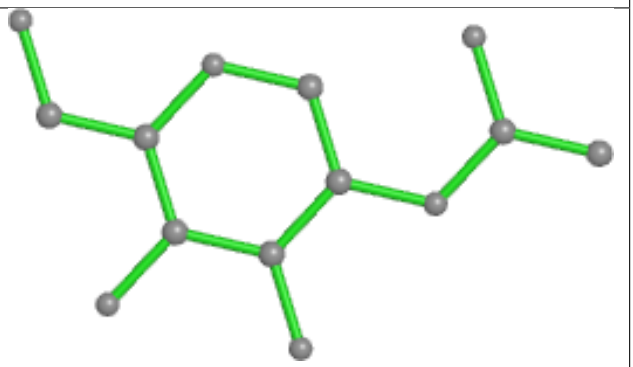


Rings

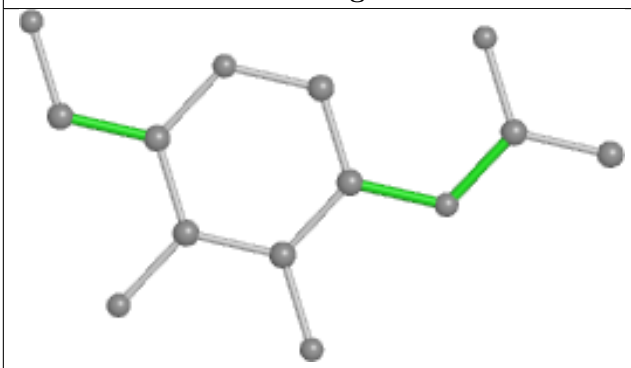
Ligand NAG A 502



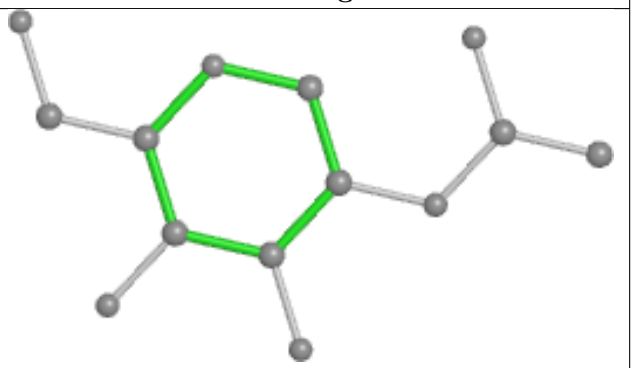
Bond lengths



Bond angles

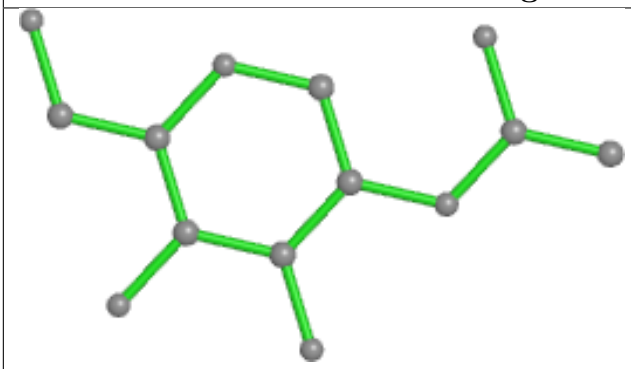


Torsions

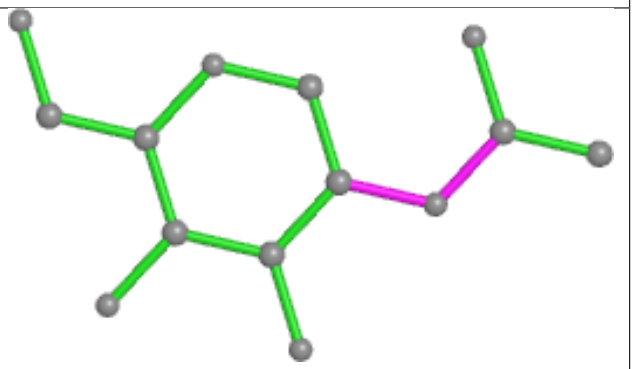


Rings

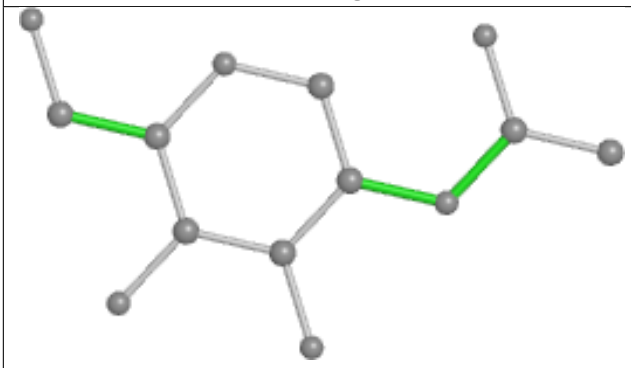
Ligand NAG B 203



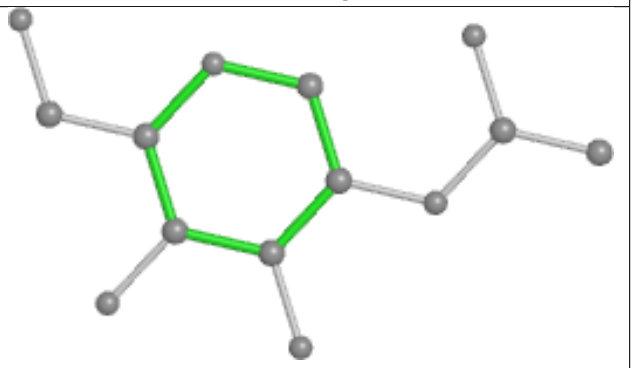
Bond lengths



Bond angles

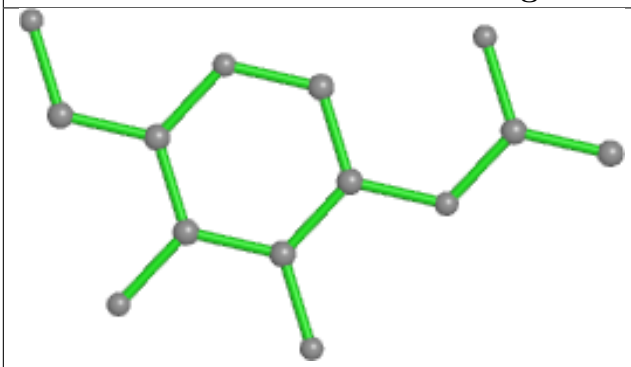


Torsions

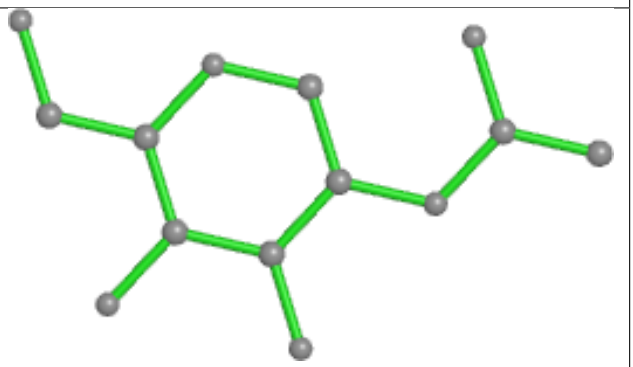


Rings

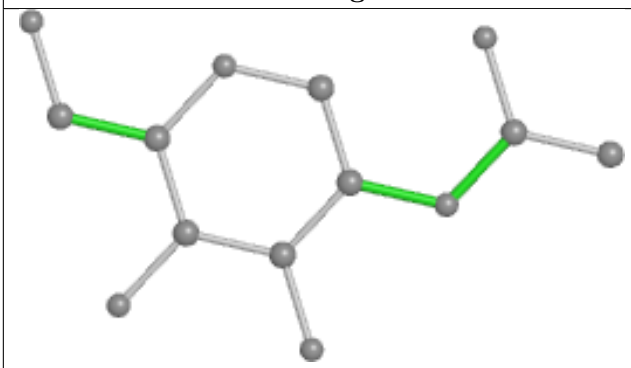
Ligand NAG B 204



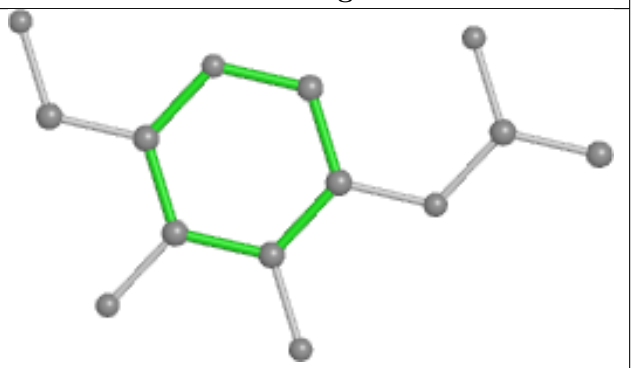
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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^2)	Q<0.9
1	A	474/484 (97%)	0.14	31 (6%)	18	18	15, 27, 63, 86	0
1	B	474/484 (97%)	0.18	21 (4%)	34	32	18, 31, 50, 89	0
2	C	17/18 (94%)	0.60	3 (17%)	1	1	24, 35, 49, 52	0
2	D	17/18 (94%)	0.47	2 (11%)	4	4	30, 35, 61, 62	0
All	All	982/1004 (97%)	0.18	57 (5%)	23	22	15, 30, 57, 89	0

All (57) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	137	PHE	8.5
1	B	137	PHE	7.2
1	A	125	ASN	6.1
1	B	115[A]	LYS	5.8
1	B	-162	ASN	5.4
1	B	112[A]	CYS	5.4
1	A	64	ALA	5.1
1	B	-336	MET	5.0
2	D	16	LEU	4.6
1	A	61	GLN	4.6
1	A	122	PRO	4.4
1	B	89[A]	VAL	4.4
1	A	-334	LYS	4.3
1	A	-333	ILE	4.0
1	A	63	PRO	3.9
1	A	68	GLU	3.7
1	A	-335	ALA	3.7
1	B	88[A]	GLY	3.6
1	B	114[A]	GLU	3.6
1	B	86[A]	PRO	3.5
1	B	72[A]	CYS	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	-162	ASN	3.5
1	A	85	THR	3.5
2	C	29	SER	3.4
1	A	-194	ALA	3.2
1	A	111	TYR	3.2
1	A	-336	MET	3.1
1	A	65	TYR	3.1
1	A	124	ASN	3.1
1	B	-163	GLU	3.0
1	B	87[A]	ALA	3.0
1	A	71	TYR	2.9
1	A	121	HIS	2.8
2	D	17	HIS	2.8
1	B	136	ALA	2.8
1	A	-163	GLU	2.7
1	B	70[A]	PRO	2.6
1	B	68[A]	GLU	2.6
1	A	115	LYS	2.5
1	A	119	PHE	2.4
1	B	-334	LYS	2.4
1	B	-156	LYS	2.4
2	C	24	ARG	2.4
1	A	81	CYS	2.4
1	A	127	THR	2.4
1	A	128	TRP	2.4
1	A	114	GLU	2.3
1	B	90[A]	LEU	2.3
1	B	-282	THR	2.2
2	C	17	HIS	2.2
1	A	66	GLN	2.1
1	A	-96	LYS	2.1
1	A	86	PRO	2.1
1	B	-151	ASP	2.1
1	B	113[A]	ASP	2.1
1	A	112	CYS	2.0
1	A	123	GLU	2.0

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)

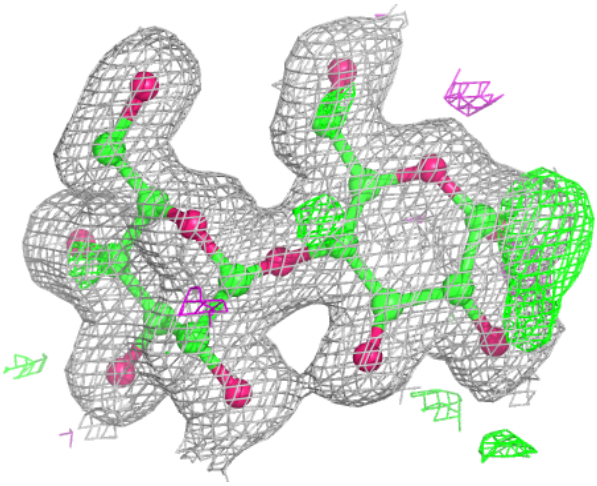
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

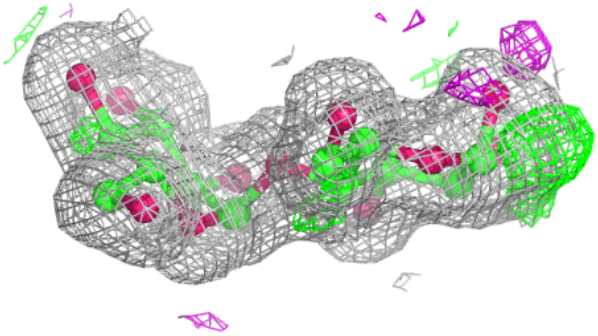
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	GLC	E	1	12/12	0.96	0.09	15,17,21,25	0
3	GLC	F	1	12/12	0.97	0.10	21,23,25,25	0
3	GLC	E	2	11/12	0.98	0.11	15,15,16,16	0
3	GLC	F	2	11/12	0.98	0.08	20,21,21,22	0

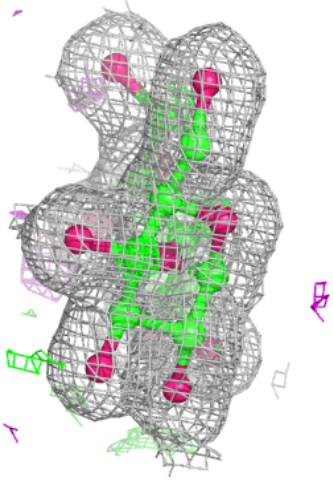
The following is a graphical depiction of the model fit to experimental electron density for oligosaccharide. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around Chain E:

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

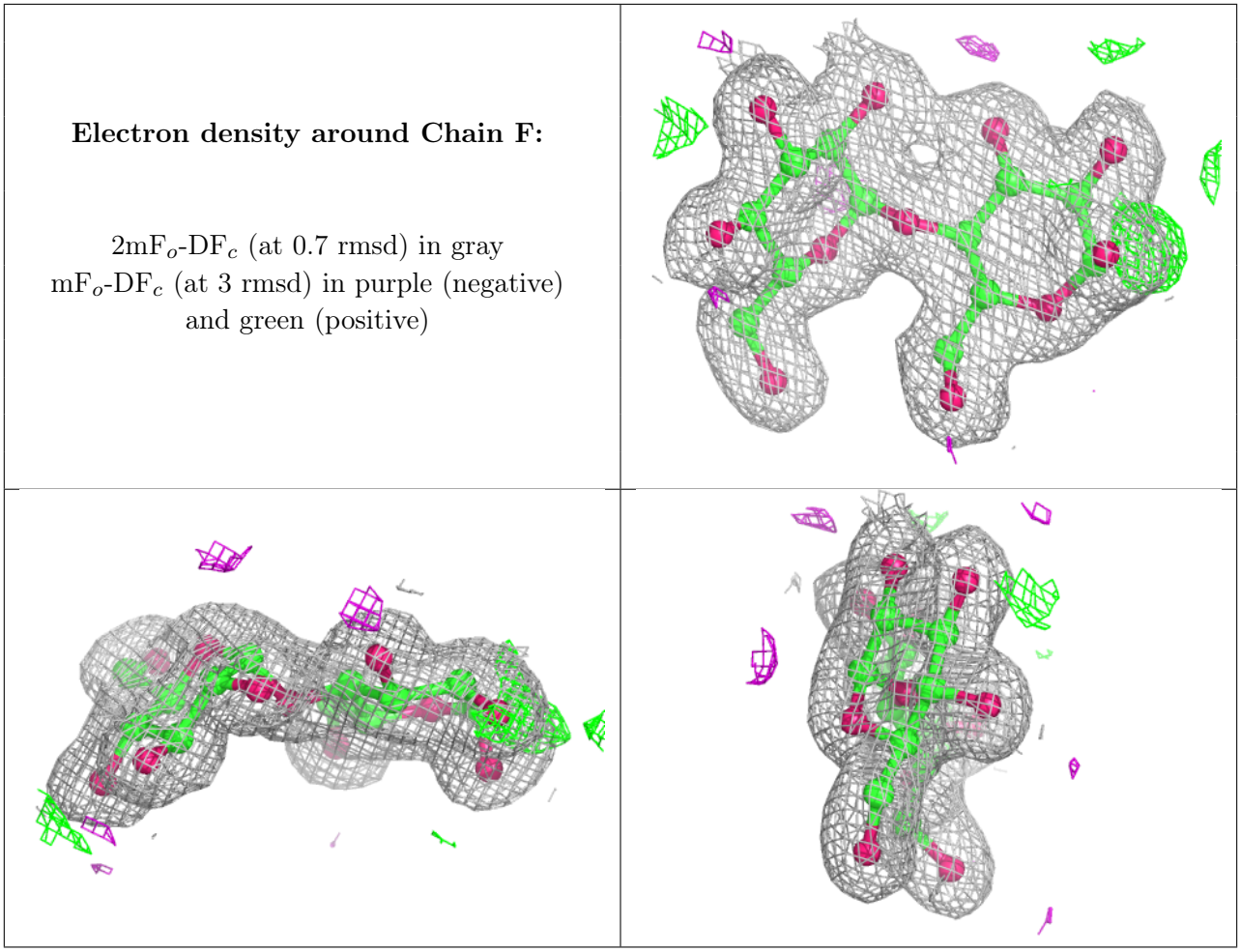






Electron density around Chain F:

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



6.4 Ligands

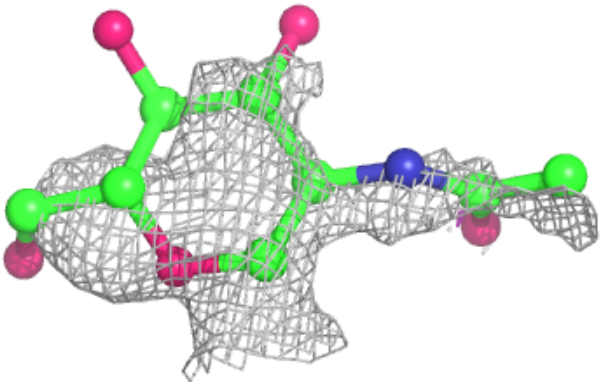
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

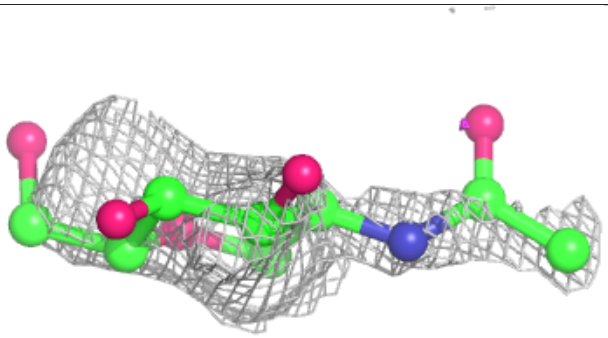
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	NAG	A	501	14/15	0.77	0.44	84,92,95,97	0
4	NAG	B	203	14/15	0.85	0.17	51,60,63,65	0
4	NAG	A	502	14/15	0.90	0.22	46,50,54,55	0
4	NAG	B	204	14/15	0.90	0.16	38,42,44,45	0
5	SO4	B	201	5/5	0.95	0.10	33,38,42,46	0

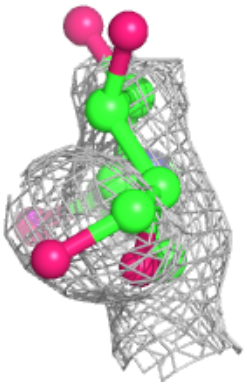
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around NAG A 501:

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

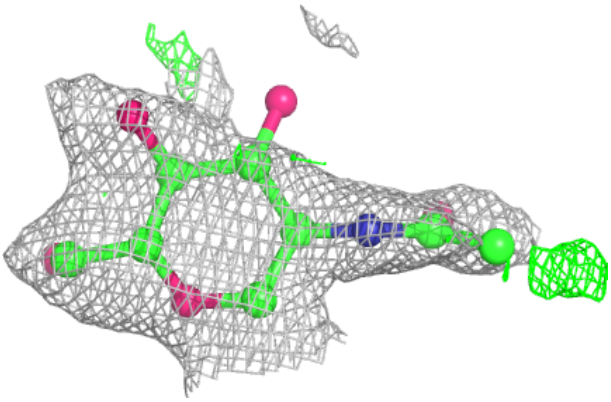


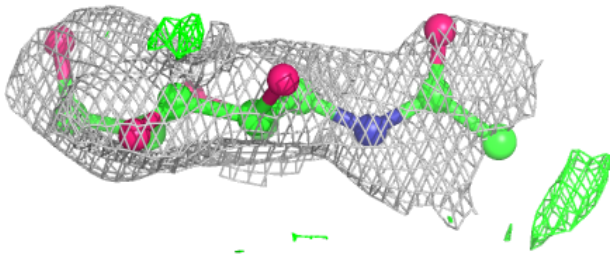


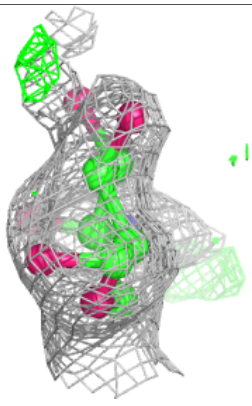


Electron density around NAG B 203:

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

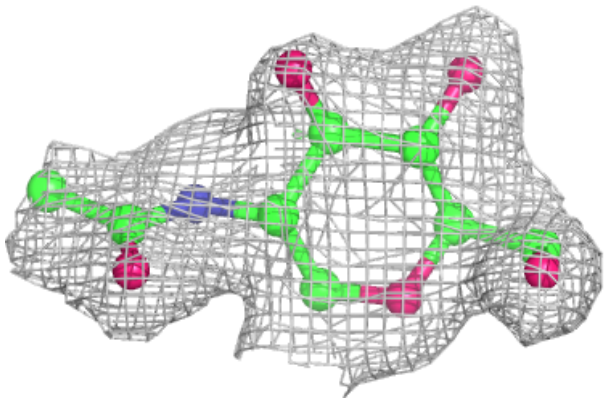


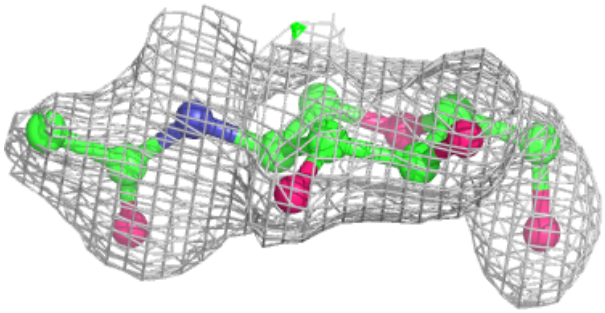


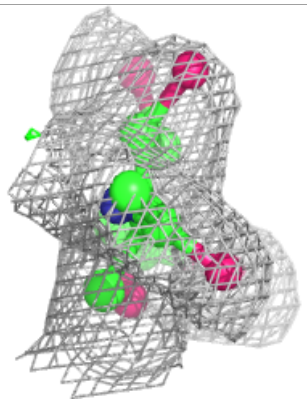


Electron density around NAG A 502:

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

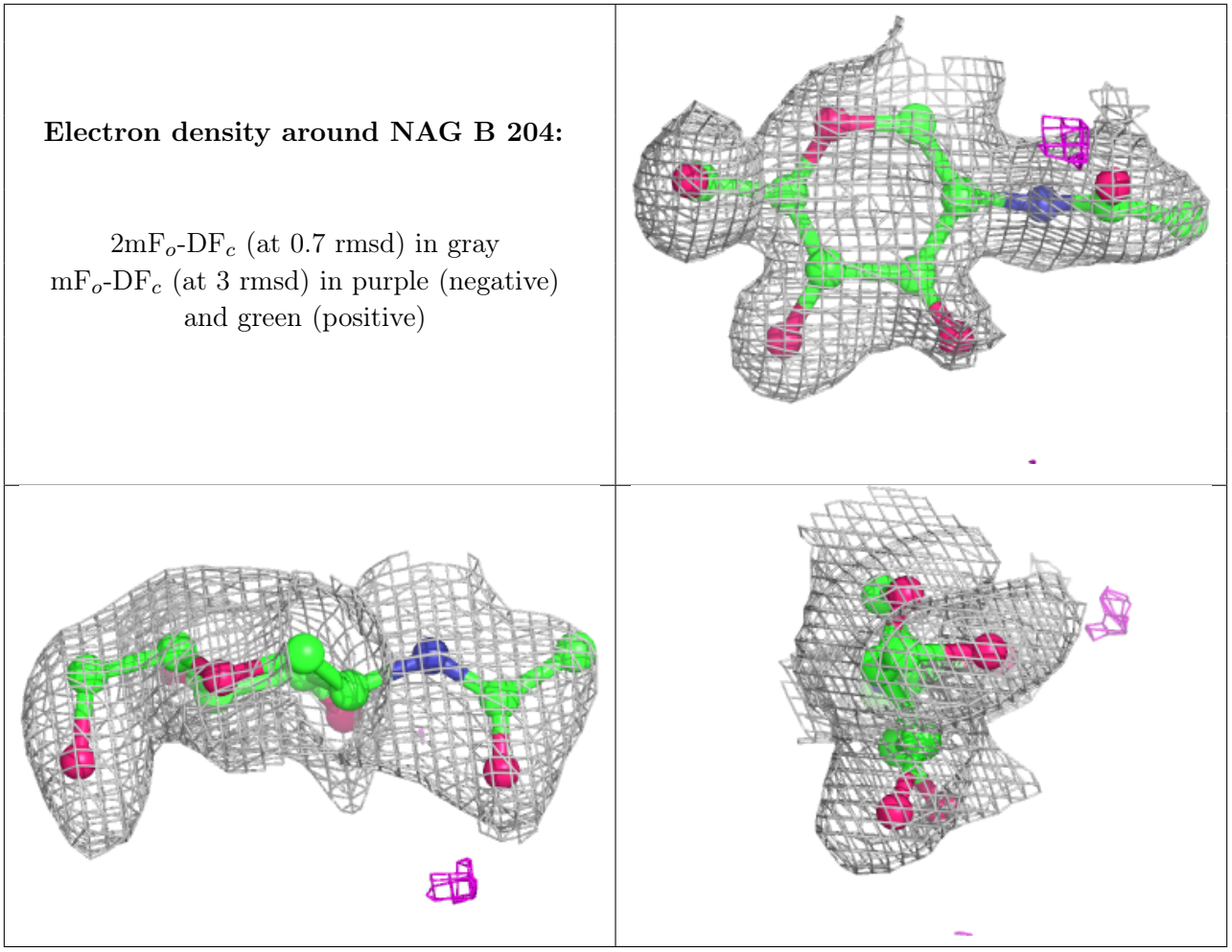






Electron density around NAG B 204:

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



6.5 Other polymers

There are no such residues in this entry.